available at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejconline.com

Clinical and psychometric validation of a questionnaire module, the EORTC QLQ-OG25, to assess health-related quality of life in patients with cancer of the oesophagus, the oesophago–gastric junction and the stomach

Pernilla Lagergren[a,b], Peter Fayers[c,d], Thierry Conroy[e], Hubert J. Stein[f], Orhan Sezer[g], Richard Hardwick[h], Eva Hammerlid[i], Andrew Bottomley[j], Eric Van Cutsem[k], Jane M. Blazeby[a,l,], on behalf of the European Organisation for Research Treatment of Cancer Gastrointestinal and Quality of Life Groups*

[a]*Departments of Social Medicine and Clinical Sciences at South Bristol, University of Bristol, Bristol, United Kingdom*
[b]*Unit of Esophageal and Gastric Research (ESOGAR), Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden*
[c]*Department of Public Health, University of Aberdeen, Aberdeen, United Kingdom*
[d]*Department of Cancer Research and Molecular Medicine, Norwegian University of Technology and Science, Trondheim, Norway*
[e]*Department of Medical Oncology, Centre Alexis Vautrin, Vandoeuvre-Lès-Nancy, France*
[f]*Department of Surgery, Parscelsus Medical University, Salzburg, Austria*
[g]*Charité – Universitaetsmedizin Berlin, Berlin, Germany*
[h]*Addenbrookes NHS Hospital, Cambridge, United Kingdom*
[i]*Department of Otorhinolaryngology Head and Neck Surgery, Sahlgrenska University Hospital, Gothenburg, Sweden*
[j]*Quality of Life Unit, European Organisation for Research and Treatment of Cancer, Brussels, Belgium*
[k]*Department of Internal Medicine, University Hospital Gasthuisberg, Leuven, Belgium*
[l]*United Bristol Healthcare Trust, Bristol, UK*

ARTICLE INFO

Article history:
Received 1 July 2007
Accepted 4 July 2007
Available online 15 August 2007

Keywords:
European Organisation for Research and Treatment of Cancer
EORTC QLQ-C30
EORTC QLQ-OES18
EORTC QLQ-STO22
Questionnaires

ABSTRACT

Aim: To combine and test the EORTC questionnaires for assessing quality of life (HRQL) for oesophageal (QLQ-OES18) and stomach cancer (QLQ-STO22), into a single questionnaire for tumours of the oesophagus, oesophago–gastric junction or stomach.

Methods: The QLQ-OES18, QLQ-STO22 and seven modified items were administered to 300 patients with oesophageal ($n = 148$), junctional ($n = 66$), or gastric cancer ($n = 86$). Semi-structured interviews assessed item and scale preference and multi-trait scaling analyses confirmed the scale structure of the new module (QLQ-OG25). This was further tested for validity.

Results: The QLQ-OG25 has six scales, dysphagia, eating restrictions, reflux, odynophagia, pain and anxiety. Scales have good reliability (α range 0.67–0.87) and they distinguish between tumour sites and disease stage. Scales do not correlate highly with scores from the core questionnaire, thus indicating that the module was addressing separate HRQL aspects.

* *Corresponding author:* Present address: Clinical Sciences at South Bristol, Bristol Royal Infirmary, Marlborough St., Bristol BS2 8HW, United Kingdom. Tel.: +44 0 117 928 7231; fax: +44 0 117 928 7292.
E-mail address: j.m.blazeby@bristol.ac.uk (J.M. Blazeby).
0959-8049/$ - see front matter
doi:10.1016/j.ejca.2007.07.005

Conclusion: The QLQ-OG25 is recommended to supplement the EORTC QLQ-C30 when assessing HRQL in patients with oesophageal, junctional or gastric cancer.

1. Introduction

The epidemiology of oesophageal and gastric tumours has shown marked changes over the past two decades.[1] There are increasing numbers of distal oesophageal adenocarcinomas and cancers of the proximal stomach. Treatment for oesophago–gastric junctional tumours is difficult, and five year survival may only be achieved in 40%.[2] Important outcomes for patients with upper gastrointestinal cancers, therefore, include survival, treatment-related morbidity and toxicity, and health-related quality of life (HRQL). Accurate and valid assessment of HRQL may be performed using generic cancer measures such as the FACT-G or EORTC QLQ-C30, and both have supplementary site specific modules for oesophageal and gastric cancers.[3–8] It is unclear, however, which questionnaire module is most appropriate for patients with oesophago–gastric junctional tumours. The aim of this study, therefore, was to improve assessment of HRQL in patients with upper gastrointestinal cancer by producing a single EORTC questionnaire module to measure HRQL in patients with oesophageal or gastric cancer, including tumours of the oesophago–gastric junction.

