

ORIGINAL ARTICLE

# Psychometric Validation of the EORTC QLQ-PAN26 Pancreatic Cancer Module for Assessing Health Related Quality of Life after Pancreatic Resection

Anne A Eaton<sup>1</sup>, Paul Karanicolas<sup>2</sup>, Colin D Johnson MChir<sup>3</sup>, Andrew Bottomley<sup>4</sup>, Peter J Allen<sup>5</sup>, Mithat Gonen<sup>1</sup>

Department of <sup>1</sup>Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, 485 Lexington Ave, 2nd Floor, New York, NY, 10017, USA

<sup>2</sup>Sunnybrook Health Sciences Centre, 2075 Bayview Ave, Room T2 16, Toronto, ON, M4N 3M5, Canada

<sup>3</sup>University of Southampton, University Road, Southampton, SO17 1BJ, UK

<sup>4</sup>European Organization for Research and Treatment of Cancer Quality of Life Department, Ave. E. Mounier 83, B. 11, 1200 Brussels, Belgium

<sup>5</sup>Department of Surgery, Memorial Sloan Kettering Cancer Center, 1275 York Ave, New York, NY 10065, USA

## ABSTRACT

**Context** The European Organization for Research and Treatment of Cancer has developed the PAN26 instrument to measure quality of life in patients with pancreatic cancer. Its use has been increasing, but it has not yet undergone psychometric validation in a large cohort or in the setting of pancreatic resection. **Objective** We aimed to validate the PAN26 in patients undergoing pancreatic resection using a high-quality Phase III clinical trial dataset. **Methods** The European Organization for Research and Treatment of Cancer core questionnaire and pancreatic cancer module were administered pre-operatively and at 14 and 60 days post-operatively to 300 patients enrolled in a Phase III trial of pasireotide to prevent pancreatic fistula. Multi-trait scaling analysis was performed; construct validity and internal consistency were assessed. **Results** With the exception of the hepatic scale, the PAN26 scales had adequate internal consistency (Cronbach's alpha = 0.69 - 0.97), and items were more correlated with their own scale than other scales, indicating appropriate aggregation. Adenocarcinoma diagnosis was associated with worse scores on multiple scales. As expected, the PAN26 and C30 pain scales were highly correlated (>0.7). **Conclusions** In the largest psychometric analysis to date of the PAN26, we demonstrated that the scales are reliable and valid, although the appropriateness of the hepatic scale in the post-operative setting may need more examination. We observed differences by final diagnosis (adenocarcinoma or benign), and have shown previously that scores on symptom scales were worse post-operatively than at baseline, confirming the sensitivity of the PAN26 to detect clinically meaningful differences in quality of life.

## INTRODUCTION

Pancreatic cancer and its treatment can have severe negative effects on patients' health related quality of life (QOL), some of which are perceived differently by doctors and patients and are best assessed by patients themselves [1]. The prognosis for pancreatic cancer patients is generally poor; less than 20% of patients are

eligible for potentially curative pancreatic resection, and this treatment is associated with substantial morbidity and measurable mortality. More research on the effects of pancreatic cancer and its treatment on health related QOL and how these effects can be accurately measured is urgently needed in order to provide the best care for these patients.

In response to rising demand, in the European Organization for Research and Treatment of Cancer (EORTC) in 1999 developed the PAN26 instrument, which is a patient-reported measure of health related QOL issues specifically arising in pancreatic cancer patients and is intended to supplement the EORTC Core Quality of Life Questionnaire [2]. The PAN26 has completed Phase III of EORTC module development, meaning that it can be used in clinical trials with permission. Phase IV of development, which requires psychometric testing in a large international group of patients, is ongoing [3, 4]. The

Received July 15th, 2016 - Accepted September 12th, 2016

**Keywords** Pancreatectomy; Pancreatic neoplasms; Psychometrics; Quality of life; Surveys and questionnaires; Validation Studies

**Abbreviations** C30 core questionnaire; EORTC European Organization for Research and Treatment of Cancer; PAN26 pancreatic cancer module

**Correspondence** Anne A Eaton

Memorial Sloan Kettering Cancer Center

485 Lexington Ave, 2nd Floor

New York, NY 10017

Tel + 646 888-8334

Fax + 929 321-1516

E-mail eatona@mskcc.org

tool has been translated using robust EORTC translation processes, ensuring its use is valid across multiple nations [5, 6]. To our knowledge, only a limited amount of previous work describing data collected with the PAN26, and specifically validating it, has been published. We recently described the trajectory of PAN26 and QLQ-C30 scores in the 60 days following surgery but the psychometric validation presented in that paper was not comprehensive and consisted only of confirming that the PAN26 scales had adequate internal consistency [7]. It is critical that this instrument be validated to ensure its appropriateness to measure the most important concerns arising for pancreatic cancer patients in a way that is reliable, interpretable, and sufficiently sensitive to detect health related QOL differences that are meaningful to patients. In addition, the usefulness of the instrument in the pancreatic resection setting needs to be confirmed. Previous validation studies have been according to EORTC procedures for module development [3], on small populations and in the setting of chronic pancreatitis.

