Acta Oncologica Vol. 31, No. 3, pp. 311-321, 1992

PSYCHOMETRIC VALIDATION OF THE EORTC CORE QUALITY OF LIFE QUESTIONNAIRE, 30-ITEM VERSION AND A DIAGNOSIS-SPECIFIC MODULE FOR HEAD AND NECK CANCER PATIENTS

KRISTIN BJORDAL and STEIN KAASA

A cancer-specific self-reporting quality of life questionnaire has been validated. The questionnaire is designed to assess physical functioning, role functioning, cognitive functioning, emotional functioning, social functioning, pain, fatigue, emesis and quality of life by means of multi-item scales, and other disease- and treatment-related symptoms by means of single items. The questionnaire was completed by 126 head and neck cancer patients with a mean age of 67 years. The internal consistency (scale reliability) was satisfactory for all scales but one. Correlations between scales and items assessing the same underlying dimension were also satisfactory. The questionnaire discriminates between patient subgroups and between acute, subacute and late toxicity. Patient compliance was high. The questionnaire provided valuable information, and most of the scales/items functioned well. A few problems were found, especially with the modified visual analogue scales, and minor modifications will be made.

Combined treatment consisting of high-voltage radiotherapy and surgery has increased the cure rate for patients with early stage head and neck cancer. Patients with large primary tumours and nodal disease have lower change of cure (1). The incidence of head and neck cancer in Norway was 13 per 100 000 people in 1989 (2). Seventy per cent of cases were located in the oral cavity or pharynx, 20% in the larynx, and 9% in the nasal cavity and paranasal sinuses. The average five-year survival for all locations together was 60% ranging from 80% in stage I disease to 10% in stage IV. Patients with advanced disease often had local relapses. Head and neck cancer is associated with physical morbidity in the face, mouth, and neck. The morbidity may be caused by the disease or by the treatment. Acute toxicity during radiotherapy is well documented (3). However, the prevalence of late toxicity and its influence on patients quality of life (QOL) has not been systematically investigated. Information about the QOL of head and neck cancer patients after curative treatment is useful since it may improve their rehabilitation.

Paper presented at a symposium in Copenhagen, October 1990, arranged by the Nordic Cancer Union (NCU).
Submitted 20 December 1991.
Accepted 19 January 1992.
Correspondence to: Dr Kristin Bjordal, Dept of Oncology, The Norwegian Radium Hospital, Montebello, N-0310 Oslo, Norway.
Address: See above for both authors.

Defining quality of life

In the literature of oncology the term 'quality of life' has been used in different ways, both as a concept and as an instrument of measurement (4, 5). It is rarely defined explicitly (6). QOL may be regarded as a positive term, convening a person's subjective impression of his life as a whole (7). It includes an assessment of general health, satisfaction, fulfilment, ability to cope, happiness, being in control and degree of independence (8, 9). Assessment of QOL may be regarded as an assessment of cognitive judgements and affective reactions (10).

In medical research, QOL is usually approached as a multidimensional construct where the dimensions are broadly related to health (11, 12). It may be linked with the World Health Organization's (WHO) definition of health as 'a state of complete physical, mental and social well-being, and not merely the absence of disease and infirmity' (13).

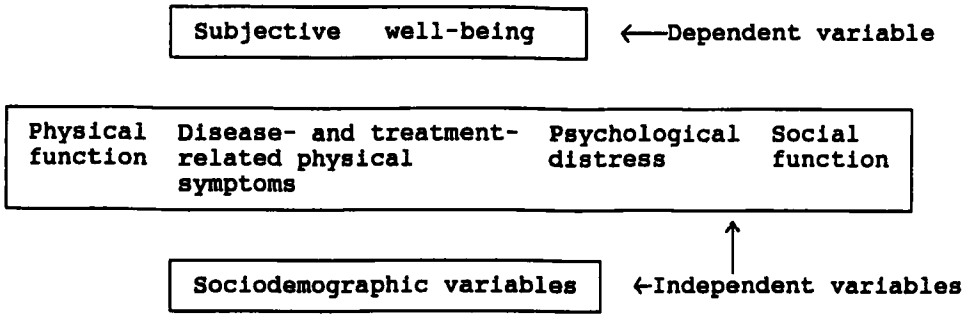

Figure. Independent variables describing subjective well-being.

In the social sciences, subjective well-being (also referred to as QOL) is usually measured in terms of both positive and negative affects (14). In some analyses it is treated as a dependent variable influenced by dimensions such as social network, prevalence of disease, and economy. A similar approach may be used in clinical oncology by measuring subjective well-being as a separate dimension (Figure). In such analyses, disease- and treatment-related symptoms, physical function, social function, psychological distress, and sociodemographic factors may be treated as independent variables of QOL (15).

In the present study, however, QOL is used as a general multidimensional term.

Measuring quality of life

The dimensions of the QOL construct are abstract and subjective and are therefore not directly observable or measurable (7). Indirect methods of measurement or indicators must be used. QOL can be assessed by observation, interviews, or questionnaires.