2. Patients and methods

2.1. Patients

Prospectively recruited patients were entered into this study between August 2005 and August 2006. Patients were eligible if they had a histological diagnosis of oesophageal or gastric cancer, including tumours of the oesophago–gastric junction. The definition of oesophageal cancer was adenocarcinoma or squamous cell carcinoma of the oesophagus. Tumours of the oesophago–gastric junction included Siewert I–III.[9] The definition of gastric cancer was adenocarcinoma of the stomach, and included all tumours of the fundus, body and antrum and linitis plastica. For this study, a heterogeneous sample of patients was required including patients undergoing a variety of treatments and at different stages of their treatment. Patients were staged and selected for treatment according to local policies and classified into three groups: (1) oesophageal, (2) oesophago–gastric junction, and (3) stomach. Patients were excluded if they had concurrent malignancies, except for basal cell skin cancer, or if they were unable to understand and complete the questionnaire. Patients with oesophageal small cell carcinoma and patients with gastric lymphoma or maltoma were excluded. There were no restrictions for either age or performance status. Written informed consent was obtained from all enrolled patients and local or national ethical approval was acquired. Patients were recruited from two hospitals in France, Germany and the United Kingdom, and from one hospital in Sweden.

2.2. Data collection and questionnaires

Sociodemographic and clinical data were collected according to a written study protocol by extraction from the case notes, or obtained from the patients directly.

The methods used to combine the existing validated modules are summarised in Table 1. All patients completed the EORTC QLQ-C30 (version 3.0),[6] the provisional combined module including items from the EORTC QLQ-OES18 oesophageal module[7] and the EORTC QLQ-STO22 stomach module,[8] and seven additional modified items. The latter included a proposed new three-item dysphagia scale, three eating items and a modified reflux item. All seven items addressed issues already in both modules, but with slightly different wording. The structure of the QLQ-C30 has been reported in detail elsewhere and its validity and reliability has been examined in several diagnostic groups, including in cancer of the oesophagus.[10] The QLQ-OES18 module contains 18 items and incorporates four symptom scales measuring dysphagia, eating restrictions, reflux and pain and six single items measuring dry mouth, taste problems, difficulties in speaking, choking, coughing and difficulties swallowing saliva. The QLQ-STO22 includes similar scales assessing dysphagia, eating restrictions, pain and reflux and single items for dry mouth and taste, but it has an additional scale addressing anxiety and single items for body image and the impact of hair loss. All multi-item scales and single items were scored on a one to four-point Likert scale ('not at all', 'a little', 'quite a bit', 'very much').

During the completion of the provisional combined module, a semi-structured interview explored patients' opinions and preferences for all items and scales. The questionnaire items and the interview were performed in the patients' first language. For example, after the patients had completed the dysphagia scales (QLQ-OES18, QLQ-STO22 and new scale), they were asked which scale they would retain in the updated questionnaire and why they selected this option. Observer records of dysphagia grade were also kept. Where modified items were added to the eating and reflux scale, patients were asked to identify which items they would select preferentially for the final questionnaire and reasons for their choices were recorded. Patients also commented whether to include or remove single items.

2.3. Statistical analysis

2.3.1. Defining HRQL scales and items

Multi-trait scaling analysis was used to examine whether the individual items of the combined oesophageal and gastric questionnaires and the seven additional items could be aggregated as hypothesized into a more limited set of multi-item scales. This technique, based on an examination of item-scale correlations, has been used extensively in evaluating the

Table 1 – Phases of EORTC Quality of Life Group module development and method for combining existing questionnaires

Phase	Standard module development	Modified process for combining two EORTC modules[a]
I	Generation of quality of life issues • Literature search • Interviews with patients • Interviews with healthcare professionals	Generation of 'provisional combined module' • Combine two modules and remove one of each identical item • Add 'new' items only to improve wording of existing items • Patients complete the 'temporary combined module' • Interview patients to understand item preference
	A list of relevant issues	
II	Construction of a provisional module • Items identified from QOL issues • Items selected from EORTC item bank • New items created in EORTC format	Define scale structure of final module • Use questionnaire responses to define scales with multi-trait scaling analyses • Remove items not meeting scaling criteria • Retain all single items from both modules
	Review by QOL Group Module Development Committee	Review by EORTC Quality of Life GI Group
III	Pre-testing of new module • Patients complete module and are interviewed • Responses prevalence and variance analysed • Site specific module development review group results	Test psychometric properties of new scale structure • Known group comparisons • Associations with QLQ-C30
	Review by QOL Group Module Development Committee	Review of QOL Group Module Development Committee

a This modified process is proposed by the senior author of the module and has not yet been adopted by the EORTC Quality of Life Group.

