ELSEVIER

European Journal of Cancer 40 (2004) 2260–2268

European Journal of Cancer

www.ejconline.com

Clinical and psychometric validation of a questionnaire module, the EORTC QLQ-STO 22, to assess quality of life in patients with gastric cancer

J.M. Blazeby [a,*], T. Conroy [b], A. Bottomley [c], C. Vickery [a], J. Arraras [d], O. Sezer [e], J. Moore [f], M. Koller [g], N.S. Turhal [h], R. Stuart [i], E. van Cutsem [j], S. D'haese [c], C. Coens [c], on behalf of the European Organisation for Research and Treatment of Cancer Gastrointestinal and Quality of Life Groups

[a] *Clinical Sciences at South Bristol and Department of Social Medicine, Bristol Royal Infirmary, University of Bristol, Bristol, UK*
[b] *Department of Medical Oncology, Centre Alexis Vautrin, Vandoeuvre-Lès-Nancy, France*
[c] *EORTC Quality of Life Unit, EORTC Data Centre, Brussels, Belgium*
[d] *Department of Oncology, Hospital De Navarra, Pamplona, Spain*
[e] *Department of Oncology & Hematology, Universitätsklinikum Charité, Berlin, Germany*
[f] *Department of Surgery, St James's Hospital, Trinity College, Dublin, Ireland*
[g] *Institute of Theoretical Surgery, Philipps-University, Marburg, Germany*
[h] *Marmara University Hospital, Istanbul, Turkey*
[i] *Glasgow Royal Infirmary, University of Glasgow, Glasgow, UK*
[j] *Department Internal Medicine, University Hospital Gasthuisberg, Leuven, Belgium*

Received 19 May 2004; accepted 20 May 2004
Available online 10 August 2004

Abstract

The purpose of this study was to define the measurement properties and clinical validity of the European Organisation for Research and Treatment of Cancer (EORTC) questionnaire module to assess health-related quality of life (HRQL) in gastric cancer. The EORTC gastric cancer module, QLQ-STO 22, was administered with the QLQ-C30, core questionnaire, to 219 patients undergoing treatment with curative or palliative intent before and after treatment. Reliability and validity of the module was tested and patients' debriefing comments analysed. Compliance rates were high, questionnaires well accepted and less than 4% of items had missing data. Multi-trait scaling analyses and face validity refined the module to five scales and four single items. Scales distinguished between clinically distinct groups of patients and demonstrated treatment-induced changes over time. Test–retest scores demonstrated good reliability. The EORTC QLQ-STO 22 demonstrates psychometric and clinical validity that supports its use to supplement the EORTC QLQ-C30 to assess quality of life in patients with gastric cancer undergoing surgery, surgery and chemoradiotherapy, palliative chemotherapy, palliative surgery and best supportive care.

Keywords: Quality of life; Gastric cancer; Questionnaires; QLQ-STO 22; QLQ-C30; Psychometrics

1. Introduction

Most patients in the Western part of the world with gastric cancer have locally inoperable or metastatic disease at presentation, with only 20–30% of patients being

* Corresponding author. Tel.: +44-117-9283495; fax: +44-117-9252736.
E-mail address: j.m.blazeby@bristol.ac.uk (J.M. Blazeby).

0959-8049/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ejca.2004.05.023

suitable for potentially curative surgery. Although surgery offers a hope of cure, overall five year survival after gastrectomy is approximately 25% [1]. Efforts to improve survival include using pre- or post-operative chemotherapy or chemoradiotherapy or extended resection with D2 lymphadenectomy. Improvements in survival with multi-modal treatment and more radical resection may also be associated with increased toxic side-effects and peri-operative morbidity [2–4]. Palliative treatment for gastric cancer may include palliative surgery or chemotherapy and there may be a role for chemotherapy to downstage locally advanced disease and render surgical resection possible [5–7]. Full evaluation of new treatments and combination treatment for gastric cancer is essential and patient-based outcome measures as well as assessment of biomedical outcomes are required.

Patient-based outcomes include survival data, symptom assessment and health-related quality of life (HRQL) data [8]. Measurement of HRQL has developed over the past decade and there are several generic questionnaires designed for patients with cancer, such as the Functional Assessment of Cancer Therapy-G (FACT-G) and European Organisation for Research and Treatment of Cancer Core Quality of Life – questionnaire (EORTC QLQ-C30) [9,10]. Generic cancer questionnaires may be supplemented by site-specific modules to increase sensitivity and specificity. The EORTC Quality of Life Group has developed a questionnaire module for patients with gastric cancer that assesses HRQL issues related to dysphagia, eating restrictions, reflux, and abdominal pain as well as specific symptoms that may occur during chemotherapy or radiation treatment [11]. The aim of this study was to test the psychometric properties and clinical validity of the EORTC gastric cancer module in an international group of patients undergoing treatment for gastric cancer.

