available at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejconline.com

An international field study of the EORTC QLQ-PR25: A questionnaire for assessing the health-related quality of life of patients with prostate cancer

George van Andel[a], Andrew Bottomley[b], Sophie D. Fosså[c], Fabio Efficace[b], Corneel Coens[b], Stephane Guerif[d], Howard Kynaston[e], Paolo Gontero[f], George Thalmann[g], Atif Akdas[h], Sven D'Haese[b], Neil K. Aaronson[i,], on behalf of the EORTC Genito-Urinary Tract Cancer Group and the EORTC Quality of Life Group*

[a]*Department of Urology, Onze Lieve Vrouwe Gasthuis, The Netherlands*
[b]*Quality of Life Department, European Organization for Research and Treatment of Cancer (EORTC) Headquarters, Brussels, Belgium*
[c]*Department of Oncology and Radiotherapy, Rikshospitalet, Oslo, Norway*
[d]*Service de Radiotherapie, Hopital Jean Bernard, Poitiers, France*
[e]*Department of Urology, University Hospital of Wales, Cardiff, United Kingdom*
[f]*Department of Urology, University of Studies of Torino, Torino, Italy*
[g]*Department of Urology, University of Bern, Anna Seiler Haus, Inselspital, Bern, Switzerland*
[h]*Department of Urology, Marmara University Hospital, Istanbul, Turkey*
[i]*Division of Psychosocial Research and Epidemiology, The Netherlands Cancer Institute – Antoni van Leeuwenhoek Hospital, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands*

ARTICLE INFO

Article history:
Received 14 July 2008
Accepted 17 July 2008
Available online 5 September 2008

Keywords:
Prostate cancer
Quality of life
Patient-reported outcome
Questionnaire
Bladder function
Urinary function
Sexuality

ABSTRACT

Aim: To evaluate the psychometrics of the EORTC QLQ-PR25, a questionnaire assessing the health-related quality of life of prostate cancer patients.
Methods: The QLQ-PR25 and the QLQ-C30 were administered to 642 prostate cancer patients from 13 countries treated with curative or palliative intent. The QLQ-PR25 assesses urinary, bowel and sexual symptoms and functioning, and the side-effects of hormonal treatment.
Results: Five hundred and nine patients were available for the final analysis. Multitrait scaling analyses confirmed the hypothesised scale structure of the QLQ-PR25. Internal consistency reliability was good (coefficient $\alpha = 0.70$–0.86) for the urinary symptoms and sexual function scales, but lower for the bowel function and side-effects of hormonal treatment scales ($\alpha < 0.70$). The module discriminated clearly between clinically distinct patient subgroups, and was responsive to changes in health status over time.
Conclusion: In general, the QLQ-PR25 demonstrates acceptable psychometric properties and clinical validity. Some caution should be used in interpreting the bowel function and side-effects of hormonal therapy scales; results can be reported at the individual item and scale level.

* *Corresponding author:* Tel.: +31 20 512 2481; fax: +31 20 512 2322.
E-mail address: n.aaronson@nki.nl (N.K. Aaronson).
0959-8049/$ - see front matter
doi:10.1016/j.ejca.2008.07.030

1. Introduction

Worldwide, after lung cancer, prostate cancer is the most commonly diagnosed malignant tumour in men, with an estimated 40 million affected individuals in North America, Europe and Japan.[1,2] Patients with localised prostate cancer without distant metastasis (T1–2, N0, M0) can be treated with curative intent with either surgery or radiotherapy, with or without hormonal therapy. To date, there is no convincing evidence demonstrating survival superiority of any one of these approaches to curative treatment.[3–5] Alternative disease management strategies include active surveillance and watchful waiting.[6]

Patients with locally advanced prostate cancer, with lymph node metastases or with distant (bone) metastases, cannot be cured. The first-line palliative treatment is hormonal. Patients with locally advanced prostate cancer typically receive external radiation combined with hormonal treatment.[4,7]

In choosing amongst the various treatment options, not only are the issues of disease-free and overall survival important, but also the effect of the disease and its treatment on patients' health-related quality of life (HRQOL). Due to the range of treatment options available, and the differential effects of various treatments on patients' symptoms and functional health, HRQOL considerations may play an even greater role in treatment decision-making than is the case for some other types of cancer.