Assessments of patients' physical functions by a clinician have often been used as a measure of QOL (Karnofsky's or ECOG/WHO performance status) (16, 17). Spitzer's Quality of Life Index (18) has also been designed to be completed by physicians. This index consists of five categories of statements measuring activity, daily living, health, support, and outlook. The answers are scored on a three-point scale and added to an unweighted sum. Physicians' ratings require little time and effort compared with other types of QOL assessments. However, in several trials, the physicians' compliance has been shown to be poor and a low agreement has been found between ratings by physicians and ratings by patients (19).

Generic self-report questionnaires have been designed for use in a wide range of chronic disease populations. The General Health Questionnaire (GHQ) (20), the Sickness Impact Profile (SIP) (21), Nottingham Health Profile (NHP) (22), and McMaster Health Index Questionnaire (MHIQ) (23) are all examples of ways of measuring general health applied to QOL.

These instruments are not specifically relevant to cancer patients, and do not allow for the assessment of factors such as disease- and treatment-related symptoms. Their psychometric properties are well documented, but their validity and reliability, when applied to cancer populations, have not been investigated in detail.

Specific cancer-related instruments have been designed for use in a wide range of diagnostic groups of cancer patients. The European Organization for Research and Treatment of Cancer Core Quality of Life Questionnaire, 36-item version (EORTC QLQ-C36) (11) has been recently developed and validated. It is suitable for repeated testing. This first version has been revised and a 30-item version (EORTC QLQ-C30) is currently validated in several clinical trials in Europe and the USA. The short form of the Cancer Rehabilitation Evaluation System (CARES), (24, 25) has also been subjected to clinical trials. These two instruments use a multidimensional approach and employ multi-item scales without making any total score.

The Rotterdam Symptom Checklisk (RSCL) (26) and a visual analogue self-assessment (VAS) scale by Coates et al. (27) focus on physical symptoms and physchological problems. They do not include functional or social health. The RSCL is scored on four-point Likert scales (eg. 'not at all', 'a little'), the other on VAS scales (i.e. 10 cm line segments with descriptive anchors at each end).

The Function of Living Index Cancer (FLIC) (28) and the Linear Analogue Self Assessment (LASA) (29) do not have subscales, and the scores are added without any weighting of the different dimensions. Both instruments are scored on VAS scales. Special disease-specific instruments have been designed e.g. for breast cancer patients (30, 31).

None of the QOL questionnaires for cancer patients are in general use. However, a basic set of criteria have been established (11, 32). The instrument should be cancer-specific, and the patients should be the primary source of data collection. The instrument should be brief and easy to understand and respond to. The approach should be multidimensional, and must include subscales assessing physical, functional, psychological and social health. The time frame should be specific and the items should have categorical

responses rather than VAS scales (33). Finally, the instrument should have acceptable validity and reliability, be sensitive to changes over time, and be able to differentiate between patient subgroups (34).

Reliability and validity

Reliability is a measure of random error. It can be assessed by repeated applications of the same test (test-retest reliability), or indirectly by analysing the internal consistency (scale reliability), which is a measure of the relationships between a set of indicators for the same underlying variable (i.e. multiple indicators for the construct happiness). It may be difficult to measure the test-retest reliability for a set of variables that fluctuates. Therefore, measures of internal consistency are often used to determine the reliability of QOL instruments.

Validity can be defined as the extent to which the result of a method measures the underlying construct in question. It is connected to the degree of systematic measurement error or bias. There are three basic types of validity: criterion, content, and construct validity (35). In laboratory medicine a criterion or a standard is often used to calibrate a new instrument. Since there is no golden standard for measuring cancer patients' QOL, indirect approaches need to be used, such as interview-based ratings, and/or analysis of the psychometric properties of the given instrument (how particular instruments relates to each other and their consistency with prior expectations or theory) (construct validity) (36).

Aim of the study

The aim of the present study was to evaluate the psychometric properties of the EORTC QLQ-C30, and of a diagnosis-specific module for head and neck cancer patients which has been developed at the Norwegian Radium Hospital (NRH). The clinical properties of the instrument were assessed and compared with other sources of information such as the General Health Questionnaire (GHQ) (20) and questions about subjective well-being previously used in the North Trøndelag study in Norway (37).