2. Patients and methods

2.1. Patients

This study opened in April 2001 and closed in May 2003. It was co-ordinated at the Quality of Life Unit at the EORTC Data Centre in Brussels, Belgium (Protocol 15001/40003). Patients were prospectively registered before treatment and were eligible to participate if they had a diagnosis of gastric adenocarcinoma that had been confirmed histologically and they had an expected survival time of at least four weeks. Written informed consent was obtained. Exclusion criteria were: concurrent malignancies, inability to understand and complete the questionnaire and participation in another HRQL study that would interfere with this protocol. There were no limitations for either age or performance status. Investigators had to obtain local or national ethical committee approval to be eligible to recruit patients into this field study.

Patients were staged and selected for treatment according to local policies and entered into two predetermined groups for the purpose of questionnaire validation: Group A consisted of patients who were selected for potentially curative treatment (subgroups: (1) surgery alone; (2) neoadjuvant chemotherapy and/or radiotherapy and surgery; (3) surgery and adjuvant chemotherapy and/or radiotherapy; (4) endoscopic mucosal resection or laparoscopic wedge excision of an early gastric cancer). Group B consisted of patients selected for treatment with palliative intent, (subgroups: (1) surgery with palliative intent; (2) endoscopic procedure (e.g. stent insertion); (3) chemotherapy and/or radiotherapy with palliative intent; (4) best supportive measures available in each participating centre).

2.1.1. Questionnaires and data collection

Quality of life assessments were conducted at various time points. In Group A, patients undergoing surgery alone (subgroups 1 and 4), completed a baseline assessment within 4 weeks before the first day of treatment and a second assessment 3 months after the day of surgery (plus/minus 3 weeks). In the case of United Kingdom (UK) patients, a third assessment was planned with patients given a set of questionnaires to take home and asked to complete them 3–5 days later and return them by post (test–re-test study). Patients undergoing neoadjuvant chemo/radiotherapy and surgery (subgroup 2), completed questionnaires within 4 weeks before starting chemotherapy and/or radiotherapy, a second assessment within 4 weeks of surgery and a final assessment 3 months after the day of surgery (plus/minus 3 weeks). Subgroup 3, however, completed questionnaires within 4 weeks before starting adjuvant chemotherapy/radiotherapy and a second assessment 6 weeks (plus/minus 3 weeks) after completion of chemotherapy/radiotherapy.

Patients in Group B completed the questionnaire within three weeks before starting palliative treatment and a second assessment four weeks after the start of the treatment (plus/minus 2 weeks). Patients with HRQL assessments outside of these timeframes were excluded from the data analyses.

At all time-points, the EORTC QLQ-C30 (version 3.0), and the EORTC QLQ-STO 22 were administered to patients. At the baseline assessment, patients were asked to complete a short debriefing questionnaire covering questions about the time taken to complete the EORTC QLQ-C30 and quality of life questionnaire (Stomach module) (QLQ-STO 22), the need for help in completing the questionnaires and querying if any of the items appeared confusing, difficult to answer or upsetting. Sociodemographic data were recorded. This

included: gender, age (date of birth), marital status, cohabitants, education and employment status. In addition, the following clinical data were collected for each patient: anatomical position of tumour, site of metastases (at first presentation or recurrence) co-morbid disease, weight measured at each interview using the same scales each time, Karnofsky Performance Status and dysphagia grade at each assessment date and nature of surgery, date(s) and number of fractions and total dose of external radiotherapy, date, dose and type of chemotherapy and number of treatments, date and type of other supportive measures employed, date and type of endoscopic treatment.

Compliance with HRQL measures and other general aspects of this field study were monitored using standard EORTC procedures and reviewed by investigators every six months at bi-annual quality of life group meetings.

The EORTC QLQ-STO 22 module contains 22 items in a similar layout and response format to the EORTC QLQ-C30. The initial stages of the development of this measure in four countries has been detailed previously in [11]. The hypothesised scale structure of the module consists of five scales (dysphagia, eating restrictions, pain, reflux and anxiety) and three single items (dry mouth, body image and hair loss).