Table 2 – Socio-demographic and clinical details of 300 patients with cancer of the oesophagus, the oesophageal-gastric junction (OGJ) or the stomach

	All, *n* = 300	Oesophagus, *n* = 148	OGJ, *n* = 66	Stomach, *n* = 86
Age, median (range)	63 (27–85)	64 (27–85)	63 (30–83)	62 (34–84)
Gender, male (%)	209 (69)	109 (74)	54 (82)	46 (53)
Co-habitants (%)				
Alone	58 (19)	29 (20)	13 (20)	16 (19)
With other adults	223 (74)	113 (76)	46 (70)	64 (74)
Unknown	14 (5)	5 (3)	6 (9)	3 (3)
Marital status (%)				
Single	19 (6)	10 (7)	5 (8)	4 (5)
Married	205 (68)	102 (69)	41 (62)	62 (72)
Separated, divorced, widowed	60 (20)	28 (19)	15 (23)	17 (20)
Unknown	16 (5)	8 (5)	5 (8)	3 (3)
Education (%)				
Less than compulsory school	12 (4)	10 (7)	1 (2)	1 (1)
Compulsory school	157 (52)	79 (53)	37 (56)	41 (48)
Post compulsory school	94 (31)	40 (27)	21 (32)	33 (38)
University	30 (10)	14 (9)	6 (9)	10 (12)
Employment (%)				
Full-time or part-time	107 (36)	53 (36)	28 (42)	26 (30)
Homemaker	15 (5)	3 (2)	3 (5)	9 (10)
Unemployed	10 (3)	4 (3)	2 (3)	4 (5)
Retired	163 (54)	85 (57)	33 (50)	45 (52)
Student	1 (1)	0 (0)	0 (0)	1 (1)
Advanced disease, yes (%)	85 (28)	36 (24)	16 (24)	33 (38)
Started treatment, yes (%)	189 (63)	98 (66)	39 (59)	52 (60)
Place of questionnaire completion (%)				
At home	12 (4)	8 (5)	2 (3)	2 (2)
In out patient clinic	46 (15)	18 (12)	12 (18)	16 (19)
In hospital (as resident)	236 (79)	118 (80)	50 (76)	68 (79)
Country (%)				
France	68 (23)	48 (32)	2 (3)	18 (21)
Germany	129 (43)	43 (29)	36 (55)	50 (58)
Sweden	25 (8)	17 (11)	6 (9)	2 (2)
United Kingdom	78 (26)	40 (27)	22 (33)	16 (19)

structure of health status measures.[11] The combined hypothesized scales from the original two modules were used to start the analyses. Evidence of item convergent validity is defined as a correlation of 0.40 or greater between an item and its own scale (corrected for overlap).[11] A scaling success for an item is seen when the correlation between an item and its own scale is significantly higher than its correlation with other scales. Items that consistently correlated more highly with another scale than the hypothesized one were considered as scaling error and revisions were made. After revising the scales, the psychometric properties of the new scales in the module were re-explored by use of the revised correlations between items and scales. Thereafter, analyses of internal consistency (reliability) and clinical and construct validity were performed.

2.3.2. *Reliability*

The reliability (internal consistency) of the multi-item questionnaire scales was assessed by Cronbach's alpha coefficient. As recommended by Fayers and Machin,[11] internal consistency estimates of a magnitude of >0.70 are considered acceptable for group comparisons.

2.3.3. *Validity*

Correlations between the QLQ-C30 and the new module, QLQ-OG25, were examined using Pearson's product moment correlation. It was expected that those scales that are conceptually related (e.g., dysphagia and eating restrictions) would correlate substantially with one another (i.e. Pearson's $r > 0.40$). Conversely, those scales with less in common (e.g. specific emotional problems and upper gastrointestinal symptoms) were expected to exhibit lower correlations (i.e. Pearson's $r < 0.40$).[11]

2.3.4. *Clinical validity*

Clinical validity and sensitivity were examined by known-group comparisons, i.e. to explore the extent to which the questionnaire scores are able to discriminate between groups of patients differing in clinical status.[11] Known groups used for these comparisons were disease stage (advanced disease versus local/locally advanced disease), physical function scores (poor physical function versus good physical function), and disease site (oesophageal, oesophago–gastric junction and stomach). For comparisons of disease stage and physical function the *t*-test was used, while one-way analysis of variance (ANOVA) was used to compare disease sites.