The EORTC employs a modular approach to assessing cancer patients' HRQOL, including a core questionnaire, the QLQ-C30, intended for use across a wide spectrum of patient populations, and supplemental questionnaire modules, designed to assess HRQOL issues most relevant to a specific patient population.[8] The EORTC has established standardised procedures for the development of condition-specific HRQOL questionnaire modules.[9] Although other questionnaires are available for the assessment of prostate cancer patients' HRQOL, including the Functional Assessment of Cancer Therapy – Prostate (FACT-P),[10] the University of California, Los Angeles – Prostate Cancer Index (PCI)[11] and the Expanded Prostate Cancer Index Composite (EPIC),[12] it was considered desirable and appropriate to generate a questionnaire that is compatible and consistent with the measurement strategy employed within the EORTC. This paper reports the results of the international psychometric field testing of the prostate cancer-specific HRQOL questionnaire module, the QLQ-PR25.

2. Patients and methods

2.1. Patients

Patients were recruited from March 2002 to December 2004, with follow-up data collected through 2005. The study was coordinated at the Quality of Life Department of the EORTC Headquarters in Brussels, Belgium (EORTC Protocol 15011-30011). Patients were prospectively registered before treatment, and were eligible to participate if they had a histologically confirmed diagnosis of prostate cancer, had no previous treatment for their prostate cancer and did not have cerebral metastases or concurrent malignancies except basal cell carcinoma of the skin. Patients were excluded if they did not have basic fluency in one of the languages in which the questionnaire was available, suffered from serious cognitive or psychiatric problems or were participating in other HRQOL studies. No age limit was imposed. Local or national ethics committees approved the study, and all patients provided written informed consent.

Patients were selected for treatment according to local policies, and were entered into two pre-determined groups for the purpose of questionnaire validation. Group 1 consisted of patients selected for potentially curative treatment and staged as T1–3, G1–3, N0 and M0 (subgroups: A surgery, B external radiation therapy and C external radiation therapy combined with hormonal therapy). Group 2 consisted of patients selected for treatment with palliative intent and staged as T1–4, G1–3, N0–2 and M1. These patients were all hormonally treated.

2.2. Data collection schedule and content

HRQOL assessments were conducted at pre-specified time points. Group I patients, including all subgroups, completed the questionnaire twice: before the start of treatment (time window of 14 d before the start of treatment) and 3 months after the start of primary treatment (±14 d). Group II patients completed the questionnaire three times: before the start of treatment (time window of 14 d before the start of treatment) and three and 6 months after the start of primary treatment (±14 d).

Karnofsky performance status was rated by the treating clinician at the pre-treatment and on-treatment time points. Common toxicity criteria ratings were provided by the clinician at the time of the second and third questionnaire administration. Sociodemographic and other clinical information were recorded pre-treatment.

The QLQ-C30 (version 3.0), and the QLQ-PR25 were administered at all time points. Additionally, at the baseline assessment only, patients were asked to complete a short debriefing questionnaire covering questions about the time taken to complete the two questionnaires, the need for help in completing them and whether any of the items were confusing, difficult to answer or upsetting. Compliance with questionnaire completion and other general aspects of this field study were monitored using standard EORTC procedures, and were reviewed by the principal investigators every 6 months.