2.2. Statistical analysis

The primary objective of the study is to test the scale structure, reliability and validity of the gastric cancer module (EORTC QLQ-STO 22) and its sensitivity to change in clinical health status. Multi-trait scaling analyses were used to examine whether the individual items composing the EORTC QLQ-STO 22 may be pooled into a hypothesised clinical scale structure with a more limited set of multi-item scales. This technique that is based on an examination of item-scale correlations has been used extensively in evaluating the structure of health status measures. Evidence of item convergent validity will be defined as correlation of 0.40 or greater between an item and its own scale (corrected for overlap) [12].

2.2.1. Validity

Discriminant validity was examined with item convergent validity for each scale being assessed using the correlation between each item and its own scale corrected for overlap. A scaling success for an item is seen when the correlation between an item and its own scale was significantly higher than its correlation with any other scale.

2.2.2. Clinical validity

Known group comparisons were examined by exploratory analysis of the clinical validity of the QLQ-STO 22 using the method of known-group comparison i.e. to explore the extent to which the questionnaire scores are able to discriminate between subgroups of patients differing in terms of their clinical status. Known groups used for this comparison were baseline treatment groups (potentially curative vs. palliative) and baseline Karnofsky scores (< 80 vs. > 80). Group differences were assessed using the Wilcoxon rank sum test.

The availability of two sets of questionnaires, one prior to the start of treatment and one during the treatment period, allows a preliminary evaluation of the responsiveness of the QLQ-STO 22 to changes in health status over time. Improvement or deterioration in health status was used on the basis of a shift of at least one level on the Karnofsky Performance Status Scale and weight loss of more than 10% between the baseline and follow-up assessment. Repeated measures ANOVA (analysis of variance) was used to test for the significance of changes in HRQL scores as a function of observed changes in clinical status over time. All tests were performed using statistical analysis software (SAS).

2.2.3. Reliability

The test–retest reliability of scales and single item measures was assessed using intraclass correlations between the second and retest assessments in the patients undergoing surgery alone in Group A in centres in the UK. These patients completed a few additional items on the debriefing questionnaire that assess whether health had changed over the days that the test–retest study was performed. Patients who reported a change in health status (e.g. development of a concurrent illness) were excluded from test–retest analyses.

3. Results

3.1. Patient characteristics

A total of 267 gastric cancer patients from 14 institutions in eight different countries were entered in the study. The potentially curative treatment group contained 127 and 137 patients underwent treatment with palliative intent. A further three patients had no assigned group. Of the 267 patients, 11 were classed as ineligible (e.g. identified with a concurrent cancer, prognosis was less than four weeks) and of the remaining 256 patients, eight patients were considered not evaluable (most frequently related to different treatment received than originally expected). Of the remaining 248 patients, 219 (88.3%) had at least one HRQL form present. For 29 patients, the HRQL form was not present or outside the time window so they were excluded. Sociodemographic and clinical characteristics of evaluable patients are shown in Table 1.

Table 1
Clinical and sociodemographic characteristics of the evaluable patients

	Potentially curative Group A ($n = 108$)	Palliative treatment Group B ($n = 111$)
Gender (%)		
Male	71 (65.7)	75 (67.6)
Female	37 (34.3)	36 (32.4)
Age (in years)		
Median (range)	65.7 (36.7–87.3)	67.1 (29.6–92.0)
Co-habitants		
Alone	19 (17.6)	17 (15.3)
With family or others	89 (82.4)	94 (84.7)
Marital status		
Single	8 (7.4)	12 (10.8)
Married	75 (69.4)	72 (64.9)
Separated, divorced, widowed	25 (23.1)	27 (24.3)
Education		
Less than compulsory	13 (12.0)	15 (13.5)
Compulsory	64 (59.3)	64 (57.7)
Post-compulsory	18 (16.7)	17 (15.3)
University	13 (12.0)	12 (10.8)
Unknown	0 (0)	3 (2.7)
Employment		
Full- or part-time	32 (29.6)	34 (30.6)
Homemaker	8 (7.4)	6 (5.4)
Unemployed	3 (2.8)	10 (9.0)
Retired	65 (60.2)	59 (53.2)
Other	0 (0.0)	2 (1.8)
Anatomical location of tumour		
Proximal stomach & cardia	42 (38.9)	30 (27.0)
Body of stomach	33 (30.6)	35 (31.5)
Distal stomach	28 (25.9)	23 (20.7)
Overlapping	4 (3.7)	15 (13.5)
Unknown	1 (0.9)	8 (7.2)
Country		
UK	34 (31.5)	48 (43.2)
France	20 (18.5)	10 (9.0)
Spain	19 (17.6)	13 (11.7)
Germany	12 (11.1)	19 (17.1)
Republic of Ireland	10 (9.3)	5 (4.5)
Australia	1 (0.9)	1 (0.9)
Turkey	11 (10.2)	8 (7.2)
Belgium	1 (0.9)	7 (6.3)
Karnofsky status (%)		
< 60	1 (0.9)	14 (12.6)
60–80	28 (25.9)	51 (46.0)
> 80	79 (73.2)	46 (41.4)
Treatment subgroup (%)		
Total or partial gastrectomy	59 (54.6)	6 (5.4)
Neoadjuvant treatment + surgery	7 (6.5)	0 (0)
Surgery + adjuvant treatment	26 (24.1)	0 (0)
Endoscopic mucosal resection	3 (2.8)	0 (0)
Surgical bypass	0 (0)	8 (7.2)
Endoscopic treatment	0 (0)	4 (3.6)
Palliative chemotherapy/radiotherapy	0 (0)	64 (57.7)
Best supportive care	0 (0)	29 (26.1)
Other	13 (12.0)	0 (0)