Material and Methods

Study sample

The QOL questionnaire was administered consecutively to hospitalized patients and out-patients at the head and neck ward at NRH from May to July 1990, or mailed to patients treated for head and neck cancer 1/2 to 1 year previously. There were no age or diagnosis limitations in patient selection. The patients were divided into five groups according to treatment:

1. patients prior to treatment
2. patients early in treatment (less than half-way through the radiotherapy course)
3. patients late in treatment (more than half-way through the radiotherapy course)
4. patients visiting the outpatient clinic after completed treatment
5. patients at home 1/2 to 1 year after treament

Instruments used in the study

The self-administered QOL questionnaire consisted of 73 items and a standardized form for sociodemographic data (education, profession, marital status, and size of household). The QOL questionnaire had four main parts:

1. EORTC QLQ-C30
2. a diagnosis-specific module for head and neck cancer
3. general satisfaction and physical health questions from a large population survey in Norway
4. General Health Questionnaire, 20-item version (GHQ-20)

Part 1 was the second generation of the EORTC Core QOL Questionnaire (Table 1). The first generation of the Core Questionnaire, EORTC QLQ-C36 exhibited satisfactory to excellent psychometric properties (38). The EORTC QLQ-C30 consists of six functioning scales: physical functioning (5 items), role functioning (2 items), cognitive functioning (2 items), emotional functioning (4 items), social functioning (2 items) and quality of life (2 items), and three symptom scales: pain (2 items), fatigue (3 items) and emesis (2 items). Single items concerning appetite, constipation, diarrhoea, dyspnoe, sleeping disturbances and economic consequences of the disease were also included. The items are scored on Likert scales. The higher the score on the functioning scales the higher the level of functioning, thus the scoring on the scales have been reversed (with 4 being the best score). A high score on a symptom scale indicates a high degree of symptoms. The time frame was one week.

Part 2 was a diagnosis-specific module for head and neck cancer patients which consisted of 21 items measuring disease-related symptoms and treatment-related side-effects (Table 1). The design of the module was based on the EORTC lung cancer module, literature, clinical experience, and interviews with patients. Most items had previously been used in a study on head and neck cancer patients at NRH (39).

Part 3 included two items taken from a population survey, the North Trøndelag Study (37). These items measure satisfaction with life in general and strength and fitness (Table 2).

Part 4 consisted of GHQ-20. It is designed to measure non-psychotic psychiatric illness in a general population,

Table 1

Scores for scales and single items in EORTC QLQ-C30 and the diagnosis-specific module

	No. of items	Score	
Functioning scales			
Physical functioning	5	No/yes	1–0
Role function	2	No/yes	1–0
Cognitive functioning	2	Likert	4–1§
Emotional functioning	4	Likert	4–1§
Social functioning	2	Likert	4–1§
Quality of life	2	Modified VAS	7–1†
Symptom scales			
Pain	2	Likert	1–4*
Fatigue	3	Likert	1–4*
Emesis	2	Likert	1–4*
Single Items			
Diarrhoea/Constipation	2	Likert	1–4*
Dyspnoe	1	Likert	1–4*
Sleeping disturbancies	1	Likert	1–4*
Appetite	1	Likert	1–4*
Economic consequences	1	Likert	1–4*
Diagnosis-specific module	21	Likert	1–4*

§ = 4: 'not at all', 3: 'a little', 2: 'quite a bit', 1: 'very much'
* = 1: 'not at all', 2: 'a little', 3: 'quite a bit', 4: 'very much'
† = Modified visual analogue scale (VAS), anchored by 1 ('very poor') at one end and 7 ('excellent') at the other

but has also been used to measure subjective well-being in populations with somatic illness (40). GHQ-20 focuses on the patient's normal function and measures changes over time (Table 2).

The construct validity of the questionnaire was assessed in three ways. An item/scale correlation matrix was made to test the validity in discrimination between items. Each item in the EORTC QLQ-C30 was correlated to each of the scales, including its own (corrected for overlap). Each item was supposed to correlate highest with its own scale as a support for the distinctiveness of the QOL dimentions assessed by the scale. A definite scaling error occurred when an item correlated more highly with another scale than with its own. In order to correct for random error, the correlation was to exceed two standard errors of the correlation. A probable scaling error occurred when the correlation was within two standard errors (38).

The scales and the two items in the QOL-scale in the EORTC QLQ-C30 were correlated with the total GHQ-20 score and the two questions from the population survey. It was hypothesized that scales and/or items measuring the same underlying dimension should correlate highly with one another, whereas scales and/or items with less in common should have low correlations.

Furthermore, the disease and treatment items were tested with respect to their ability to discriminate between groups of patients before, during, and after treatment (groups 1–5). It was expected that some items would refer to acute disease- and treatment-related symptoms and toxicity, while others would represent subacute and late disease- and treatment-related symptoms and toxicity.

Table 2

Scoring of items from the North Trøndelag Study and GHQ-20

Module	No. of items	Score	
The North Trøndelag Study	2	Likert	1–7*
GHQ-20	20	Likert	0–3†

* = 1: 'very satisfied', 2: 'satisfied', 3: 'fairly satisfied', 4: 'mixed', 5: 'rather dissatisfied', 6: 'dissatisfied', 7: 'very dissatisfied'
† = 0: 'not at all', 1: 'no more than usual', 2: 'rather more than usual', 3: 'much more than usual'

Statistical procedures

The mean scores were calculated for all items and scales. The internal consistency (scale reliability) was measured with Chronbach's alpha coefficient (41). Pearson's correlation coefficient was used in the correlation matrices. The questionnaire's discrimination ability was assessed by one-way analyses of variance. The SPSS/PC+, version 4.0 (42), was used in the majority of the statistical analyses. For the multitrait analyses the Revised Multitrait Analysis Program Software (MAP) were used.