2.3.5. *Sample size*

The sample size was planned before the start of the study and was based on the recommendation of Tabachnik and Fidell[12] that for multivariable analysis techniques to obtain reliable estimates, the number of observations should be 5–10 patients times the number of variables in the model. The study was closed after inclusion of 300 patients.

2.3.6. *HRQL scores*

All scale and item scores from the QLQ-C30 and the QLQ-OG25 were linearly transformed to a 0–100 score according to the EORTC QLQ-C30 scoring manual.[13] Mean scores with standard deviation were calculated. High scores in the multi-item function scales and the global quality of life scale indicate better levels of function and quality of life, respectively

Table 3 – Reliability of the multi to item scales in the QLQ-OG25 using multi to trait scaling analysis presented for the total group and divided by disease site

QLQ-OG25	All patients, $n = 292$			Oesophagus, $n = 142$			Oesophago-gastric junction, $n = 64$			Stomach, $n = 86$		
	Item correlation within scale[a]	Item correlation with other scales	α	Item correlation within scale[a]	Item correlation with other scales	α	Item correlation within scale[a]	Item correlation with other scales	α	Item correlation within scale[a]	Item correlation with other scales	α
OGDYS	0.55–0.77	0.05–0.55	0.79	0.48–0.74	0.11–0.55	0.76	0.62–0.89	–0.10 to 0.66	0.86	0.62–0.72	–0.04 to 0.54	0.81
OGEAT	0.44–0.70	0.03–0.64	0.81	0.48–0.65	–0.02 to 0.61	0.80	0.44–0.74	–0.13 to 0.52	0.81	0.32 to 0.77	0.14–0.65	0.82
OGRFX	0.51–0.51	0.14–0.31	0.67	0.47–0.47	0.06–0.38	0.64	0.58–0.58	0.11–0.31	0.73	0.51–0.51	0.09–0.54	0.67
OGODYN	0.64–0.64	0.22–0.62	0.78	0.59–0.59	0.22–0.59	0.74	0.60–0.60	0.12–0.67	0.75	0.77–0.77	0.03–0.64	0.86
OGP& D	0.78–0.78	0.17–0.49	0.87	0.69–0.69	0.15–0.53	0.81	0.81–0.81	–0.04 to 0.52	0.89	0.85–0.85	0.02–0.53	0.91
OGANX	0.75–0.75	0.09–0.25	0.86	0.77–0.77	0.18–0.42	0.87	0.75–0.75	–0.05 to 0.23	0.85	0.72–0.72	–0.01 to 0.20	0.84

OGDYS, dysphagia; OGEAT, eating restrictions; OGRFX, reflux; OGODYN, odynophagia; OGP and D, pain and discomfort; OGANX, anxiety.
α, Cronbach's alpha.
a Corrected for overlap

(QLQ-C30), whereas high scores in the symptom scales and items represent more symptoms (QLQ-C30, QLQ-OG25).

All analyses were performed using the statistical software Stata 9.2 (Stata Corporation, College Station, TX, USA, 2005). To minimize the risk of multiple testing, a p value of <0.01 was regarded as statistically significant.

3. Results

3.1. Patient characteristics

A total of 303 patients were recruited from seven institutions in four different countries. Of these, two were excluded because of previous cancer and one was excluded due to incomplete data resulting in 300 patients for the final analyses. In Table 2, socio-demographic data and clinical details are presented. Most of the patients had started treatment (63%) including patients who had already finished their treatment, and had local or locally advanced disease (72%).

3.2. Defining scales and items

Both the multi-trait scaling analyses and patient interviews showed that the dysphagia scale from the QLQ-STO22 was the preferred scale for assessing swallowing problems. Most patients (43%) favoured this scale compared to the scale from the QLQ-OES18 (32%) and the QLQ-STO22 (25%) because of ease of understanding, and observers noted that patients completed this scale correctly. The combined eating items from the QLQ-OES18, the QLQ-STO22 and the new items showed poor scale reliability. Three overlapping and similar items were therefore removed and a single item addressing difficulties eating in front of other people was separated as a single item, and a new four-item scale assessing eating restrictions was created with adequate scaling properties (Table 3).