The EORTC QLQ-C30 contains scales and items addressing the functional aspects of HRQOL and symptoms that commonly occur in cancer patients. There are extensive reports on the validity and reliability of the QLQ-C30 used in a wide range of diagnostic and treatment settings.[8]

The QLQ-PR25 was developed as a joint project of the EORTC Quality of Life and Genitourinary Cancer Groups. During phase 1 of questionnaire development, 45 health care professionals, and 56 patients from nine countries with varying stages of disease were interviewed to identify the most prevalent and important symptoms and functional problems. In phases 2 and 3, a provisional module was developed and pilot

tested, amongst 41 patients from France, the Netherlands and the United Kingdom. The resulting questionnaire, the QLQ-PR25, consists of 25 items assessing urinary and bowel symptoms, sexual activity and functioning and side-effects of treatment. The module was translated into 14 languages according to the EORTC Quality of Life Group guidelines, including iterative forward–backward methods and pilot testing.[13] More detailed information pertaining to phase 1 through 3 developments of the QLQ-PR25 is available from the EORTC Quality of Life Department at the EORTC Headquarters (www.eortc.be).

2.3. Hypothesised scale structure of the QLQ-PR 25

The items of the QLQ-PR25 were hypothesised to cluster into five multi-item scales (one of which is conditional), and one single, conditional item: (1) urinary symptoms (URI; eight items); (2) bother due to the use of incontinence aid (AID, conditional on using an incontinence aid; single item); (3) bowel symptoms (BOW; four items); (4) hormonal treatment-related symptoms (HTR; six items); sexual activity (SAC; two items) and (5) sexual functioning (SFU, conditional on being sexually active; four items).

All items and scale scores of the QLQ-PR25 are linearly transformed to a 0–100 scale, with higher scores reflecting either more symptoms (urinary, bowel, hormonal treatment-related symptoms) or higher levels of functioning (sexual).

2.4. Statistical analysis

Multitrait scaling analysis was used to examine whether the individual items of the QLQ-PR25 could be aggregated into a more limited set of multi-item scales as hypothesised. Evidence of item convergent validity was defined as a correlation of 0.40 or greater between an item and its own scale (corrected for overlap). Discriminate item validity was supported when the correlation between an item and its hypothesised scale (corrected for overlap) was significantly higher than its correlation with any other scale.

Reliability was assessed with Cronbach's coefficient alpha. As recommended by Nunnally, alpha coefficients of a magnitude of 0.70 or greater were considered acceptable for the purpose of group comparisons.[14]

Three approaches were taken to evaluate the validity of the QLQ-PR25. First, the method of known-group comparisons was used to determine the extent to which the questionnaire scores were able to discriminate between subgroups of patients differing in clinical status. Known groups used for these comparisons were treatment intent (potentially curative versus palliative) and Karnofsky performance status scores (<90 versus 90 versus 100). Group differences were assessed using the Wilcoxon rank sum test. A p-value of 0.05 or less was used to define statistical significance.

Second, we used the two sets of questionnaires, one administered prior to the start of treatment and one during treatment, to evaluate the responsiveness of the QLQ-PR25 to changes in health status and symptoms over time. In this study, improvement or deterioration in health status was defined as a shift of at least one level upwards or downwards on the Karnofsky performance status scale. Repeated measures ANOVA was used to test for significant changes in questionnaire scores as a function of observed changes in performance status.

Third, convergent and divergent validity of the QLQ-PR25 was assessed by evaluating the correlations between that module and the core questionnaire, the QLQ-C30. It was hypothesised that the QLQ-PR25 assesses HRQOL domains distinct from those assessed by the QLQ-C30, and thus Pearson's correlation coefficients of less than 0.40 were sought. All statistical tests were performed using SAS software (version 8.02).

2.5. Sample size estimate

To carry out the requisite multivariate psychometric and statistical tests, a minimum sample size of 375 patients was needed (taking into account the stratification by disease stage). However, to obtain greater balance within the geographical/linguistic categories (Southern Europe, Northern Europe, Anglo-Saxon) and within treatment arms, the following target sample sizes were set: in the curative surgery group, 100 patients for each of the three geographical/linguistic regions; in the curative radiotherapy group, 75 patients in each of the three geographical/linguistic categories; and in the palliative treatment group, a total sample of 100 patients.