Results

Patients' characteristics

One hundred and forty-six patients were asked to participate in the study. Ten patients refused, two patients did not participate due to psychiatric disorders, seven did not return the questionnaire and one questionnaire was not completed. One hundred and twenty-six patients were thus eligible for analysis. The mean age was 67 years (Table 3). Seventy-five per cent of the patients were over 60 years of age and 77% were men. Only 11% of the patients had had a university education. Sixty-seven per cent were married, 24% lived alone. Thirty-three per cent of the patients had cancer in the oral cavity, 13% in the pharynx, 19% in the larynx, 18% of the skin and the remaining 18% in salivary glands, paranasal sinuses, the thyroid gland or in cervical lymph nodes with unknown primary tumour. The analysis showed no major differences between the five subgroups on sociodemographic or disease-related variables. All patients were therefore considered as a group for subsequent analysis. In the following, the numbers of patients vary owing to missing data.

Acceptability of the questionnaire

Sixty per cent (76/122) of the patients spent less than 30 min on completing the questionnaire. Ten per cent (13/122) needed more than 45 min. Thirty-nine per cent (48/125) needed help, and half of them (27/48) received support from a family member. A small minority (5/126) indicated that some questions were difficult to answer.

EORTC QLQ-C30

Functioning scales. Forty-two to fifty-two per cent of the patients had trouble with physically demanding activities, while only 6% needed help for self-care activities (Table 4a). Thirty-seven per cent were limited in work or household jobs and 21% were unable to perform any of these activities. The mean scores of the physical functioning and role functioning scales were 0.74 and 0.70 respectively, and the reliability for both scales was 0.74. Cognitive functioning were assessed by a two-item scale. Ten to twenty per cent reported 'quite a bit' or 'very much' problems, and the mean scale score was 3.46. The scale reliability was 0.28. Fourteen to twenty-one per cent reported 'quite a bit' or 'very much' problems with emotional functioning. Here the mean scale score was 3.28 and the reliability 0.83. The disease had interfered in family life and social activities 'quite a bit' or 'very much' in 17% and 21% of the patients respectively. The mean scale score of the social functioning scale was 3.35, with a reliability of 0.77.

The majority of patients scored positively (4 or more) on the two questions about global physical functioning and global QOL. Six per cent scored 2 or less, and 37% scored 6 or more on the question about global physical functioning, with a mean of 4.78. The global QOL scored 2 or less by 3% of the patients, and 6 or more by 42%, and the mean was 4.92. The reliability of the QOL scale was 0.90 and the mean scale score was 4.85.

Symptom scales and single items. Twenty-two to twenty-seven per cent of the patients reported 'quite a bit' or 'very much' problems on the fatigue scale (Table 4b), 3–7% on the emesis scale, and 10–18% on the pain scale. The mean scale scores for the fatigue, emesis, and pain scales were 1.99, 1.24, and 1.55 respectively, with reliabilities of 0.70 or higher. On average, 6–28% of the patients reported 'quite a bit' or 'very much' problems with dyspnoe, sleeping, constipation, diarrhoea, appetite and impact on financial situation (single items, Table 4c). The mean single item score varied from 1.35 to 1.80.

The diagnosis-specific module. The mean scores of these items varied from 1.19 to 2.23. Only two questions concerning dryness in the mouth and salivation/mucus production had a mean score above 2. The frequency of patients reporting 'quite a bit' or 'very much' problems varied from 4% (5/123) (Have you had pain in the chest?) to 37% (47/122) (Have you been bothered with a lot of salivation/mucus production?).

Construct validity of the EORTC QLQ-C30. Two definite scaling errors and 14 probable scaling errors were found in the cognitive function scale (Appendix 1). All other items had a higher correlation with their own scale than to the other scales. Eighty per cent of these correlations were significant (exceeding two standard errors of the correlation). The physical functioning scale had one item that showed low correlation (0.41) within the scale. The remaining items except those in the cognitive functioning scale, correlated 0.52 or higher within the scales. The correlations between scales in the EORTC QLQ-C30 are

Table 3

Sociodemographic characteristics of the study sample (n = 126)

	No. of patients (%)
Age (years)	
Mean (range) 67 (20–99)	
<60	32 (25)
⩾60	94 (75)
Sex	
Female	29 (23)
Male	97 (77)
Education	
Compulsory only	61 (48)
Upper secondary	52 (41)
University < 4 years	11 (9)
University ⩾ 4 years	2 (2)

presented in Table 5. All correlations were in the right direction. The negative correlations are due to different directions of scorings on functioning and symptom scales. Physical functioning correlated 0.63 with role functioning and 0.55 with the QOL scale. Emotional functioning correlated −0.53 with fatigue, −0.57 with pain, and 0.56 with social functioning. The correlation between QOL and fatigue was −0.52. The other correlations were below 0.48. The correlation between GHQ-20 and the emotional functioning scale and QOL scale were −0.70 and −0.58 respectively (Table 6). The correlation between GHQ-20 and pain and fatigue were 0.52, between GHQ and global physical functioning, global QOL and social functioning from −0.54 to −0.64. The satisfaction with life question from the North Trøndelag study had a correlation of −0.45 with the global QOL question and −0.49 with the QOL scale but of −0.51 with social functioning. The strength and fitness question had a correlation −0.54 with the physical functioning scale, −0.64 with the QOL scale.