The four-item reflux scale was reduced to two items by removing the belching item which did not fit in the scale and removing the additional acid reflux item that did

Table 4 – Mean scores with standard deviation (SD) of scales and items in the QLQ-C30 and the QLQ-OG25 for clinically distinct groups of patients

	Advanced disease[a]			Physical function[b]		
	Mean scores (SD)		p-value[c]	Mean scores (SD)		p-value[c]
	No, n = 187	Yes, n = 85		Good, n = 158	Poor, n = 142	
QLQ-C30						
PF	82 (21)	72 (29)	0.00	–	–	–
RF	71 (34)	59 (37)	0.01	87 (36)	45 (36)	0.00
EF	73 (24)	68 (27)	0.15	79 (25)	64 (25)	0.00
CF	84 (19)	80 (25)	0.14	89 (22)	76 (22)	0.00
SF	75 (29)	63 (35)	0.00	83 (32)	57 (32)	0.00
QOL	65 (24)	53 (24)	0.00	71 (24)	50 (24)	0.00
PA	23 (26)	31 (31)	0.02	15 (28)	36 (28)	0.00
FA	34 (34)	47 (30)	0.00	21 (29)	57 (29)	0.00
NV	12 (21)	23 (27)	0.00	9 (24)	23 (24)	0.00
QLQ-OG25						
OGDYS	26 (30)	27 (29)	0.72	22 (29)	31 (29)	0.01
OGEAT	38 (30)	43 (32)	0.20	31 (29)	50 (29)	0.00
OGRFX	18 (26)	18 (22)	0.89	13 (25)	24 (25)	0.00
OGODYN	27 (30)	25 (30)	0.72	22 (30)	30 (30)	0.02
OGP& D	20 (28)	28 (27)	0.04	16 (29)	29 (29)	0.00
OGANX	62 (33)	59 (35)	0.52	55 (34)	67 (34)	0.00
OGEO	17 (31)	18 (32)	0.78	16 (31)	19 (32)	0.44
OGDM	32 (36)	44 (34)	0.01	27 (36)	45 (36)	0.00
OGTA	17 (27)	24 (35)	0.08	10 (30)	29 (30)	0.00
OGBI	18 (30)	29 (36)	0.01	14 (32)	30 (32)	0.00
OGSV	14 (28)	11 (26)	0.48	10 (26)	15 (26)	0.11
OGCH	13 (26)	14 (26)	0.90	11 (26)	17 (26)	0.04
OGCO	12 (25)	11 (24)	0.79	5 (26)	20 (26)	0.00
OGSP	9 (20)	17 (32)	<0.01	8 (24)	14 (24)	0.04
OGWL	30 (36)	42 (38)	0.01	26 (38)	42 (38)	0.00
OGHAIR	7 (21)	8 (20)	0.75	4 (20)	9 (20)	0.04

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): PA, pain; FA, fatigue; NV, nausea and vomiting, QLQ-OG25 scales and items (high scores = more problems): OGDYS, dysphagia; OGEAT, eating restrictions; OGRFX, reflux; OGODYN, odynophagia; OGP & D, pain and discomfort; OGANX, anxiety; OGEO, eating with others ; OGDM, dry mouth; OGTA, sense of taste; OGBI, body image; OGSV, saliva; OGCH, choking; OGCO, cough; OGSP, speech; OGWL, weight loss; OGHAIR, hair loss.

a Patients with M1 and/or disease recurrence, 29 patients with missing data.

b Poor physical function was defined as less than median (<80) and good as more than median (⩾80) on the QLQ-C30 physical function scale.

c t-test was used to determine if differences were statistically significant at the 1% level.

not add any precision to the original scale from the QLQ-OES18.

The combined pain items from the QLQ-OES18 and QLQ-STO22 pain scales demonstrated poor validity as a single large scale. The scale was therefore separated into two scales to address pain and discomfort when eating (odynophagia) and pain or discomfort in the stomach area.

Only the gastric module included items about anxiety and when this 3-item scale was tested in this patient population, item correlations were poor, therefore the item addressing worries about weight loss was separated as a single item and based on the patients comments of keeping items about anxiety, a two-item anxiety scale was created. This scale was retained in the final module because patients felt it was very important in addition to the emotional function scale in the QLQ-C30.

Although details are not reported here, exploratory analysis with item response theory (IRT) was used to confirm that the most useful (most discriminatory and most informative) items were retained. Differential item functioning (DIF), also unreported in detail, was used to confirm that, apart from the pain items mentioned in the text, all items in the multi-item scales performed equally well and consistently across the three sites.

So the final module, therefore, has six scales; dysphagia (3 items); eating restrictions (4 items); reflux (2 items); odynophagia (2 items); pain and discomfort (2 items); anxiety (2 items), and 10 single items; eating with others; dry mouth; sense of taste; body image; saliva; choking; cough; speech; weight loss; hair loss. Results from the multi-trait scaling analyses of the resultant QLQ-OG25 are presented in Table 3. The remaining results in this paper use the QLQ-OG25 scale and item structure.