3. Results

3.1. Patient characteristics

Between March 2002 and December 2004, 642 patients from 34 institutions in 13 countries were entered into the study. Of these, 53 patients were determined not to have met the eligibility criteria. The primary reasons for ineligibility included the presence of concurrent malignancies, prior hormonal treatment and prior non-hormonal treatment within 2 years of study entry. Of the remaining 589 patients, 80 could not be included in the final analysis due to missing treatment forms, use of the incorrect version/translation of the questionnaires or the absence of at least one valid completed questionnaire. The final study sample, for the purpose of this analysis, included 509 patients. The sociodemographic and clinical characteristics of the sample are shown in Table 1.

3.2. Compliance rates and questionnaire acceptability

Compliance with questionnaire completion was 85% at baseline, 77% at 3 month follow-up and 57% at 6 month follow-up. Most patients (85%) completed the questionnaires in less than 15 min, and 83% did not require any help. In the 17% of cases requiring assistance, this typically involved administering the questionnaire orally. The percentage of patients requiring assistance was significantly greater in the palliative group than in the potentially curative group (28% versus 15%, $p = 0.02$). The majority of patients (79%) found that the questions were clear. No item was reported to be upsetting or confusing by more than 7% of the patients, with the majority of the negative feedback generated by the questions about sexual functioning.

Table 1 – Baseline sociodemographic and clinical characteristics of the sample (n = 509 patients)

Variable	Subgroup		Total (*N* = 509)
	Group I (*N* = 451) *N* (%)	Group II (*N* = 58) *N* (%)	*N* (%)
Age			
Median	66.2	73.3	66.9
Range	41.8–83.7	53.7–93.1	41.8–93.1
Cohabitants			
1. Living alone	43 (9.5)	9 (15.5)	52 (10.2)
2. Living with family	376 (83.4)	47 (81.0)	423 (83.1)
3. Living with other adults	28 (6.2)	2 (3.4)	30 (5.9)
Unknown	4 (0.9)	0 (0.0)	4 (0.8)
Education level			
1. <Compulsory education	29 (6.4)	4 (6.9)	33 (6.5)
2. Compulsory education	193 (42.8)	25 (43.1)	218 (42.8)
3. Post-compulsory below university level	134 (29.7)	23 (39.7)	157 (30.8)
4. University level	75 (16.6)	4 (6.9)	79 (15.5)
5. Others	2 (0.4)	1 (1.7)	3 (0.6)
Unknown	18 (4.0)	0 (0.0)	18 (3.5)
Missing	0 (0.0)	1 (1.7)	1 (0.2)
Marital status			
1. Single	23 (5.1)	5 (8.6)	28 (5.5)
2. Married	388 (86.0)	49 (84.5)	437 (85.9)
3. Separated, divorced or widower	36 (8.0)	4 (6.9)	40 (7.9)
Unknown	4 (0.9)	0 (0.0)	4 (0.8)
TNM: T category			
1. T1	181 (40.1)	1 (1.7)	182 (35.8)
2. T2	209 (46.3)	17 (29.3)	226 (44.4)
3. T3	60 (13.3)	26 (44.8)	86 (16.9)
4. T4	0 (0.0)	13 (22.4)	13 (2.6)
9. Tx	1 (0.2)	1 (1.7)	2 (0.4)
TNM: N category			
0. N0	407 (90.2)	17 (29.3)	424 (83.3)
1. N1	2 (0.4)	14 (24.1)	16 (3.1)
9. Nx	42 (9.3)	27 (46.6)	69 (13.6)
TNM: M category			
0. M0	416 (92.2)	7 (12.1)	423 (83.1)
1. M1	0 (0.0)	47 (81.0)	47 (9.2)
9. Mx	35 (7.8)	4 (6.9)	39 (7.7)
PSA			
Median	8.0	82.0	8.7
Range	0.1–228.0	10.1–948.7	0.1–948.7
Country			
Australia	14 (3.1)	0 (0.0)	14 (2.8)
Belgium	86 (19.1)	0 (0.0)	86 (16.9)
France	61 (13.5)	0 (0.0)	61 (12.0)
Germany	22 (4.9)	0 (0.0)	22 (4.3)
Italy	74 (16.4)	0 (0.0)	74 (14.5)
Norway	25 (5.5)	1 (1.7)	26 (5.1)
Portugal	2 (0.4)	9 (15.5)	11 (2.2)
Romania	0 (0.0)	3 (5.2)	3 (0.6)
Russia	0 (0.0)	0 (0.0)	0 (0.0)
Switzerland	29 (6.4)	5 (8.6)	34 (6.7)
The Netherlands	60 (13.3)	20 (34.5)	80 (15.7)
Turkey	22 (4.9)	9 (15.5)	31 (6.1)
United Kingdom	56 (12.4)	11 (19.0)	67 (13.2)
Karnofsky performance status			
60	0 (0.0)	2 (3.4)	2 (0.4)
70	2 (0.4)	4 (6.9)	6 (1.2)
80	15 (3.3)	7 (12.1)	22 (4.3)
90	61 (13.5)	20 (34.5)	81 (15.9)
100	367 (81.4)	25 (43.1)	392 (77.0)
Unknown	2 (0.4)	0 (0.0)	2 (0.4)
Missing	4 (0.9)	0 (0.0)	4 (0.8)