Disease and treatment variables may be divided on the basis of clinical experience into measurements of acute and subacute treatment-related symptoms or toxicity. There were significant mean score differences ($p < 0.05$) between patient subgroups (1–5) in nine of the diagnosis-specific

Table 4a

EORTC QLQ-C30. Item and scale scores and distribution of patients reporting problems with physical functioning or role functioning

Scale/item	n	%	Mean	Mean scale score (SD)	alpha (†)
Physical functioning				0.74 (0.28)	0.74
Trouble doing strenuous activ.	52/123	42	0.58		
Trouble taking a long walk	66/126	52	0.48		
Trouble taking a short walk	20/125	16	0.84		
Bed- or chair-ridden	15/124	12	0.88		
Help with eating, dressing	7/125	6	0.94		
Role functioning				0.70 (0.40)	0.74
Limited in work/housework	47/122	37	0.61		
Unable to do work/housework	26/123	21	0.69		

n refers to patients answering yes to the question. (The higher the score the higher the level of functioning.)
† = Chronbach's alpha

Table 4b

EORTC QLQ-C30. Item and scale scores and distribution of patients reporting 'quite a bit' or 'very much' problems with functioning

Scale/item	n	%	Mean	Mean scale score (SD)	alpha (†)
Cognitive functioning				3.46 (0.61)	0.28
Difficulty in concentrating	13/125	10	3.65		
Difficulty in remembering	25/124	20	3.26		
Emotional functioning				3.28 (0.71)	0.83
Did you feel tense?	19/126	15	3.30		
Did you worry?	23/124	18	3.22		
Did you feel irritable?	17/124	14	3.40		
Did you feel depressed?	26/123	21	3.20		
Social functioning				3.35 (0.86)	0.77
Disease interfered w.fam.life	21/124	17	3.43		
Disease interfered w.soc.act.	27/125	21	3.27		
Quality of life				4.85 (1.47)	0.90
Global physical functioning	27/123	22	4.78		
Global quality of life	23/124	19	4.92		

n refers to patients answering 'quite a bit' or 'very much'. For the last two items, n refers to patients answering 1, 2, 3, or 4 on the modified VAS scale. (The higher the score the higher the level of functioning.)
† = Chronbach's alpha

Table 4c

EORTC QLQ-C30. Item and scale scores and distribution of patients reporting 'quite a bit' or 'very much' symptoms

Scale/item	n	%	Mean	Mean scale score (SD)	alpha (†)
Fatigue				1.99 (0.74)	0.84
Did you need to rest?	28/124	22	1.96		
Have you felt weak	32/125	25	2.00		
Were you tired?	34/124	27	2.02		
Emesis				1.24 (0.53)	0.82
Have you felt nauseated?	9/125	7	1.30		
Have you vomited?	4/126	3	1.17		
Pain				1.55 (0.75)	0.70
Have you had pain?	23/121	18	1.66		
Did pain interf.w.activities?	13/121	10	1.45		

Single items	n	%	Mean	(SD)
Were you short of breath?	15/124	12	1.50	(0.79)
Have you had troub. sleeping?	20/126	16	1.63	(0.88)
Have you been constipated?	35/126	28	1.80	(1.04)
Have you had diarrhoea?	8/123	6	1.35	(0.68)
Have you lacked appetite?	26/126	21	1.70	(0.99)
Disease caused financ. diff.?	18/124	14	1.51	(0.92)

n refers to patients answering 'quite a bit' or 'very much'. (The higher the score the higher the level of symptoms.)
† = Chronbach's alpha

Table 5

Inter-scale correlation in the EORTC QLQ-C30

	PF	RF	CF	EF	SF	QOL	F	E
Scale								
Physical functioning (PF)								
Role functioning (RF)	0.63							
Cognitive functioning (CF)	0.26	0.23						
Emotional functioning (EF)	0.21*	0.21*	0.35					
Social functioning (SF)	0.28	0.33	0.34	0.56				
Quality of life (QOL)	0.55	0.44	0.27	0.36	0.45			
Fatigue (F)	−0.38	−0.39	−0.27	−0.53	−0.47	−0.52		
Emesis (E)	−0.10*	−0.14*	−0.06*	−0.30	−0.32	−0.25	−0.28	
Pain (P)	−0.21*	−0.21*	−0.23	−0.57	−0.42	−0.32	−0.38	−0.38

* = $p > 0.01$

items (Table 7a and b). Similar differences were found in the EORTC QLQ-C30 for two scales (fatigue, social functioning) and for two of the single items (constipation, global physical functioning). Problems with soreness in the mouth, with swallowing and with salivation/mucus production had the highest mean score in group 3 (after half way through the radiotherapy course), while change of taste and dryness in the mouth had the highest mean scores in group 4 (at the out-patient clinic). Group 4 also had the highest mean score on fatigue, and the lowest mean score on social functioning (reversed scale).