This paper describes a psychometric validation study assessing the reliability and validity of the PAN26 carried out using a high-quality dataset from a Phase III clinical trial of 300 patients undergoing surgical resection for pancreatic and peri-pancreatic neoplasms. We also aim to confirm the presence of expected health related QOL differences between clinically different groups and to provide reference data (here and in related publications) to inform future sample size calculations and for use in interpretation [7, 8, 9]. Only by understanding the psychometric properties of the PAN26 will we be able to make the most accurate interpretations of the data collected using this instrument and ensure that clinical decisions are based on solid data rather than solely on expectations and expert opinions.

## METHODS

### Patients and Study Design

Three hundred patients undergoing pancreatic resection were enrolled in a Phase III, single-center, randomized, double-blind, placebo-controlled trial of pre-operative pasireotide to reduce postoperative pancreatic fistula. Primary clinical results have been reported previously [9]. Patients completed the EORTC core cancer module (QLQ-C30) and pancreatic cancer module (QLQ-PAN26) at three time points: prior to surgery (baseline), 14 and 60 days after surgery [10]. The numbers of questionnaires returned at each time point were 299, 273 and 265 respectively. Surveys were completed in person at scheduled clinic visits with the assistance of research study assistants if needed.

### Questionnaire

The QLQ-C30 consists of 28 four-level Likert items, and two seven-level Likert items which are scored according to the EORTC scoring guidelines into the 15 domains. Dyspnea, insomnia, appetite loss, constipation and financial difficulties are single-item scales while the

remaining scales are composed of two to five items [10]. The QLQ-PAN26 consists of 26 four-level Likert items which are scored according to draft scoring procedures supplied by EORTC to obtain seven multi-item scale scores consisting of two to four items and nine single-item scores (Supplemental Table 1).

Domain scoring can be done when less than half the items within a domain are missing and is done by taking the mean of the component items and scaling so that zero corresponds to the lowest possible score and 100 corresponds to the highest possible score [11, 12, 13]. The QLQ-C30 pain scale items ask about general pain while the PAN26 pancreatic pain scale items refer specifically to abdominal discomfort, back pain, pain during the night, and discomfort in certain positions. To distinguish between the pain scales, we will refer to them as QLQ-C30 pain and pancreatic pain. For most scales, higher scores indicate worse symptoms and worse health related QOL. The QLQ-C30 physical, role, emotional, cognitive and social functioning scales, the QLQ-C30 global health status scale, and the PAN26 satisfaction with health care scale are scored as functional scales where higher scores indicate better function and better health related QOL.

### Ethics

This study was approved by the institutional review board at Memorial Sloan Kettering Cancer Center and all subjects provided written consent.

### Statistics

Validity and reliability of the QLQ-PAN26 scales were assessed separately for the three time points due to the trial design and because a previous report demonstrated that scores changed across time points [7]. Correlations between four-level ordinal items were estimated using polychoric correlation and correlations between items and multi-item scale scores were estimated using polyserial correlation. Multi-trait scaling analysis was performed to confirm appropriate scale aggregation, that is, that multi-item scales consisted of similar, inter-correlated items and that items were more correlated with their own scale than with other scales. Item-total correlations were corrected for overlap by leaving the item of interest out of the total. Internal consistency of multi-item scales was assessed using Cronbach's alpha and coefficient omega, a measure of internal consistency that relies on fewer assumptions than alpha, with values  $>0.7$  considered acceptable [14, 15, 16]. For scales with only two items, reliability measures were based on polychoric correlation, which leads to alpha and omega being equivalent because the two factor loadings are equal [17].

Convergent and discriminant validity were assessed using a correlation matrix containing Pearson's correlations between all PAN26 domains versus all PAN26 and QLQ-C30 domains. To assess known-groups validity, we compared mean domain scores by final diagnosis and gender using t-tests ( $p < 0.05$  considered significant). Statistical analysis was performed in SAS 9.3 (SAS Institute, Cary, NC) and

**Table 1.** Multi-trait scaling analysis of PAN26 multi-item scales. Two-item scales yield a single item-total correlation.