3.3. *Validity*

Relations between the QLQ-OG25 and the core questionnaire QLQ-C30 were examined for all patients. Most scales in the QLQ-OG25 had low correlations with the core questionnaire, reflecting that the module is assessing clinically distinct HRQL issues. Anticipated exceptions were associations between the QLQ-OG25 pain and discomfort scale and the odynophagia scale that were moderately correlated with the QLQ-C30 pain scale ($r = 0.58$ and 0.43, respectively). The QLQ-OG25 anxiety scale was also moderately associated with the emotional function scale of the QLQ-C30 ($r = 0.60$). There were weak correlations between the body image and social function scales ($r = 0.42$), and between the eating restriction scale of the QLQ-OG25 and the global quality of life scale, role function, pain, fatigue, and nausea or vomiting scales of the QLQ-C30 ($r = 0.43$–0.47). These correlations demonstrate clinical overlap between the subscales and were expected.

3.4. *Clinical validity*

Patients with advanced disease reported significantly worse functional aspects of HRQL and more symptoms than patients with localised disease (Table 4). Patients with advanced disease also reported more problems with dry mouth, body image, speech and weight loss (Table 4). A similar pattern was observed among patients with poor physical function who reported significantly worse scores in all scale items in the QLQ-C30 and QLQ-OG25 modules compared to patients with good physical function (Table 4). Statistically significant differences between patients with oesophageal, junction and stomach tumours were seen in three scales from the QLQ-OG25 (pain and discomfort, problems with saliva and problems with speech) (Table 5). Fig. 1 summarises the results.

Table 5 – Mean scores with standard deviation (SD) of scales and items in the QLQ-C30 and the QLQ-OG25 for patients with cancer of the oesophagus, the oesophago–gastric junction (OGJ) or the stomach

	Mean scores (SD)			*p*-value[a]
	Oesophagus *n* = 148	OGJ *n* = 66	Stomach *n* = 86	
QLQ-C30				
PF	79 (23)	80 (23)	74 (27)	0.19
RF	71 (34)	65 (37)	64 (37)	0.26
EF	73 (25)	71 (22)	71 (27)	0.80
CF	81 (23)	86 (19)	82 (21)	0.33
SF	74 (28)	73 (30)	62 (36)	0.01
QOL	63 (23)	63 (24)	57 (25)	0.16
PA	24 (27)	22 (26)	30 (31)	0.13
FA	36 (29)	38 (29)	42 (30)	0.35
NV	14 (20)	16 (25)	19 (27)	0.35
QLQ-OG25				
OGDYS	29 (29)	28 (31)	19 (27)	0.02
OGEAT	40 (30)	43 (31)	39 (32)	0.71
OGRFX	17 (24)	18 (27)	21 (26)	0.45
OGODYN	28 (31)	31 (31)	19 (26)	0.04
OGP& D	17 (25)	28 (30)	28 (32)	0.00
OGANX	56 (36)	67 (31)	64 (32)	0.07
OGEO	21 (34)	20 (32)	9 (26)	0.01
OGDM	35 (36)	33 (35)	37 (36)	0.75
OGTA	17 (27)	22 (33)	21 (33)	0.43
OGBI	18 (30)	23 (32)	26 (33)	0.19
OGSV	16 (30)	14 (26)	5 (17)	0.01
OGCH	16 (29)	14 (27)	9 (20)	0.16
OGCO	15 (29)	11 (24)	9 (22)	0.27
OGSP	15 (28)	7 (21)	6 (19)	<0.01
OGWL	34 (38)	32 (35)	34 (38)	0.94
OGHAIR	21 (34)	14 (26)	23 (29)	0.85

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): PA, pain; FA, fatigue; NV, nausea and vomiting, QLQ-OG25 scales and items (high scores = more problems): OGDYS, dysphagia; OGEAT, eating restrictions; OGRFX, reflux; OGODYN, odynophagia; OGP and D, pain and discomfort; OGANX, anxiety; OGEO, eating with others ; OGDM, dry mouth; OGTA, sense of taste; OGBI, body image; OGSV, saliva; OGCH, choking; OGCO, cough; OGSP, speech; OGWL, weight loss; OGHAIR, hair loss.

a ANOVA was used to determine if differences were statistically significant at the 1% level.