UK, United Kingdom.

3.2. Compliance rates and debriefing questionnaire

Patient compliance of the 219 patients was 100% at the baseline assessment and 86.8% at the follow-up. For the test–re-test planned only for patients in the UK, some 34 patients were included and compliance was 71% with 24 patients returning the HRQL forms. Most patients (82%) completed both of the questionnaires in less than 15 min, and 53% did not require any help. In the 45% of patients who needed help, it was minimal and mostly in order to help read the items. The proportion of patients requiring help was significantly greater in the palliative group than in the potentially curative group ($P = 0.02$). Most patients ($n = 195$, 89%) found that the questions were clear and less than 4% found any items upsetting.

3.3. Defining HRQL scales and items

Results of the multi-trait scaling analyses are shown in Table 2. Item scale correlations in the reflux (STORFX), pain (STOPAIN) and anxiety (STOANX) scale exceeded 0.69. These scales were therefore retained in their original form. In the hypothesised dysphagia scale (STODYS), correlations exceeded 0.60, but item 34 addressing discomfort when eating was also highly correlated with the pain scale. It was therefore including in the pain scale to retain the clinical focus of each scale. Item 45, addressing taste problems was part of the eating restrictions scale. However, when tested as a single item, the item scale correlations of the eating scale improved. Item 45, was therefore maintained as a single item in the final module. The single items (hair loss, dry mouth and body image) were retained in their original form. Therefore, the final module (QLQ-STO 22) has five scales and four single items (Fig. 1). The following results use the scales and items in the QLQ-STO 22.

3.4. Reliability

Cronbach's α coefficient was lowest in the reflux and anxiety scales (0.72 and 0.73, respectively). In the remaining scales, it was 0.80 (Table 2). About 24 patients returned a third HRQL assessment for the test–retest study. The pain, eating restrictions and anxiety scales showed good reproducibility with interclass correlations above 0.70 and single item interclass correlations were also high (> 0.79). Interclass correlations for the dysphagia and reflux scales were 0.60 and 0.63, respectively.

3.5. Relationships between the STO 22 module and the QLQ-C30 core questionnaire

Most scales in the QLQ-STO 22 were weakly correlated with the QLQ-C30 scales. The gastric dysphagia scale (STODYS) was moderately correlated with QLQ-C30 appetite loss, nausea and vomiting, physical function scale and fatigue. Likewise, the gastric eating restrictions scale was moderately correlated with QLQ-C30 appetite loss, fatigue and physical function. The gastric pain scale was moderately correlated with the QLQ-C30 pain, emotional and social function scales. These correlations demonstrate the clinical overlap between the subscales and were expected. The scales were not changed because of the need to focus on eating-related issues in this gastric module.

3.6. Clinical validity

Patients in clinically distinct groups reported significant differences in baseline HRQL scores in several scales and items (Table 3). Differences were statistically significant in the dysphagia, and eating scales in single items assessing taste and dry mouth ($P < 0.01$). The reflux scale did not demonstrate statistical differences in either clinical subgroup and the single item assessing hair loss did not show differences between these selected groups of patients.