Domain	Item	Item-total correlation	Item-other scale correlation (range)	Item-total correlation	Item-other scale correlation (range)	Item-total correlation	Item-other scale correlation (range)
		Baseline		14 days		60 days	
Pancreatic pain	31	0.65	0.09-0.55	0.59	0.05-0.30	0.65	-0.05-0.36
	33	0.72	0.09-0.34	0.55	0.00-0.31	0.56	0.04-0.27
	34	0.92	0.05-0.35	0.79	-0.02-0.35	0.82	-0.01-0.37
	35	0.88	0.11-0.44	0.61	-0.02-0.29	0.74	0.00-0.46
Digestive	36	0.85	0.12-0.46	0.68	0.05-0.26	0.70	-0.02-0.44
	37		0.17-0.54		-0.04-0.36		0.08-0.38
Altered bowel habit	46	0.69	0.13-0.38	0.52	-0.10-0.05	0.66	-0.06-0.24
	47		0.11-0.42		-0.06-0.22		0.01-0.31
Hepatic	44	0.69	0.11-0.47	0.05	-0.03-0.42	0.28	-0.03-0.27
	45		0.05-0.40		-0.02-0.23		-0.21-0.20
Body image	48	0.64	0.21-0.50	0.72	0.02-0.35	0.79	-0.01-0.45
	49		0.09-0.31		-0.01-0.38		0.00-0.52
Health care satisfaction	53	0.90	0.09-0.18	0.85	-0.02-0.24	0.87	-0.09-0.08
	54		0.09-0.18		-0.07-0.15		-0.07-0.07
Sexuality	55	0.92	0.10-0.38	0.94	0.05-0.39	0.90	0.00-0.45
	56		0.09-0.35		-0.01-0.44		0.00-0.50

R 3.1.1 (R Foundation, Vienna, Austria) with the polycor, psych, and GPArotation packages [18, 19, 20]. This study is registered as ClinicalTrials.gov number NCT00994110.

**RESULTS**

Characteristics of the 300 evaluable patients and means and standard deviations of domain scores at each time point have been presented previously [7, 9]. Average age was 65 years and 45.0% of patients were female. Eighty patients (27%) had a distal pancreatectomy and 220 patients (73%) had a pancreaticoduodenectomy. Forty-five patients (15.0%) developed pancreatic complications (grade three or higher postoperative pancreatic fistula, leak or abscess within 60 days). A drain was utilized to manage pancreatic complications in the 45 patients who experienced complications and the median duration of drainage was seven days (range 3-152 days). One hundred and sixty patients had a final diagnosis of pancreatic adenocarcinoma and 84 had a final diagnosis of benign pancreatic neoplasm.

Response rates for the baseline, 14 day and 60 day questionnaires respectively were 99.7%(299/300), 91.0% (273/300) and 88.3% (265/300). On the baseline questionnaire, non-completion rates by item ranged from 0-14.0% depending on the item and non-completion rates by domain ranged from 0-11.0% (Supplemental Table 2). The least-completed question was item 56, "Have you felt less sexual enjoyment during the past week?" and least-completed domain was sexuality.

Multi-trait scaling analysis revealed that all items were more correlated with their own scale than with other scales except for the hepatic scale items at 14 days (**Table 1**). Hepatic item-total correlations were low at 14 days (0.05) and 60 days (0.28). Correlations between PAN26 single-item scales and PAN26 multi-item scales were low to moderate (Supplemental Table 3). The correlation between Q32, "Did you have a bloated feeling in your abdomen?" and pancreatic pain was fairly high: 0.70, 0.49 and 0.68 at baseline, 14 and 60 days. This relationship was

driven by high correlations between item 32 and item 31 ("Have you had abdominal discomfort?"): 0.78, 0.63 and 0.77 at the three time points respectively.

Cronbach's alpha and omega values for baseline measurements ranged from 0.78 to 0.96 (**Table 2**). Internal consistency remained good at the second and third time points with the exception of the hepatic domain, which had values of 0.09 and 0.43 at 14 and 60 days. The hepatic domain consists of two items: "Have you had itching?" and "To what extent was your skin yellow?" At 14 and 60 days, few patients reported any extent of yellow skin (9.8% and 2.8% respectively), but more than one out of four patients had itching and 2-3% of patients had "very much" itching.