4. Discussion

This international study indicates that the new module, the EORTC QLQ-OG25, is a reliable and valid questionnaire to sup-

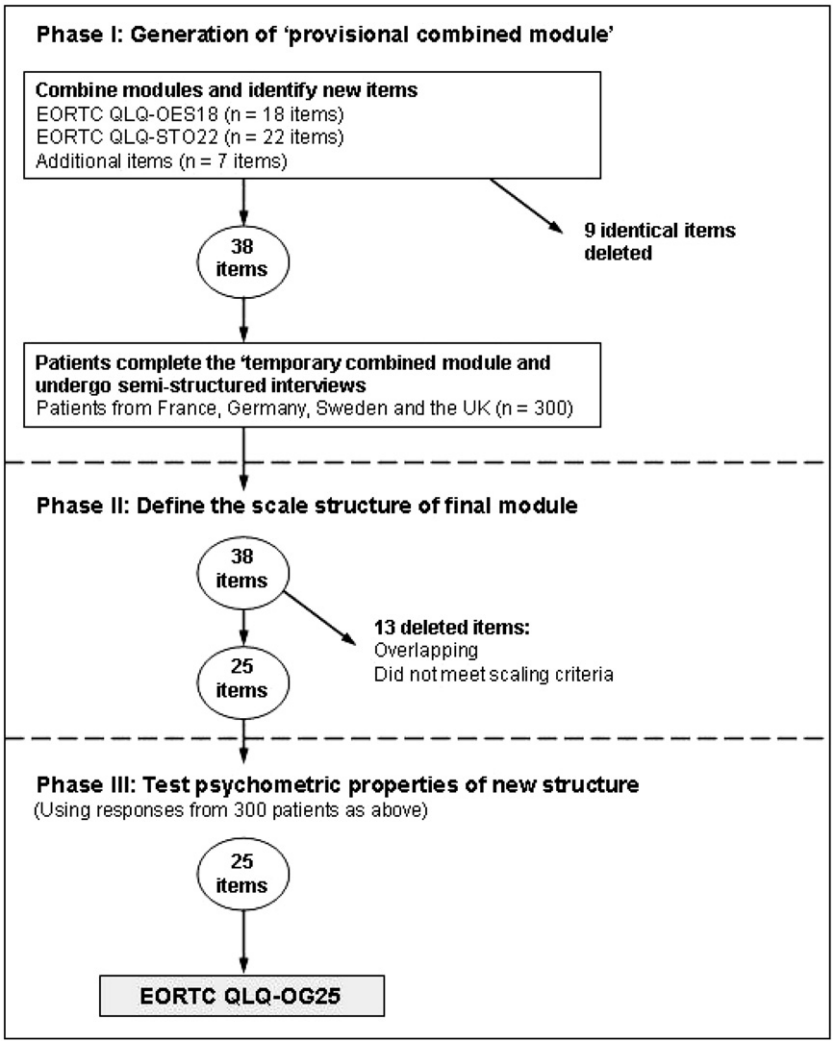

Fig. 1 – Flowchart of the QLQ-OG25 development.

plement the EORTC QLQ-C30 to assess HRQL among patients with cancer of the oesophagus, oesophago–gastric junction or the stomach.

Measurement of HRQL in patients with oesophageal or gastric cancer, including junctional tumours, has become an important outcome in clinical trials, clinical practice and in longitudinal studies. Using the core instrument remains important because it will detect changes in general aspects of HRQL such as deteriorations in the social and role function in advanced disease. To detect HRQL issues of importance to patients within specific treatment groups, however, it is necessary to use a cancer-specific questionnaire with a site-specific module because patients with upper gastrointestinal cancer often have major problems related to eating and drinking. These issues are not addressed in the generic measures of HRQL, and might otherwise be overlooked. Indeed several studies have shown that although general aspects of HRQL may deteriorate after treatment, patients simultaneously report improvements or relief of site specific issues, such as relief of dysphagia.[14,15] Evidence for this was also supported in this study as few associations between scales in the QLQ-C30 and the QLQ-OG25 were demonstrated. This is probably due to that the core questionnaire and the module assess different aspects of HRQL. Using a modular approach for assessment of HRQL is therefore essential to detect both benefits and disadvantages of the treatment.

The QLQ-OG25 has six symptom scales (dysphagia, eating restrictions, reflux, odynophagia, pain and discomfort, and a scale assessing anxiety) and 10 single items relevant to patients undergoing palliative or potentially curative treatments and follow-up for upper gastrointestinal cancer. The choice of scales and items was based on multi-trait scaling analyses, reliability and validity test and patients opinions through the semi-structured interview. For example, most patients preferred the dysphagia scale from the QLQ-STO22 module because the items were easy to understand, and for the eating scale, patients reported that some of the items were overlapping.