3.3. *Multitrait scaling analysis and internal consistency*

Results from the multitrait scaling analyses are shown in Table 2. Item-scale correlations for the urinary symptom, bowel symptom, sexual activity and sexual functioning scales exceeded the 0.40 criterion in all cases, both at baseline and at follow-up. For the hormonal treatment-related symptom scale, one of the six items did not meet the 0.40 criterion at baseline, and three of these six items did not meet the criterion at follow-up. At baseline and follow-up, the item on sexual enjoyment correlated more highly (0.67) with the sexual activity scale than with the hypothesised sexual function scale (0.60). At follow-up, the question on urinary leakage correlated more highly with the use of incontinence aid item (0.57) than with the urinary symptom scale (0.52). All other items correlated much higher with their own scale than with other scales at baseline and follow-up.

Several alternative item combinations were investigated in an attempt to improve the psychometric properties of the bowel symptoms and hormonal treatment-related symptoms scales. However, the psychometric properties of these alternative scales were no better than the original, hypothesised scales.

Scale reliability, as assessed by Cronbach's alpha coefficient, met the 0.70 criterion for the urinary symptoms, sexual activity and sexual functioning scales at baseline, and for the urinary symptoms and sexual activity scales at follow-up. For the remaining two scales (bowel symptoms and treatment-related symptoms), the alpha coefficients were below the 0.70 level (Table 2).

3.4. *Scales descriptive statistics*

Scale descriptive statistics are provided in Table 3. The mean bowel symptom (4.5–5.4) and treatment-related scale scores (6.9–11.9) before and after treatment were very low, whilst the standard deviations were fairly large. The percentage of respondents at the ceiling was generally very low. The highest percentage ceiling effect was for the sexual functioning scale before treatment (8.5%). For the urinary, bowel and hormonal treatment-related scales, the maximum possible scores were never observed. For the bowel symptoms scale, the lowest possible scores (floor) were observed for approximately 60% of patients both before and after treatment. Except for the bowel symptom scale and the conditional item on incontinence aid, scores on the QLQ-PR25 scales demonstrated very close to

Table 2 – Scale description, multitrait scaling results and reliability (baseline n = 472 – first row; after treatment n = 463 – second row in cursive)