**Table 3** shows correlations between each of the PAN26 multi-item scales and the other PAN26 scales, and the QLQ-C30 scales. Correlation between PAN26 domains was generally low (<0.5). As expected, pancreatic pain and QLQ-C30 pain had a high correlation (>0.7). Correlations between QLQ-C30 and PAN26 symptom scales were positive or very close to zero with one exception (correlation between constipation and altered bowel habit at 14 days = -0.21), while correlations between QLQ-C30 function scales and PAN26 symptom scales were negative. Correlations between the PAN26 single-item scales were positive and low to moderate in magnitude (Supplemental Table 4). As expected, and similarly to the PAN26 multi-item symptom scales, the PAN26 single-item scales were negatively correlated with QLQ-C30 function scales and (with two exceptions of correlations close to zero) positively correlated with QLQ-C30 symptom scales. Correlations of high magnitude were observed between item 52 ("Were you limited in planning activities in advance (e.g. meeting friends)?") and social functioning and global health status, and between item 42 ("Did your arms and legs feel weak?") and fatigue. In general, correlations between the health care satisfaction scale and all other scales were close to zero (range: -0.17 to 0.17). Correlations between scales generally had a smaller magnitude at 14 days compared to pre-op and 60 days.

**Table 2.** Omega and Cronbach's alpha measures of reliability. For two-item scales, alpha and omega are equivalent.

Domain	Baseline	14 days	60 days
Pancreatic pain	Omega: 0.84 Alpha: 0.83	Omega: 0.78 Alpha: 0.78	Omega: 0.79 Alpha: 0.78
Digestive	0.92	0.81	0.82
Altered bowel habit	0.82	0.69	0.80
Hepatic	0.82	0.09	0.43
Body image	0.78	0.84	0.88
Health care satisfaction	0.95	0.92	0.93
Sexuality	0.96	0.97	0.95

At baseline, female gender was significantly associated with worse outcomes on the pancreatic pain and body image scales and diagnosis of pancreatic adenocarcinoma was associated with worse outcomes on the digestive, altered bowel habit, hepatic and sexuality scales (**Table 4**, all  $p < 0.05$ ).

## DISCUSSION

This paper reports on the psychometric properties of the PAN26 pancreatic cancer module of the EORTC Quality of Life Group in patients undergoing pancreatic resection. To our knowledge, only a limited amount of research has been done assessing the reliability and validity of the PAN26. A previous development study by Fitzsimmons *et al.* assessed the appropriateness of the PAN26 instrument in a cross-cultural population of 66 patients with chronic pancreatitis, 36 of whom underwent resection [11]. They found that the instrument had adequate internal consistency, that correlations between conceptually related scales were high, and that the instrument detected differences in health related QOL based on both performance status and the requirement for opiate analgesia. Shaw *et al.* assessed quality of life in 40 patients who had undergone pancreaticoduodenectomy at a median of 42 months after surgery, demonstrating that the instrument can distinguish between long term survivors and matched controls on some problems associated with exocrine insufficiency (upper gastrointestinal symptoms, weight loss, muscular weakness) [21]. The levels of other pancreatic-specific symptoms were not significantly different in the two groups, probably due to the long time period between surgery and assessment. Recently, the PAN26 has been shown to be reliable in patients receiving chemoradiotherapy [22]. As in the present study, Cronbach's alpha was  $>0.7$  for all scales except hepatic. The instrument was able to detect changes in health related QOL after radiotherapy, which recovered at later assessment. Comparison of known groups identified clinically significant differences; however, numbers in that study were insufficient for full psychometric assessment of scale structure. The PAN26 has also been reported after total pancreatectomy [23, 24].

During Phase I-III of development of the PAN26 module, the instrument was translated into 10 European languages and it is now available in over 30 languages other than English [2]. The PAN26 was translated into Lithuanian, following EORTC translation procedures, by Vanagas *et al.* who also performed a preliminary assessment of the internal consistency of the scales based

on limited validation data from 13 patients with pancreatic cancer [6, 25]. Although Cronbach's alpha was quite low for the digestive and health care satisfaction scales, they concluded that these deviations were insignificant due to the small number of patients and the small number of questions (two) comprising the scales.

In the largest psychometric analysis of the PAN26 to date, we found that the instrument was generally valid and sensitive to clinically relevant differences in health related QOL including impairment on multiple scales prior to surgery in patients with cancerous versus benign diagnoses. Previous analyses have shown that the instrument can detect the short term morbidity associated with pancreatic resection, with significantly lower scores on all scales except hepatic at 14 days post resection, compared to baseline, as well as the negative impact of pancreatic complications on quality of life [7].