Although this was a large prospective international study with excellent patient compliance, the numbers of patients within each country are insufficient for exploration of any

cross-cultural differences. Further tests are also needed to confirm the anticipated sensitivity to changes over time. Finally, this module has also not yet been tested in patients with benign oesophageal or gastric disease.

In conclusion, the new module, EORTC QLQ-OG25, together with the core questionnaire, QLQ-C30, can be recommended to measure HRQL in patients with cancer of the oesophagus, oesophago–gastric junction or the stomach undergoing curative or palliative treatment or follow-up. The module is now available from the Quality of Life Unit, at the EORTC data centre (www.eortc.be/home/qol/) and is already available in eight major European languages following rigorous translation procedures.[16]

Conflicts of interest statement

None declared.

Acknowledgement

P.L. was supported by the Karolinska Institutet (Center for Health Care Sciences), the Swedish Society for Medical Research and the Swedish Cancer Society. O.S. was supported by a research grant from Deutsche Krebshilfe, Bonn, Germany. J.M.B. was supported by the UK Medical Research Council Clinician Scientist Award. We thank the Upper GI multi-disciplinary team at United Bristol Healthcare Trust, Bristol UK for allowing us to study patients under their care. We thank Joanna Nicklin and Linda Bycroft for interviewing patients in the UK and Isabella Leonard for interviewing patients in France.

REFERENCES

1. Devesa SS, Blot WJ, Fraumeni Jr JF. Changing patterns in the incidence of esophageal and gastric carcinoma in the United States. *Cancer* 1998;**83**(10):2049–53.
2. Lerut T, Coosemans W, Decker G, De Leyn P, Nafteux P, Van Raemdonck D. Cancer of the esophagus and gastro-esophageal junction: potentially curative therapies. *Surg Oncol* 2001;**10**(3):113–22.
3. Webster K, Cella D, Yost K. The functional assessment of chronic illness therapy (FACIT) measurement system: properties, applications, and interpretation. *Health Qual Life Outcomes* 2003;**1**(1):79.
4. Darling G, Eton DT, Sulman J, Casson AG, Celia D. Validation of the functional assessment of cancer therapy esophageal cancer subscale. *Cancer* 2006;**107**(4):854–63.
5. Eremenco SL, Cashy J, Webster K, et al. FACT-gastric: a new international measure of QOL in gastric cancer. *J Clin Oncol* 2004;**22**(14S):8123 [2004 ASCO annual meeting proceedings (Post-Meeting Edition)].
6. Aaronson NK, Ahmedzai S, Bergman B, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst* 1993;**85**(5):365–76.
7. Blazeby JM, Conroy T, Hammerlid E, 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. *Eur J Cancer* 2003;**39**(10):1384–94.
8. Blazeby JM, Conroy T, Bottomley A, et al. Clinical and psychometric validation of a questionnaire module, the EORTC QLQ-STO 22, to assess quality of life in patients with gastric cancer. *Eur J Cancer* 2004;**40**(15):2260–8.
9. Siewert JR, Stein HJ. Classification of adenocarcinoma of the oesophagogastric junction. *Br J Surg* 1998;**85**(11):1457–9.
10. Blazeby JM, Williams MH, Brookes ST, Alderson D, Farndon JR. Quality of life measurement in patients with oesophageal cancer. *Gut* 1995;**37**(4):505–8.
11. Fayers P, Machin D. *Quality of life: the assessment analysis and interpretation of patient reported outcomes*. second ed. Wiley: Chichester; 2007.
12. Tabachnik BJ, Fidel LS. *Using multivariate statistics*. London: Harper & Row; 1993.
13. Fayers PM, Aaronson NK, Bjordal K, Groenvold M, Curran D, Bottomley Aon behalf of the EORTC Quality of Life Study Group. *The EORTC QLQ-C30 scoring manual*. third ed. Brussels: European Organisation for Research and Treatment of Cancer; 2001.
14. Blazeby JM, Farndon JR, Donovan J, Alderson D. A prospective longitudinal study examining the quality of life of patients with esophageal carcinoma. *Cancer* 2000;**88**(8):1781–7.
15. Reynolds JV, McLaughlin R, Moore J, Rowley S, Ravi N, Byrne PJ. Prospective evaluation of quality of life in patients with localized oesophageal cancer treated by multimodality therapy or surgery alone. *Br J Surg* 2006;**93**(9):1084–90.
16. Cull A, Sprangers M, Bjordal K, Aaronson N, West K, Bottomley Aon behalf of the EORTC Quality of Life Study Group. *Quality of life group translation procedure*. second ed. Brussels: European Organisation for Research and Treatment of Cancer; 2002.