Clinical parameters were used to assess the responsiveness to treatment over time in HRQL scores before and after treatment. The reflux scale of the QLQ-STO 22 demonstrated sensitivity to changes in weight loss over time, $P = 0.003$ (Table 4). Dysphagia scores as rated by clinicians were compared with changes in HRQL scales in the STO 22 before and after treatment. Scales assessing dysphagia, pain, reflux and eating were all sensitive to changes in observer-rated dysphagia

Table 2
Reliability – Convergent and discriminant validity of multi-item scales ($n = 219$ all evaluable forms)

STO 22	Item own scale correlation[a]	Item other scale correlation	Cronbach's α
STODYS (dysphagia)	0.77–0.88	0.24–0.66	0.80
STOPAIN (pain)	0.66–0.85	0.24–0.70	0.80
STORFX (reflux)	0.73–0.82	0.24–0.49	0.72
STOEAT (eating)	0.60–0.85	0.22–0.65	0.80
STOANX (anxiety)	0.70–0.85	0.22–0.55	0.73

[a] Corrected for overlap.

EORTC QLQ – STO 22

Patients sometimes report that they have the following symptoms or problems. Please indicate the extent to which you have experienced these symptoms or problems during the past week. Please answer by circling the number that best applies to you.

During the past week:	Not at All	A Little	Quite a Bit	Very Much
31. Have you had problems eating solid foods?	1	2	3	4
32. Have you had problems eating liquidised or soft foods?	1	2	3	4
33. Have you had problems drinking liquids?	1	2	3	4
34. Have you had discomfort when eating?	1	2	3	4
35. Have you had pain in your stomach area?	1	2	3	4
36. Have you had discomfort in your stomach area?	1	2	3	4
37. Did you have a bloated feeling in your abdomen?	1	2	3	4
38. Have you had trouble with acid or bile coming into your mouth?	1	2	3	4
39. Have you had acid indigestion or heartburn?	1	2	3	4
40. Have you had trouble with belching?	1	2	3	4
41. Have you felt full up too quickly after beginning to eat?	1	2	3	4
42. Have you had trouble enjoying your meals?	1	2	3	4
43. Has it taken you a long time to complete your meals?	1	2	3	4
44. Have you had a dry mouth?	1	2	3	4
45. Did food and drink taste different from usual?	1	2	3	4
46. Have you had trouble with eating in front of other people?	1	2	3	4
47. Have you been thinking about your illness?	1	2	3	4
48. Have you worried about your weight being too low?	1	2	3	4
49. Have you felt physically less attractive as a result of your disease or treatment?	1	2	3	4
50. Have you worried about your health in the future?	1	2	3	4
51. Have you lost any hair?	Yes	No		
52. Answer this question only if you lost any hair: If so, were you upset by the loss of your hair?	1	2	3	4

Fig. 1. The EORTC QLQ-STO22.

scores ($P < 0.01$). Likewise, the physical function scale of the QLQ-C30 was strongly related to changes in Karnofsky scores over time ($P < 0.01$).

Changes of HRQL scores were examined in relation to treatment group. Three months after gastrectomy, decreased physical function was reported and increased fatigue, diarrhoea and poor body image ($P < 0.01$) (Table 5). After palliative treatment, significant deteriorations were also reported in physical function, taste and hair loss (Table 5).

Table 3
Mean baseline mean scores (SD) for clinically distinct groups of patients

HRQL scales	Treatment intent			Karnofsky score		
	Curative-*n* = 108	Palliation-*n* = 111	*P* value[a]	< 80-*n* = 94	> 80-*n* = 125	*P* value[a]
STODYS (dysphagia)	14 (21)	25 (29)	< 0.01	26 (28)	15 (24)	< 0.01
STOPAIN (pain)	21 (20)	28 (25)	0.09	26 (24)	24 (23)	0.72
STORFX (reflux)	17 (21)	21 (25)	0.42	22 (24)	17 (23)	0.06
STOEAT (eating)	21 (23)	34 (27)	< 0.01	33 (25)	25 (26)	< 0.01
STOANX (anxiety)	45 (27)	50 (28)	0.20	50 (29)	45 (27)	0.19
STODM (dry mouth)	29 (33)	40 (38)	0.04	42 (37)	29 (34)	< 0.01
STOTA (taste)	15 (28)	26 (34)	< 0.01	26 (33)	17 (17)	0.01
STOBI (body image)	10 (25)	18 (29)	< 0.01	14 (25)	14 (29)	0.60
STOHAIR (hair loss)	3 (13)	6 (17)	0.23	5 (15)	4 (15)	0.45

HRQL, Health – related quality of life; SD, standard deviation.
[a] Wilcoxon rank sum test.