Multi-trait scaling analysis generally supported appropriate aggregation of items as almost all items were more correlated with their own scale than any other scale. (Note that the scales used here are hypothesized scales; confirmed scales will be released when Phase IV of validation is complete.) The one exception to this was the hepatic scale, where item-total correlation was only 0.05 at 14 days and remained low at 60 days. The hepatic scale also had low internal consistency at postoperative time points. Similar findings in patients treated with chemoradiation therapy suggest that the hepatic scale is probably most relevant at the time of presentation and for those treated by palliative stenting, in whom obstructive jaundice is more common [22]. Fitzsimmons *et al.* also observed low internal validity for the hepatic scale (Cronbach's alpha=0.18) [11]. The hepatic scale encompasses two symptoms, itching and yellow skin. The low incidence of yellow skin (which may be expected to resolve following surgery) and other causes of itching (opioid analgesia, time spent in bed, wound healing) postoperatively are likely the main reasons why this scale has poor internal consistency in postoperative patients. Nevertheless, the hepatic scale was able to identify significant differences at baseline between benign and malignant tumors (which are more likely to cause jaundice), supporting the construct validity of this scale. The reliability problem may be limited to the post-surgery and chronic pancreatitis settings. Internal consistency was good for all other scales.

Single items that were not part of a scale generally had low correlations with existing scales and thus there is

**Table 3.** Correlation between PAN26 scales and between PAN26 and QLQ-C30 scales. The three values in a cell correspond to the time points in chronological order.

QLQ-PAN26 multi-item scales	QLQ-PAN26 Multi-Item Scales						
	Pancreatic pain	Digestive	Altered bowel habit	Hepatic	Body image	Health care satisfaction	Sexuality
Pancreatic pain	1.00	0.47 0.31 0.39	0.23 0.00 0.21	0.13 0.34 0.23	0.32 0.33 0.34	0.13 0.13 0.02	0.26 0.19 0.21
Digestive	—	1.00	0.33 0.00 0.17	0.28 0.08 0.07	0.35 0.23 0.17	0.14 0.07 0.09	0.25 0.22 0.17
Altered bowel habit	—	—	1.00	0.39 0.05 0.14	0.23 0.14 0.27	0.14 -0.06 -0.06	0.17 0.10 0.25
Hepatic	—	—	—	1.00	0.23 0.14 0.08	0.09 0.07 -0.09	0.13 0.03 0.15
Body image	—	—	—	—	1.00	0.13 0.01 0.02	0.33 0.37 0.45
Health care satisfaction	—	—	—	—	—	1.00	0.10 0.15 0.01
Sexuality	—	—	—	—	—	—	1.00
<b>QLQ-C30 scales</b>							
Physical functioning	-0.30 -0.36 -0.48	-0.26 -0.26 -0.43	-0.13 -0.10 -0.10	-0.11 -0.15 -0.12	-0.36 -0.36 -0.37	-0.07 -0.14 -0.04	-0.28 -0.31 -0.32
Role functioning	-0.40 -0.32 -0.43	-0.43 -0.31 -0.43	-0.19 -0.09 -0.05	-0.21 -0.18 -0.13	-0.38 -0.18 -0.36	-0.11 -0.09 -0.10	-0.35 -0.25 -0.29
Emotional functioning	-0.27 -0.31 -0.48	-0.21 -0.18 -0.31	-0.27 -0.05 -0.24	-0.10 -0.12 -0.09	-0.29 -0.50 -0.50	-0.06 -0.02 -0.06	-0.20 -0.34 -0.32
Cognitive functioning	-0.27 -0.40 -0.35	-0.21 -0.25 -0.33	-0.18 -0.08 -0.25	-0.11 -0.12 -0.13	-0.33 -0.39 -0.33	-0.15 -0.09 0.03	-0.24 -0.39 -0.21
Social functioning	-0.33 -0.31 -0.43	-0.48 -0.32 -0.50	-0.32 -0.06 -0.24	-0.22 -0.18 -0.10	-0.49 -0.40 -0.47	-0.06 -0.17 -0.07	-0.26 -0.46 -0.48
Global health status	-0.48 -0.35 -0.49	-0.49 -0.34 -0.37	-0.36 -0.07 -0.21	-0.39 -0.15 -0.18	-0.37 -0.35 -0.46	-0.17 0.02 0.02	-0.30 -0.39 -0.41
Fatigue	0.47 0.43 0.53	0.51 0.34 0.42	0.30 0.04 0.26	0.36 0.16 0.15	0.40 0.36 0.36	0.17 0.09 0.08	0.34 0.35 0.42
Nausea and vomiting	0.34 0.36 0.35	0.44 0.19 0.18	0.17 -0.02 0.09	0.22 0.14 0.07	0.19 0.24 0.20	0.04 0.08 -0.01	0.12 0.23 0.10
QLQ-C30 Pain	0.72 0.70 0.75	0.47 0.23 0.36	0.21 0.00 0.11	0.10 0.30 0.17	0.22 0.33 0.32	0.10 0.12 0.02	0.18 0.28 0.18
Dyspnea*	0.38 0.3 0.35	0.16 0.22 0.27	0.07 0.14 0.22	0.15 0.12 0.07	0.31 0.26 0.37	0.08 -0.15 -0.05	0.18 0.26 0.35
Insomnia*	0.42 0.44 0.58	0.24 0.05 0.24	0.16 -0.01 0.19	0.15 0.21 0.21	0.26 0.21 0.32	0.16 0.01 -0.05	0.24 0.21 0.19
Appetite loss*	0.47 0.38 0.49	0.53 0.24 0.51	0.21 0.02 0.19	0.32 0.11 0.14	0.32 0.36 0.38	0.08 0.1 0.06	0.37 0.34 0.33
Constipation*	0.5 0.39 0.4	0.22 0.23 0.3	0.01 -0.21 -0.09	0.11 0.1 0.02	0.21 0.26 0.08	0.01 0.04 -0.09	0.23 0.05 0.14
Diarrhea*	0.15 0.2 0.22	0.28 0.21 0.22	0.58 0.44 0.5	0.36 0.17 0.05	0.21 0.34 0.39	0.1 -0.01 -0.03	0.14 0.13 0.21
Financial difficulties*	0.25 0.27 0.35	0.27 0.34 0.29	0.34 -0.01 0.31	0.02 0.08 0.09	0.37 0.33 0.41	0.14 -0.02 0.05	0.24 0.14 0.25