PR25 – scales/single items – name	Abbreviation	Number of items	Item – own scale correlation	Item – other scale correlation	Cronbach's alpha
Urinary symptoms	PRURI	8	0.53–077	–0.32–0.61	0.86
			0.50–0.78	*–0.14–0.57*	*0.84*
Incontinence aid (conditional)	PRAID	1	1.00	–0.16–0.54	
			1.00	*–0.06–0.51*	
Bowel symptoms	PRBOW	4	0.52–0.69	–0.17–0.41	0.53
			0.48–0.74	*–0.11–0.39*	*0.55*
Hormonal treatment-related symptoms	PRHTR	6	0.30–0.68	–0.35–0.33	0.41
			0.28–0.62	*–0.40–0.31*	*0.39*
Sexual active	PRSAC	2	0.93–0.93	–0.26–0.41	0.85
			0.87–0.92	*–0.06–0.27*	*0.77*
Sexual functioning (conditional)	PRSFU	4	0.60–0.80	–0.34–0.67	0.70
			0.42–0.78	*–0.30–0.17*	*0.59*

Table 3 – Scales descriptive statistics

PR 25 scales	Number of forms	Mean	SD	*N* (%) floor	*n* (%) ceiling	Normality
PR URI (urinary symptoms)	468	16.5	16.3	89 (18.9)	0 (0.0)	0.86
	457	*22.7*	*18.1*	*32 (6.9)*	*1 (0.2)*	*0.91*
PR AID (incontinence aid)	41	10.6	26.3	34 (7.2)	2 (0.4)	0.46
	146	*22.6*	*27.1*	*74 (16.0)*	*5 (1.1)*	*0.77*
PR BOW (bowel symptoms)	397	4.5	8.6	273 (57.8)	0 (0.0)	0.59
	423	*5.4*	*9.4*	*267 (57.7)*	*0 (0.0)*	*0.63*
PR HTR (hormonal treatment related symptoms)	456	6.9	9.1	206 (43.6)	0 (0.0)	0.76
	457	*11.9*	*10.7*	*112 (24.2)*	*0 (0.0)*	*0.9*
PR SAC (sexual active)	458	37.4	27.6	101 (21.4)	16 (3.4)	0.92
	454	*27.8*	*26*	*154 (33.3)*	*5 (1.1)*	*0.88*
PR SFU (sexual function)	304	75	21.4	4 (0.8)	40 (8.5)	0.88
	194	*53.6*	*25.4*	*12 (2.6)*	*5 (1.1)*	*0.97*

Baseline reported in the first row and after treatment in the second row, in *cursive*.
% floor, percentage of respondents at the lowest scale rating. % ceiling, percentage of patients at the highest scale rating. Normality is assessed by Shapiro–Wilk statistic.

the normal distribution with Shapiro–Wilk estimates ranging from 0.76 to 0.97.

3.5. *Known-groups validity and responsiveness*

As hypothesised, patients in clinically distinct groups (curative versus palliative treatment intent; higher versus lower Karnofsky performance status scores) had significantly different baseline scores for almost all scales of the QLQ-PR25 (Table 4). This was also the case for the majority of comparisons at 3 month follow-up. Responsiveness of the QLQ-PR25 to changes in health status over time was evaluated by testing for significant changes in questionnaire scores as a function of observed changes in performance status (defined as a shift of at least one level upwards or downards on the Karnofsky scale). Statistically significant differences over time in the expected direction were observed for the urinary symptom scale, the bowel symptom scale, and the sexual functioning scale, but not for the hormonal treatment-related symptom scale or the sexual activity scale (data not shown in table form).

3.6. *Relationships between the QLQ-PR25 and the QLQ-C30 questionnaire*

None of the QLQ-PR25 scales had high correlations (above 0.40) with the EORTC QLQ-C30 scales (data not shown). The highest observed correlation was between the QLQ-C30 physical function scale and the QLQ-PR25 sexual function scale (0.39). This suggests that the HRQOL issues addressed by the QLQ-PR25 are distinct from those assessed by the more general QLQ-C30.

4. Discussion

The results of this international field study of the QLQ-PR25 generally support its psychometric robustness. The questionnaire was well accepted by patients, yielded high compliance rates and proved to be acceptable to the large majority of patients. As expected items assessing sexual functioning were experienced as problematic by some elderly patients.