Discussion

The reliability of the EORTC QLQ-C30 was satisfactory (alpha > 0.69) for all scales except for the cognitive functioning scale, which had a reliability of 0.28. The correlation between the two items in this scale was low (0.17).

Several scaling errors were found in the cognitive functioning scale, which indicates that the psychometric properties of the scale are unsatisfactory. This is a new scale which replaces a single item in the previous version, EORTC QLQ-C36. The variability were very low in the same items.

The correlations between the different scales in the EORTC QLQ-C30 indicate that emotional functioning is related to pain, fatigue, and social functioning, which is as expected from clinical experience. Similar relationships were found between physical functioning and role function.

The correlation between GHQ-20 and the emotional functioning scale was 0.70. A higher correlation would have been expected if these two scales had measured exactly the

Table 6

Correlation of scales/items in the EORTC QLQ-C30 with total GHQ-20 score and items from the North Trøndelag Study

	Tot.GHQ-20 score	Satisfact. with life	Strength and fitness
Scales			
Physical functioning	−0.33	−0.24	−0.54
Role functioning	−0.33	−0.25	−0.40
Cognitive functioning	−0.40	−0.26	0.29
Emotional functioning	−0.70	−0.45	−0.30
Social functioning	−0.64	−0.51	−0.35
Quality of life	−0.58	−0.49	−0.64
Fatigue	0.52	0.48	0.56
Emesis	0.34	0.24	0.12*
Pain	0.52	0.36	0.14*
Single items			
Global physical functioning	−0.54	−0.48	−0.63
Global Quality of life	−0.56	−0.45	−0.60

Scales/items that were expected to be related are underlined in the table.
* = $p > 0.01$

same underlying dimension. The GHQ-20 probably measures several dimensions, as found in previous studies (43). The correlation between GHQ-20 and social function scale was 0.64, which support this assumption.

The correlation between the two questions in the QOL-scale, the global physical functioning question and the global QOL question was 0.81. This indicates that head and neck cancer patients do not discriminate between physical functioning and QOL. The high correlation may be a result of patients' misinterpretation of the term 'quality of life'. 'Quality of life' has been employed in advertising and in politics for the last ten or twenty years, but it is not part of everyday language. The high correlation may also be a result of the use of the modified VAS scales, and/or a design fault in the questionnaire. These two questions were placed next to each other and a contamination effect may have occurred. This possibility is strengthened by the low correlation between satisfaction with life and global QOL (r = 0.45). Patients' difficulties in scoring on VAS scales have been reported in earlier studies (33). It might be troublesome to analyse these two questions as a scale.

The discrimination between patient subgroups according to acute and subacute toxicity was satisfactory. The find-

Table 7a

Mean scores of items expected to measure acute treatment-related symptoms in the EORTC QLQ-C30 and the diagnosis-specific module for head and neck cancer patients

	Mean score				
	Gr 1	Gr 2	Gr 3	Gr 4	Gr 5
Item					
EORTC QLQ-C30					
Constipation**	1.67	1.90	2.57	2.00	1.52
Global physic. function*	5.05	4.76	3.85	4.21	5.14
Diagnosis-specific module					
Sore mouth or tongue*	1.29	1.29	2.14	2.00	1.65
Trouble swallowing*	1.10	1.57	2.92	2.29	1.57
Hoarseness**	1.62	2.05	2.86	2.30	1.44
Trouble talking on the phone**	1.52	1.62	2.43	1.90	1.41
Trouble swallowing bread**	1.38	1.57	2.85	2.32	1.68
Trouble swalling liquid**	1.00	1.29	2.14	1.35	1.12
Salivation/mucus production**	1.48	2.38	2.86	2.70	1.88

Significant variance between groups * = $p < 0.05$, ** = $p < 0.01$

Table 7b

Mean scores of scales/items expected to measure subacute treatment related symptoms in the EORTC QLQ-C30 and the diagnosis-specific module for head and neck cancer patients

	Mean score				
	Gr 1	Gr 2	Gr 3	Gr 4	Gr 5
Scales/items					
EORTC QLQ-C30					
Fatigue*	1.83	1.87	2.13	2.44	1.91
Social functioning (†)*	3.45	3.57	3.54	2.76	3.39
Diagnosis-specific module					
How much coughing**	1.57	1.62	2.14	2.15	1.62
Trouble with taste**	1.29	1.71	2.14	2.70	1.63
Dryness in the mouth**	1.76	2.05	2.15	3.11	2.19
Difficult to breath thr. nose**	1.10	1.14	1.29	1.95	1.29

Significant variance between groups * = $p < 0.05$, ** = $p < 0.01$
† = The higher the score, the higher the level of functioning.

ings are, as expected, owing to the early radiation effect on the mucous membranes and the subacute effect on the salivary glands and taste buds. This indicates that these questions will be satisfactory measures of acute, subacute and late toxicity.