\*indicates single-item scales

**Table 4.** Baseline PAN26 domain scores by gender and final pathologic diagnosis.

Domain	Male (N=165)	Female (N=134)	P-value	Pancreatic adenocarcinoma (N=160)	Benign diagnoses (N=84)	P-value
Pancreatic pain	13.43	23.98	<0.0001	18.99	19.05	0.9848
Digestive	15.96	22.01	0.0630	23.54	11.71	0.0017
Altered bowel habit	20.51	19.42	0.7038	23.44	14.66	0.0083
Hepatic	15.15	15.05	0.9716	21.56	4.37	<0.0001
Body image	11.21	19.65	0.0007	17.08	12.30	0.0881
Health care satisfaction	77.33	79.32	0.6092	77.46	78.71	0.7824
Sexuality	34.73	35.61	0.8428	38.31	27.22	0.0234

no evidence that they should be merged into the existing scales. A possible exception is question 32, "Did you have a bloated feeling in your abdomen?" which had a fairly high correlation with the pancreatic pain scale (0.70 at baseline, 0.68 at 60 days), and specifically, with item 31 ("Have you had abdominal discomfort?"), indicating that some patients may interpret items 31 and 32 as asking about very similar symptoms. The preliminary analysis in Phase III of development of the PAN26 showed that item 32 was moderately correlated with both pancreatic pain and digestive symptom, and this item was ultimately not included in any scale [2].

Correlations between the PAN26 scales and across PAN26 and QLQ-C30 scales supported the construct validity of the PAN26, with high correlation between conceptually related scales (pancreatic pain and QLQ-C30 pain), generally positive correlations between symptom scales, and negative correlations between symptom severity and function. Interestingly, correlation between health care satisfaction and symptom and function scales was generally of low magnitude, possibly indicating that satisfaction is not related to outcome but rather to parts of the disease/treatment process not measured by these instruments. Administering the instrument was feasible as we were able to achieve high survey completion rates, likely partially due to the short time needed to complete the survey [25]. Similar to others, we observed that questions related to sexuality were the most likely to be skipped; many factors can contribute to this [10, 22, 26, 27].