The results confirmed the hypothesised scale structure and reliability of 3 of the multi-item scales – urinary symptoms, sexual activity and sexual function – with reliability estimates varying between 0.70 and 0.86. In contrast, the reliability of the bowel function scale and the treatment-related symptom scale was suboptimal (0.53 and 0.41, respectively). In part, the lower reliability observed for these two scales can be explained by limited score variance. Specifically, for both scales, a restricted range of (possible) scores was observed (0–58 for the bowel symptom scale and 0–55 for the treatment-related symptom scale). Cronbach's alpha coefficient reflects the ratio of variances of the individual scale items to the variance of the total scale. The greater the variation within the individual items in relation to the variation of the total scale, the higher the Cronbach alpha. Conversely, a restricted range of responses will have a greater impact on the total scale score than on the individual items, resulting in a lower reliability estimate.

It is important to note that lower reliability estimates resulting from restricted response ranges do not necessarily negate the value of a scale or the items comprising that scale. In interviews conducted during the development of the module, both clinicians and patients rated the items making up these scales as being relevant and important.

One reason for the low prevalence of severe bowel symptoms in this study may be the use of improved radiation techniques such as three-dimensional conformal radiotherapy and/or intensity-modulated radiotherapy. The use of these more sophisticated radiotherapy techniques is known to reduce the prevalence and severity of bowel symptoms.[15]

The limited variability in scores for the treatment-related symptom scale is more difficult to explain. It may be that the older men, whilst experiencing such symptoms, tend to accept them and thus not rate them as being severe. Nevertheless, we would again argue that the items comprising this scale are clinically relevant. Therefore, our decision was not to delete these items from the questionnaire, but rather to advise analysing these domains at both the individual item and the scale level. The fact that all the scales comprising the QLQ-PR25 were able to distinguish clearly between various clinically defined patient subgroups (curative versus pallia-

Table 4 – Mean baseline and follow-up QLQ-PR25 scale scores for clinically distinct groups

HRQOL scales		Treatment intent			Karnofsky score			
		Curative	Palliative	P-value	<90	90	100	P-value
		n=418	n=54		n=32	n=70	n=370	
		n=372	n=26		n=45	n=88	n=265	
PR URI (urinary symptoms)	baseline	15.2	25.4	<.001	27.7 (23.8)	20.3 (17.9)	14.8 (14.7)	0.001
	3 months	23.1	19.5	0.188	29.8()	23.8	20.8	0.007
PR BOW (bowel symptoms)	baseline	3.2	6.6	0.007	7.5	4.4	1.6	0.000
	3 months	5.3	4.6	0.624	6.0	3.1	2.9	0.063
PR HTR (hormonal	baseline	6.8	7.6	0.548	12.6	7.9	6.9	0.024
treatment-related symptoms)	3 months	11.0	15.3	0.050	16.1	18.9	12.9	0.002
PR SAC (sexual active)	baseline	41.1	9.6	0.000	26.1	19.7	41.6	0.000
	3 months	31.8	13.2	0.000	21.6	28.7	32.4	0.035
PR SFU (sexual function)	baseline	76.9	46.8	0.000	54.6	62.2	78.3	0.000
	3 months	54.9	44.3	0.112	42.7	48.4	55.8	0.068

tive; higher versus lower performance status), and can detect change over time in health status, lends support to the value of these measures in documenting the HRQOL of prostate cancer patients.

We would note several limitations of this study. First, although the total sample size was large, the number of patients available within each participating country was insufficient to evaluate possible cross-cultural differences in the psychometric performance of the QLQ-PR25. Hopefully, such country-specific data will become available as local and national studies using the module mature. Second, approximately 20% of patients initially entered into the database had to be excluded from the analysis for various reasons. However, this was due primarily to administrative and logistical problems, rather than to patients' unwillingness to participate.

In summary, the results of this study generally support the hypothesised scale structure, reliability, validity and responsiveness of the QLQ-PR25. The limited reliability observed for several of the multi-item scales may largely reflect the limited variability observed in the scores of the items comprising these scales. We would thus recommend that these scales be interpreted with some caution, and that the results be reported at the individual item as well as at the aggregate level. The heterogeneity of the study sample in terms of geographic, cultural and linguistic background suggests that the questionnaire can be used in international prostate cancer studies. At the same time, we would recommend that future, country-specific studies examine possible cross-cultural differences in both the performance of the questionnaire and the nature, prevalence and severity of HRQOL problems experienced by men with prostate cancer.