Table 4
Clinical validity: mean difference HRQL scores between follow-up and baseline by weight loss

HRQL scales	Weight loss			
	Weight gain or no loss, *n* = 55	0–10% weight loss, *n* = 83	> 10% weight loss, *n* = 81	*P* value[a]
STODYS (dysphagia)	−3.37(22.6)	0.6 (24.4)	8.2 (27.3)	0.06
STOPAIN (pain)	0.6 (21.2)	0.9 (25.6)	2.6 (21.6)	0.89
STORFX (reflux)	−5.7 (17.5)	−2.9 (27.2)	9.3 (26.2)	< 0.01
STOEAT (eating)	2.6 (23.7)	2.7 (26.9)	10.2 (26.0)	0.21
STOANX (anxiety)	−3.4 (18.9)	−0.7 (26.8)	3.1 (37.2)	0.49
STODM (dry mouth)	0.0 (31.7)	−0.4 (34.5)	3.1 (41.6)	0.85
STOTA (taste)	12.4 (28.3)	12.9 (37.7)	18.5 (33.4)	0.57
STOBI (body image)	0.6 (22.4)	9.9 (32.3)	17.3 (32.2)	0.02
STOHAIR (hair loss)	13.5 (24.0)	9.8 (27.4)	1.7 (17.6)	0.05

[a] ANOVA (analysis of variance).

Table 5
Clinical validity: changes in mean scores over time of scales and items in the QLQ-C30 and STO 22 by treatment group

HRQL scales & items	Total or partial gastrectomy			Palliative treatment		
	Baseline, *n* = 65	Follow-up, *n* = 51	*P* valu	Baseline, *n* = 111	Follow-up, *n* = 99	*P* value
PF	85 9 (17.8)	72.5 (21.4)	< 0.01	73.3 (22.9)	60.1 (26.6)	< 0.01
RF	73.3 (32.9)	61.7 (33.4)	0.06	60.4 (35.0)	60.5 (35.0)	0.04
EF	68.4 (25.6)	72.4 (24.4)	0.39	70.4 (23.8)	70.4 (23.8)	0.82
CF	83.9 (19.9)	76.8 (27.3)	0.11	80.8 (21.2)	80.8 (21.2)	0.17
SF	80.5 (27.9)	69.7 (30.8)	0.05	70.6 (33.1)	70.6 (33.1)	0.31
QOL	65.3 (20.0)	60.1 (23.8)	0.21	51.2 (23.6)	51.2 (23.6)	0.91
FA	28.3 (25.8)	41.4 (28.1)	0.01	45.5 (29.2)	45.5 (29.2)	0.34
NV	14.6 (27.1)	19.9 (29.1)	0.31	19.7 (27.8)	19.7 (27.8)	0.61
PA	19.7 (21.4)	24.8 (29.1)	0.28	28.2 (30.6)	28.2 (30.6)	0.50
SOB	13.8 (22.7)	26.8 (43.2)	0.04	19.5 (28.6)	20.1 (28.8)	0.73
DIA	7.7 (17.5)	24.0 (28.6)	< 0.01	10.8 (20.7)	18.2 (27.1)	0.03
STODYS	11.6 (21.8)	15.8 (20.5)	0.30	14.1 (21.1)	14.1 (19.9)	0.73
STOPAIN	22.6 (22.1)	27.0 (23.9)	0.31	21.4 (20.0)	23.6 (21.9)	0.63
STORFX	16.8 (23.7)	19.0 (23.2)	0.63	16.7 (21.3)	16.5 (21.9)	0.21
STOEAT	18.2 (22.9)	27.7 (25.7)	0.04	21.2 (22.8)	25.0 (24.4)	0.84
STOANX	48.0 (27.3)	41.7 (31.7)	0.26	45.2 (27.4)	41.5 (30.0)	0.66
STODM	36.4 (35.2)	30.0 (35.1)	0.33	29.3 (33.1)	24.9 (31.0)	0.61
STOTA	13.3 (27.5)	26.0 (35.8)	0.03	14.5 (27.5)	25.6 (31.6)	< 0.01
STOBI	7.2 (21.6)	19.3 (30.2)	0.01	10.2 (24.7)	23.0 (34.1)	0.48
STOHAIR	4.6 (16.0)	3.9 (13.9)	0.82	3.0 (12.8)	6.9 (16.6)	< 0.01

QLQ-C30 scales (high score = better function): PF, physical; RF, role; EF, emotional; CF, cognitive; SF, social; QOL, overall quality of life; QLQ-C30 symptoms (high score = more problems): FA, fatigue; NV, Nausea & vomiting; PA, pain; SOB, dyspnoea; DIA, diarrhoea.
STO 22 scales and items (high score = more problems): STODYS = dysphagia; STOPAIN = chest and abdominal pain; STORFX = reflux; STO-EAT = eating restrictions; STOANX = anxieties; STODM = dry mouth; STOTA = taste problems; STOBI = body image; STOHAIR = hair loss.