In the diagnosis-specific module, items measuring pain in the head, eyes, ears and chest, breathing problems and dizziness had little variability both within and between patient subgroups. The two items concerning pain in the chest and troublesome breathing are used in the lung cancer-specific module, and do not seem relevant to head and neck cancer patients. It was difficult for the patients to distinguish between the two items measuring pain and soreness in the tongue and mouth. The 'soreness' item will later be omitted from the questionnaire.

The questionnaire was well accepted by the patients and compliance was high, even though the population had a mean age of 67 and low education. There were only minor problems in answering the questions except for the two VAS items. Low compliance has been found in an EORTC phase III study where the QOL assessment was optional. The clinicians did not trust QOL assessments and were not willing to make necessary regular and substained efforts (44). In our study, the nurses and physicians were often sceptical, whereas the patients were appreciative of our interest in measuring subjective phenomena.

The present validation study has only been done with indirect measures. It may be necessary to carry out further validation of the diagnosis-specific module.

In conclusion, the patient compliance was very high in this study. A satisfactory internal consistency were found in all scales exept in the cognitive functioning scale. There seem to be problems connected with the QOL-scale, with the wording in the global QOL question and with the use of the modified VAS scales in our population. The diagnosis-specific module was successful in discriminating between patient subgroups. Some items gave little variability, and will be omitted in later versions.

ACKNOWLEDGEMENTS

This study was supported by the Norwegian Cancer Society. We thank the members of the EORTC Study Group on Quality of Life for valuable discussions, and Dr. Neil Aaronson, The Netherlands Cancer Institute for helping us with analysing the item/scale correlation matrix (Appendix 1).

REFERENCES

1. Survival of Cancer in Norway. The Cancer Registry of Norway. Institute for Epidemiological Cancer Research, Oslo, 1990.
2. Incidence of Cancer in Norway 1989. The Cancer Registry of Norway. Institute for Epidemiological Cancer Research, Oslo, 1990.
3. Amdur RJ, Parsons JT, Medenhall WM, Million RR, Stringer SP, Cassisi NJ. Postoperative irradiation for squamous cell carcinoma of the head and neck: An analysis of treatment results and complications. Int J Radiat Oncol Biol Phys 1989; 16: 25–36.
4. Greco A. Rationale for chemotherapy for patients with advanced non-small-cell lung cancer. Semin Oncol 1986; 13: 92–6.
5. Klastersky J. Therapy with cisplatin and etoposide for non-small-cell lung cancer. Semin Oncol 1986; 13: 104–14.
6. Calman KC. Definitions and dimensions of quality of life. In: Aaronson NK, Beckmann J, eds. The quality of life of cancer patients. New York: Raven Press, 1987: 1–10.
7. Kaasa S, Mastekaasa A, Naess S. Quality of life of lung cancer patients in a randomized clinical trial evaluated by a psychosocial well-being questionnaire. Acta Oncol 1988; 27: 335–42.
8. Diener E. Subjective well-being. Psychol Bull 1984; 95: 542–75.