Conflict of interest statement

None declared.

Acknowledgements

This work was funded by the EORTC Quality of Life Group and, in part, by Grants 2U10 CA11488-36 through 5U10 CA011488-38 from the National Cancer Institute (Bethesda, MA, USA). Its contents are solely the responsibility of the authors, and do not necessarily represent the official views of the National Cancer Institute. This work was also supported, in part, by the EORTC Charitable Trust and FOCA. Fabio Efficace and Sven D'Haese undertook this work whilst employed at the EORTC Headquarters, Brussels, Belgium. We thank the members of the EORTC Genito-Urinary Tract Cancer Group and the EORTC Quality of Life Group for their support in entering patients into this study, and in particular thank the patients for their willingness to participate.

REFERENCES

1. Crawford ED. Epidemiology of prostate cancer. *Urology* 2003;**62**(Suppl. 1):3–12.
2. Jemal A, Murray T, Samuels A, Ghafoor A, Ward E, Thun MJ. Cancer statistics. *CA Cancer J Clin* 2003;**53**:5–26.
3. Heidenreich A, Aus G, Bolla M, et al. EAU Guidelines on prostate cancer. *Eur Urology* 2007;**19**:68–80.
4. Bolla M, Collette L, Blank L, et al. Long-term results with immediate androgen suppression and external irradiation in patients with locally advanced prostate cancer (an EORTC study): a phase III randomised trial. *Lancet* 2002;**360**(9327): 103–6.
5. Wilt TJ, MacDonald R, Rutks I, Shamliyan TA, Taylor BC, Kane RL. Systematic review: comparative effectiveness and harms of treatments for clinically localized prostate cancer. *Ann Intern Med* 2008;**18**:435–48.
6. Hardie C, Parker C, Norman A, et al. Early outcomes of active surveillance for localized prostate cancer. *BJU Int* 2005;**95**: 956–60.
7. Nair B, Wilt T, MacDonald R, Rutks I. Early versus deferred androgen suppression in the treatment of advanced prostatic cancer. *Cochrane Database Syst Rev* 2002;**1**:CD003506.
8. 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**:365–76.
9. Sprangers MA, Cull A, Groenvold M, Bjordal K, Blazeby J, Aaronson NK. The European Organization for Research and Treatment of Cancer approach to developing questionnaire modules: an update and overview. EORTC Quality of Life Study Group. *Qual Life Res* 1998;**7**:291–300.
10. Esper P, Mo F, Chodak G, Sinner M, Cella D, Pienta KJ. Measuring quality of life in men with prostate cancer using the functional assessment of cancer therapy-prostate instrument. *Urology* 1997;**50**:920–8.
11. Litwin MS, Hays RD, Fink A, Ganz PA, Leake B, Brook RH. The UCLA prostate cancer index: development, reliability, and validity of a health-related quality of life measure. *Med Care* 1998;**36**:1002–12.
12. Wei JT, Dunn RL, Litwin MS, Sandler HM, Sanda MG. Development and validation of the expanded prostate cancer index composite (EPIC) for comprehensive assessment of health-related quality of life in men with prostate cancer. *Urology* 2000;**56**:899–905.
13. Cull A, Sprangers M, Bjordal K, et al. Guidelines for translating EORTC questionnaires. Brussels; 2002.
14. Nunnally JC. *Psychometric theory*. 3rd ed. New York: McGraw-Hill; 1994.
15. Zelefsky MJ, Levin EJ, Hunt M, et al. Incidence of late rectal and urinary toxicities after three-dimensional conformal radiotherapy and intensity-modulated radiotherapy for localized prostate cancer. *Int J Radiat Oncol Biol Phys* 2008;**70**:1124–9.