4. Discussion

This study tested the EORTC QLQ-STO 22 in an international sample of patients with gastric cancer. Results confirmed three of the hypothesised scales (reflux, pain and anxiety) and three single items (hair loss, dry mouth, body image). Changes were made to the dysphagia and eating scales and as a result a slightly revised five scale questionnaire with four single items emerged, the EORTC QLQ-STO 22. Further testing demonstrated that this module was reliable and sensitive to changes in health status as well as being able to discriminate between clinically distinct groups of patients with gastric cancer. Debriefing information from patients did not identify any major omissions from the questionnaire and it was easily completed within 15 min by most patients. The EORTC QLQ-STO 22, is therefore recommended as a reliable and valid tool to use with the QLQ-C30 to assess HRQL in patients with gastric cancer.

Measurement of HRQL in patients with gastric cancer is important and several studies have prospectively used questionnaires after surgery, chemotherapy or combination treatment [13–15]. Questionnaires used in these studies include the Rotterdam Symptom Checklist, the Gastro Intestinal Quality of Life Index, the EORTC QLQ-C30 and the Spitzer Quality of life Index [16–18]. Although some studies have yielded useful clinical data, most have not been sufficiently powered to detect clinically significant HRQL differences between groups of patients and it is possible that generic cancer questionnaires are not sensitive to detect HRQL issues of importance to patients within specific treatment groups. The inclusion of a site-specific measure in addition to a generic measure will help address these problems.

Site-specific questionnaires to measure HRQL in patients with gastric cancer have been developed by the FACT-G system and by a German group [15,19]. Scaling data are available to support the FACT-Ga in Japanese and English, and some psychometric data from the German group have been collected, but only from patients undergoing gastrectomy [15]. The EORTC QLQ-STO 22 has been developed in four European languages and tested in eight countries. Translations into Danish, Dutch, French, German, Greek, Hungarian, Italian, Japanese, Korea, Norwegian, Portuguese, Brazilian, Russian, Spanish, Taiwanese and Turkish are now available. It has also been developed in patients undergoing surgery in the curative or palliative settings as well as in patients undergoing adjuvant chemotherapy or chemoradiation treatment, palliative chemotherapy, endoscopic treatment and best supportive care.

The use of questionnaires to assess HRQL in patients with tumours of the gastro-oesophageal junction remains an area that requires further work. At present, these may be treated as oesophageal tumours and thus we recommend using the EORTC QLQ-OES18 module [20]. However, type III junctional tumours have more features similar to gastric tumours and the EORTC QLQ-STO 22 is therefore recommended for these patients. Future work to combine these modules into a single gastro-oesophageal module is still required.

Conflict of interest

None

Acknowledgements

The authors acknowledge Mr. C.P. Barham, Mr. J. Vickers and Professor D. Alderson, (Bristol Royal Infirmary, UK), Professor J. Reynolds, (St. James's Hospital, Trinity College, Dublin, Ireland), Prof. Dr. H. Lochs, (Gastroenterology, University Hospital Charité, Berlin), Prof. Dr. J.M. Müller, (Surgery, University Hospital Charité, Berlin), Professor E. Van Cutsem, (U.Z Gasthuisberg, Leuven, Belgium), Dr. M. Ducreux, (Institut Gustave Roussy, Villejuif, France), Dr. J.M. Vieitez (Hospital General de Asturias, Oviedo, Spain), Dr. F. Lordik, (Cliniqum Rechts der Isar, Munchen, Germany), Dr. U. Yilmaz, (Dokuz, Eylul University, Izmir, Turkey), Dr. N. Spry (Sir Charles Gairdner, Hospital, Nedlands, Australia) and M. McKernan (Glasgow Royal Infirmary, University of Glasgow, Glasgow, UK) for supporting this study.

J.M. Blazeby is supported by a UK Medical Research Council, Clinician Scientist Award and O. Sezer is supported by a research grant from the Hector Stiftung, Weinheim, Germany.