Our population of interest was patients undergoing pancreatic resection. Though this population is heterogeneous in terms of pathologic diagnosis and includes patients with benign as well as malignant neoplasms, all resected patients have similar symptoms and short-term quality of life regardless of their diagnoses, thus this is an appropriate population for validation of the PAN26. Our analysis had several limitations. Only 20% of pancreatic cancer patients are suitable to undergo resection, so other studies are required to assess the instrument across the full range of stage and treatments. We could not assess the reproducibility of the PAN26 since we did not have access to test-retest data, and our trial did not include a debriefing questionnaire, which examines patients' reasons for non-compliance and whether items are upsetting or distressing and records the time

to complete the tools. An additional limitation of this study is the lack of long-term follow-up data. Lastly, the generalizability of our data may be limited as it came from a single center; we await ongoing EORTC validation studies for additional insight into psychometric issues and use of the instrument in international, cross-cultural settings. Nevertheless, the EORTC have used clinical trial data in the past to validate tools, as have other investigators reporting EORTC module validation data, and our validation study represents the largest and most thorough psychometric analysis to date of the PAN26 [26, 27].

Overall, we found that the PAN26 is valid, reliable and able to detect clinically important differences in patients undergoing pancreatic resection. The hepatic scale identified impaired health related QOL in patients with adenocarcinoma (versus benign conditions) at baseline but may not be valid in the post-operative setting; additional research is needed to determine in which situations its use is appropriate. With the exception of QLQ-C30 pain and pancreatic pain, we did not find high correlations between QLQ-C30 and PAN26 scales, indicating that the PAN26 supplements the QLQ-C30 core questionnaire by measuring aspects of health related QOL not already captured in the QLQ-C30. We look forward to a full EORTC Quality of Life Group validation study of the PAN26 module, which we anticipate will confirm our findings. In light of the extreme importance of patient-reported quality of life, especially in the setting of short life expectancy and treatments that may severely impair quality of life, reliable assessment of health related QOL in patients with pancreatic cancer is crucial. We have shown that the use of the PAN26 in studies of surgical treatment will yield reliable results, allowing patients to compare treatment options based on empirical data.

### Acknowledgements - Sources of Support

This work was supported in part by the NCI Cancer Center core grant P30 CA008748 and Novartis.

### Conflict of Interest Statement

Dr Allen has worked in a consulting/advisory role for Sanofi and received research funding from Novartis. Dr Bottomley is an author of the EORTC QLG measurement system. EORTC tools are provided free to academics;

however, a user fee is applicable for use in industry-sponsored research. These profits cover costs of future psychometric validation, testing and translation of EORTC tools.