9. Calman KC. Quality of life in cancer patients. Curr Concepts Oncol 1984; 6: 2–3.
10. Andrews FM, McKennell AC. Measures of self-reported well-being: Their affective, cognitive, and other components. Social Indicators Research 1980; 8: 125–55.
11. Aaronson NK, Bullinger M, Ahmedzai S. A modular approach to quality-of-life assessment in cancer clinical trials. Recent Results Cancer Res 1988; 3: 231–49.
12. Spitzer WO. State of science 1986: Quality of life and functional status as target variables for research. J Chron Dis 1987; 40: 465–71.
13. Breslow L. A quantitative approach to the World Health Organization definition of health: Physical, mental and social well-being. Int J Epidemiol 1972; 1: 347–55.
14. Bradburn NM, Caplovitz D. Correlations of well-being. Reports on happinesss. Chicago: Aldine, 1965: 8–60.
15. Kaasa S, Mastekaasa A. Psychosocial well-being of patients with inoperable non-small cell lung cancer. The importance of treatment- and disease-related factors. Acta Oncol 1988; 27: 829–35.
16. Karnofsky DA, Abelmann WH, Craver LF, Burchenal JH. The use of the nitrogen mustards in the palliative treatment of carcinoma. Cancer 1948; 1: 634–56.
17. WHO handbook for reporting results of cancer treatment. Geneva, 1979.
18. Spitzer WO, Dobson AJ, Hall J, et al. Measuring the quality of life of cancer patients. J Chron Dis 1981; 34: 585–97.
19. Slevin ML, Plant H, Lynch D, Drinkwater FJ, Gregory WM. Who should measure quality of life, the doctor or the patient? Br J Cancer 1988; 57: 109–12.
20. Goldberg D, Williams P. A user's guide to the General Health Questionnaire. NFER–NELSON, Windsor, Bershire: 1988.
21. Bergner M, Bobbitt RA, Carter WB, Gilson BS. The Sickness Impact Profile: Development and final revision of a health status measure. Med Care 1981; 19: 787–806.
22. Hunt SM, McKenna SP, McEwen J, Williams J, Papp E. The Nottingham Health Profile: Subjective health status and medical consultations. Soc Sci Med 1981; 15: 221–9.
23. Chambers LW, MacDonald LA, Tugwell P, Buchanan WW, Kraag G. The McMaster Health Index Questionnaire as a measure of quality of life for patients with rheumatoid disease. J Rheumatol 1982; 9: 780–4.
24. Schag CC, Heinrich RL. Development of a comprehensive quality of life measurement tool: CARES. Oncology 1990; 4: 135–8.
25. Schag CC, Heinrich RL, Ganz PA. Cancer inventory of problem situations: An instrument for assessing cancer patients' rehabilitation needs. Psychosoc Oncol 1983; 1: 11–24.
26. de Haes JCJM, van Knippenberg FCE, Neijt JP. Measuring psychological and physical distress in cancer patients: structure and application of the Rotterdam symptom checklist. Br J Cancer 1990; 62: 1034–8.
27. Coates A, Dillenbeck CF, McNeil DR, et al. On the receiving end-II. Linear Analogue Self-Assessment (LASA) in evaluation of aspects of the quality of life of cancer patients receiving therapy. Eur J Cancer Clin Oncol 1983; 19: 1633–7.
28. Schipper H, Clinch J, McMurray A, Levitt M. Measuring the quality of life of cancer patients: The Functional Living Index—Cancer: Development and validation. J Clin Oncol 1984; 2: 472–83.
29. Priestman TJ, Baum M. Evaluation of quality of life in patients receiving treatment for advanced breast cancer. Lancet 1976; 1: 899–900.
30. Selby PJ, Chapman JAW, Etazadi-Amoli J, Dalley D, Boyd NF. The development of a method for assessing the quality of life of cancer patients. Br J Cancer 1984; 50: 13–22.
31. Levine MN, Guyatt GH, Gent, M, et al. Quality of life in stage II breast cancer: An instrument for clinical trials. J Clin Oncol 1988; 6: 1798–810.
32. Moinpour CM, Feigl P, Metch B, Hayden A, Meyskens FL, Crowly J. Quality of life end points in cancer clinical trials: Review and recommendations. JNCI 1989; 81: 485–95.
33. Ganz PA, Haskell CM, Figlin RA, Lasoto N, Sian J. Estimating the quality of life in a clinical trial of patients with metastatic lung cancer using the Karnofsky Performance Status and the Functional Living Index—Cancer. Cancer 1988; 61: 849–56.
34. Ware JE. Standards for validating health measures: Definition and content. J Chron Dis 1987; 40: 473–80.
35. Kaasa S. Quality of life and survival in patients with non-small cell lung cancer. The Norwegian Cancer Society, Oslo, 1989.
36. Carmines EG, Zeller RA. Assessing reliability. In: Reliability and validity assessment. Beverly Hills/London: Stage Publications, 1979: 37–57.
37. Eriksen J, Naess S. Funksjonshemmede i Nord-Trøndelag, Notat 86:6, Institute for applied social research, Oslo, 1986.
38. Aaronson NK, Ahmedzai S, Bullinger M, et al. The EORTC Core Quality of Life Questionnaire: Interim results of an international field study. In: Osoba D, ed. Effect of cancer on quality of life. New York: CRC Press, 1991: 185–202.
39. Gotestam C, Kaasa S. The quality of life in patients with cancer in ear, nose or throat before, during, and after treatment (Abstract). 24th annual AABT (American Association of Behavior Therapy) convention, San Francisco, 1–4 Nov., 1990.
40. Malt UF, Mogstad TE, Refnin IB. Goldberg's General Health Questionnaire. Tidsskr Nor Lægeforen 1989; 109: 1391–4.
41. Chronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrica 1951; 16: 297–334.
42. Norusis MJ. SPSS/PC + ™V3.0. Chicago: SPSS Inc, 1986.
43. Worsley A, Gribbin CC. A factor analytic study of the twelve-item General Health Questionnaire. Aust N Z J Psychiatry 1977; 11: 269–72.
44. Fosså SD, Aaronson NK, Newling D, et al. Quality of life and treatment of hormone resistant metastatic prostatic cancer. Eur J Cancer 1990; 26: 1133–6.