References

1. Hundahl SA, Phillips JL, Menck HR. The national cancer data base report on poor survival of US gastric carcinoma patients treated with gastrectomy. *Cancer* 2000, **88**, 921–932.
2. Macdonald JS, Smalley SR, Benedetti J, Hundahl SA, Estes NC, Stemmermann GN, *et al*. Chemoradiotherapy after surgery compared with surgery alone for adenocarcinoma of the stomach or gastroesophageal junction. *New England Journal of Medicine* 2001, **345**, 725–730.
3. Cuschieri A, Fayers P, Fielding J, Craven J, Bancewicz J, Joypaul V, *et al*. Postoperative morbidity and mortality after D1 and D2 resections for gastric cancer. Preliminary results of the MRC randomised controlled surgical trial. *Lancet* 1996, **347**, 995–999.
4. Bonenkamp JJ, Hermans J, Saaako M, Boon MC, Bunt AMG. Meta-analysis of adjuvant chemotherapy for gastric cancer. *Proceedings of the American Society of Clinical Oncology* 1993, **12**, 554.
5. Ross P, Nicolson M, Cunningham D, Valle J, Seymour M, Harper P, *et al*. Prospective randomized trial comparing mitomycin cisplatin, and protracted venous-infusion fluorouracil (PVI 5-FU) With epirubicin, cisplatin, and PVI 5-FU in advanced

esophagogastric cancer. *Jornal of Clinical Oncolology* 2002, **20**, 1996–2004.
6. Glimelius B, Ekstrom K, Hoffman K, Graf W, Sjoden PO, Haglund U, *et al*. Randomised comparison between chemotherapy plus best supportive care with best supportive care in advanced gastric cancer. *Annals of Oncology* 1997, **8**, 163–168.
7. Waters JS, Norman A, Cunningham D, Scarffe JH, Webb A, Harper P, *et al*. Long-term survival after epirubicin, cisplatin and fluorouracil for gastric cancer: results of a randomized trial. *British Journal of Cancer* 1999, **80**, 269–272.
8. The American Society of Clinical Oncology. Outcomes of cancer treatment for technology assessment and cancer treatment guidelines. *Journal of Clinical Oncology* 1996,**14**,671–9.
9. Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, *et al*. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality of life instrument for use in international clinical trials in oncology. *Journal of the National Cancer Institute* 1993, **85**, 365–376.
10. Cella DF, Tulsky DS, Gray G, Sarafian B, Linn E, Bonomi A, *et al*. The Functional Assessment of Cancer Therapy Scale: development and validation of the general measure. *Journal of Clinical Oncology* 1993, **11**, 570–579.
11. Vickery CW, Blazeby JM, Conroy T, Arraras J, Sezer O, Koller M, *et al*. Development of an EORTC disease specific quality of life module for use in patients with gastric cancer. *European Journal of Cancer* 2001, **37**, 966–971.
12. Nunnally JC. *Psychometric Theory*. New York, McGraw-Hill, 1978.
13. Zieren HU, Zippel K, Zieren J, Muller JM. Quality of life after surgical treatment of gastric carcinoma. *European Journal of Surgery* 1998, **164**, 119–125.
14. Webb A, Cunningham D, Scarffe JH, Norman A, Joffe JK, Hughes M, *et al*. Randomised trial comparing epirubicin, cisplatin and fluorouracil versus fluorouracil, doxorubicin and methrotrexate in advanced oesophagogastric cancer. *Journal of Clinical Oncology* 1997, **15**, 261–267.
15. Thybusch-Bernhardt A, Schmidt C, Kuchler T, Schmid A, Henne-Bruns D, Kremer B. Quality of life following radical surgical treatment of gastric carcinoma. *World Journal of Surgery* 1999, **233**, 503–598.
16. Wadler S, Brain C, Catalano P, Einzig AI, Cella D, Benson III AB. Randomized phase II trial of either fluorouracil, parenteral hydroxyurea, interferon-alpha-2a, and filgrastim or doxorubicin/docetaxel in patients with advanced gastric cancer with quality-of-life assessment: eastern cooperative oncology group study E6296. *Cancer Journal* 2002, **8**, 282–286.
17. Davies J, Johnston D, Sue-Ling HM, Young S, May J, Griffith J. Total or subtotal gastrectomy for gastric carcinoma? A study of quality of life. *World Journal of Surgery* 1998, **22**, 1048–1055.
18. Wu CW, Lo SS, Shen KH, Hsieh MC, Lui WY, P'eng FK. Surgical mortality, survival, and quality of life after resection for gastric cancer in the elderly. *World Journal of Surgery* 2000, **24**, 465–472.
19. Eremenco S, Cashy J, Webster K. Linguistic validation of the fact-gastric (FACT-Ga) in Japanese and English. *Quality of Life Research* 2004, **12**, 818.
20. Blazeby JM, Conroy T, Hammerlid E, Fayers P, Sezer O, Koller M, *et al*. Clinical and psychometric validation of an EORTC questionnaire module, the EORTC QLQ-OES18, to assess quality of life in patients with oesophageal cancer. *European Journal of Cancer* 2003, **39**, 1384–1394.