## References

1. Fitzsimmons D, George S, Payne S, Johnson CD. Differences in perception of quality of life issues between health professionals and patients with pancreatic cancer. *Psychooncology* 1999; 8:135-43. [PMID: 10335557]
2. Fitzsimmons D, Johnson CD, George S, Payne S, Sandberg AA, Bassi C, Beger HG, et al. Development of a disease specific quality of life (QoL) questionnaire module to supplement the EORTC core cancer QoL questionnaire, the QLQ-C30 in patients with pancreatic cancer. EORTC Study Group on Quality of Life. *Eur J Cancer* 1999; 35:939-41. [PMID: 10533475]
3. Johnson C, Aaronson N, Blazeby JM, Bottomley A, Fayers P, Koller M, Kuliš D, et al. Guidelines for developing questionnaires modules. 3rd ed. ed. 2011, Brussels: EORTC Publications.
4. EORTC. Modules in development and available for use. [cited 2015 June 1]; Available from: <http://groups.eortc.be/qol/modules-development-and-available-use>.
5. Kuliš D1, Arnott M, Greimel ER, Bottomley A, Koller M; European Organisation for Research and Treatment of Cancer Quality of Life Group. Trends in translation requests and arising issues regarding cultural adaptation. *Expert Rev Pharmacoecon Outcomes Res* 2011; 11:307-14. [PMID: 21671700]
6. Dewolf L, Koller M, Velikova G, Scott N, Bottomley A. 2009EORTC Quality of Life Group Translation Procedure 3rd edn. EORTC Quality of Life Group.
7. Eaton AA, Gonen M, Karanicolos P, Jarnagin WR, D'Angelica MI, DeMatteo R, Kingham TP, et al. Health-related quality of life after pancreatectomy: results from a randomized controlled trial. *Ann Surg Oncol* 2016; 23:2137-45. [PMID: 26786091]
8. Scott NW, Fayers PM, Aaronson NK, Bottomley A, de Graeff A, Groenvold M, Gundy C, et al. Sprangers MAG on behalf of the EORTC Quality of Life Group. EORTC QLQ-C30 Reference Values. EORTC Quality of Life Group Publications, Brussels, 2008.
9. Allen PJ, Gönen M, Brennan MF, Bucknor AA, Robinson LM, Pappas MM, Carlucci KE, D'Angelica MI, et al. Pasireotide for Postoperative Pancreatic Fistula. *N Engl J Med* 2014; 370:2014-2022. [PMID: 24849084]
10. Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, Filiberti A, 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. [PMID: 8433390]
11. Fitzsimmons D, Kahl S, Butturini G, van Wyk M, Bornman P, Bassi C, Malfertheiner P, et al., Symptoms and quality of life in chronic pancreatitis assessed by structured interview and the EORTC QLQ-C30 and QLQ-PAN26. *Am J Gastroenterol* 2005; 100:918-26. [PMID: 15784041]
12. EORTC Quality of Life Group, Draft scoring procedure for the EORTC-PAN26. 2007.
13. Fayers PM, Aaronson NK, Bjordal K, Groenvold M, Curran D, Bottomley A. The EORTC QLQ-C30 Scoring Manual 3rd Edition ed. 2001, Brussels: European Organisation for Research and Treatment of Cancer.
14. Dunn TJ, Baguley T, Brunnsden V. From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *Br J Psychol* 2013; 105:399-412. [PMID: 24844115]
15. Peters GJY. The alpha and the omega of scale reliability and validity: Why and how to abandon Cronbach's alpha and the route towards more comprehensive assessment of scale quality. *European Health Psychologist* 2014; 16:56-69.
16. Nunnally J. *Psychometric Theory*. 1978, New York: McGraw-Hill.
17. Zinbarg RE, Revelle W, Yovel I, Li W. Cronbach's  $\alpha$ , Revelle's  $\beta$ , and McDonald's  $\omega$ H: their relationships with each other and two alternative conceptualizations of reliability. *Psychometrika* 2005; 70:123-133.
18. Fox J. *Polycor: Polychoric and Polyserial Correlations*. 2010.
19. Revelle W. *Psych: Procedures for Personality and Psychological Research*. 2015.
20. Bernaards CA, Jennrich RI. Gradient projection algorithms and software for arbitrary rotation criteria in factor analysis. *Educational and Psychological Measurement* 2005; 65:676-696.
21. Shaw CM, O'Hanlon DM, McEntee GP. Long-term quality of life following pancreaticoduodenectomy. *Hepatogastroenterology* 2005; 52:927-32. [PMID: 15966234]
22. Hurt CN, Mukherjee S, Bridgewater J, Falk S, Crosby T, McDonald A, Joseph G, et al. Health-related quality of life in scalop, a randomized phase 2 trial comparing chemoradiation therapy regimens in locally advanced pancreatic cancer. *Int J Radiat Oncol Biol Phys* 2015; 93:810-8. [PMID: 26530749]
23. Epelboym I, Winner M, DiNorcia J, Lee MK, Lee JA, Schrope B, Chabot JA, et al. Quality of life in patients after total pancreatectomy is comparable with quality of life in patients who undergo a partial pancreatic resection. *J Surg Res* 2014; 187:189-96. [PMID: 24411300]
24. Billings BJ, Christein JD, Harmsen WS, Harrington JR, Chari ST, Que FG, Farnell MB, et al. Quality-of-life after total pancreatectomy: is it really that bad on long-term follow-up? *J Gastrointest Surg* 2005; 9:1059-66; discussion 1066-7. [PMID: 16269376]
25. Vanagas T, Gulbinas A, Johnson C, Pundzius J, Barauskas G. Translation and validation of the EORTC QLQ-PAN26 quality of life questionnaire for patients with pancreatic cancer. *Acta Medica Lithuanica* 2006; 13:272-275.
26. Taphoorn MJ, Claassens L, Aaronson NK, Coens C, Mauer M, Osoba D, Stupp R, et al. An international validation study of the EORTC brain cancer module (EORTC QLQ-BN20) for assessing health-related quality of life and symptoms in brain cancer patients. *Eur J Cancer* 2010; 46:1033-40. [PMID: 20181476]
27. Blazeby JM, Hall E, Aaronson NK, Lloyd L, Waters R, Kelly JD, Fayers P. Validation and reliability testing of the EORTC QLQ-NMIBC24 questionnaire module to assess patient-reported outcomes in non-muscle-invasive bladder cancer. *Eur Urol* 2014; 66:1148-56. [PMID: 24612661]