VALUE IN HEALTH 22 (2019) 92–98

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/jval

Patient-Reported Outcomes

Assessing Emotional Functioning with Increased Power: Relative Validity of a Customized EORTC Short Form in the International ACTION Trial

Lea J. Jabbarian, MSc [1,], Mogens Groenvold, MD, PhD [2,3], Morten Aa. Petersen, MSc [2], Caroline Arnfeldt Christensen, MSc [3], Luc Deliens, MSc, PhD [4,5], Agnes van der Heide, MD, PhD [1], Marijke C. Kars, RN, MSc, PhD [6], Guido Miccinesi, MD [7], Sheila Payne, BA(Hons), DipN, RN, CPsychol, PhD [8], Andrew Wilcock, DM, FRCP [9], Ida J. Korfage, MSc, PhD [1], on behalf of the ACTION consortium*

[1]*Department of Public Health, Erasmus Medical Center Rotterdam, Rotterdam, The Netherlands;* [2]*The Research Unit, Department of Palliative Medicine, Bispebjerg Hospital, Copenhagen, Denmark;* [3]*Department of Public Health, University of Copenhagen, Copenhagen, Denmark;* [4]*End-of-life Care Research Group, Vrije Universiteit Brussel (VUB) and Ghent University, Brussels, Belgium;* [5]*Department of Medical Oncology, Ghent University Hospital, Ghent, Belgium;* [6]*Julius Center for Health Sciences and Primary Care, University Medical Center/Utrecht University, Utrecht, The Netherlands;* [7]*Clinical Epidemiology Unit, Cancer Prevention and Research Institute (ISPO), Florence, Italy;* [8]*International Observatory on End of Life Care, Division of Health Research, Faculty of Health and Medicine, Lancaster University, Lancaster, UK;* [9]*Department of Oncology, University of Nottingham and Nottingham University Hospitals NHS Trust, Nottingham, UK*

ABSTRACT

Objectives: There is a need to improve the assessment of emotional functioning (EF). In the international Advance Care Planning: an Innovative Palliative Care Intervention to Improve Quality of Life in Cancer Patients - a Multi-Centre Cluster Randomized Clinical Trial (ACTION) trial involving patients with advanced cancer, EF was assessed by a customized 10-item short form (EF10). The EF10 is based on the European Organisation for Research and Treatment of Cancer (EORTC) EF item bank and has the potential for greater precision than the common EORTC Quality of Life Questionnaire Core 30 four-item scale (EF4). We assessed the relative validity (RV) of EF10 compared with EF4. **Methods:** Patients from Belgium, Denmark, Italy, the Netherlands, Slovenia, and the United Kingdom completed EF10 and EF4, and provided data on generic quality of life, coping, self-efficacy, and personal characteristics. Based on clinical and sociodemographic variables and questionnaire responses, 53 "known groups" that were expected to differ were formed, for example, females versus males. The EF10 and EF4 were first independently compared within this known group, for example, the EF10 score of females vs the EF10 score of males. When these differences were significant, the RV was calculated for the comparison of the EF10 with the EF4. **Results:** A total of 1028 patients (57% lung, 43% colorectal cancer) participated. Forty-five of the 53 known-groups comparisons were significantly different and were used for calculating the RV. In 41 of 45 (91%) comparisons, the RV was more than 1, meaning that EF10 had a higher RV than EF4. The mean RV of EF10 compared with that of EF4 was 1.41, indicating superior statistical power of EF10 to detect differences in EF. **Conclusions:** Compared with EF4, EF10 shows superior power, allowing a 20% to 34% smaller sample size without reducing power, when used as a primary outcome measure.

Keywords: cancer, methodology, patient-reported outcomes, quality of life, relative validity

Introduction

Assessing the quality of life of patients with cancer is common practice in clinical trials and is strongly suggested as a screening instrument for detecting distress and improving care [1,2]. Commonly used questionnaires often lack precision [3,4] and may have problems with floor and/or ceiling effects, limiting their ability to detect differences between groups and change over time.

Computerized adaptive testing (CAT) is a relatively new technique that enables more efficient data collection, for example, of patient-reported outcomes. The content and the number of questions presented are selected according to the participant's

* *Address correspondence to*: Lea J. Jabbarian, Department of Public Health, Erasmus Medical Center Rotterdam, P.O. Box 2040, 3000 CA Rotterdam, The Netherlands.

E-mail: l.jabbarian@erasmusmc.nl

1098-3015/$36.00 - see front matter

https://doi.org/10.1016/j.jval.2018.07.002

previous responses; that is, they are restricted to those relevant to that specific participant [5,6]. For example, if a participant's responses indicate severe emotional problems, the next item will be one that is relevant for people with such severe problems [6]. The items used in CAT are derived from so-called item banks [6]. These items have been calibrated (estimated) to an item-response theory model [7,8], which means that scores based on any subset of the items are comparable [6].

Because CAT has a higher validity (i.e., the statistical power) than traditional measures, it has the potential to reduce trial sample size requirements without reducing power [6]. Even when it is not possible to complete questions on a computer, the item banks underlying CAT can be useful. Taking into consideration the participant's characteristics, such as age or type and stage of cancer, relevant items can be selected from the item bank and used in pen and paper questionnaires, so-called customized short forms.

Currently, several organizations work on the enhancement of clinical outcomes research by developing efficient measures of patient-reported outcomes using item banks. In the United States, the Patient-Reported Outcomes Measurement Information System (PROMIS) has developed item banks. In Europe, the European Organisation for Research and Treatment of Cancer (EORTC) Quality of Life Group has developed 14 item banks for each of the domains (excluding overall physical condition/quality of life) covered by its core quality-of-life measure, the Quality of Life Questionnaire Core 30 (QLQ-C30). This includes one of its key domains: emotional functioning (EF) [6,9]. Traditionally, EF is assessed with the QLQ-C30 four-item EF scale (EF4), measuring depression, anxiety, and general distress [10,11], or with a shortened two-item version (EF2) as included in the Quality of Life Questionnaire Core 15 Palliative Care (QLQ-C15-PAL) questionnaire, an abbreviated version of the QLQ-C30 for use in palliative care [12]. To date, the EF item bank has only been tested in the data set used for its development and no external validation has been performed. This study is the first to test a customized 10-item short form (EF10) based on the EORTC EF item bank in an external, independent, and international data set.

Our hypothesis is that the customized EF10 will provide more precise results, that is, better discrimination between groups, and thus higher relative validity (RV) and lower expected sample size requirements than the original QLQ-C30 EF scale (EF4). The primary aim was to compare the RV, which is also known as the relative efficiency, of the EF10 with that of the EF4. The secondary aims were to compare the RV of the EF10 with that of the EF2 and to compare the RV of the EF4 with that of the EF2.

Methods

Sample

The Advance Care Planning: an Innovative Palliative Care Intervention to Improve Quality of Life in Cancer Patients - a Multi-Centre Cluster Randomized Clinical Trial (ACTION) cluster randomized clinical trial investigates an adapted version of the Respecting Choices advance care planning program. Patients were recruited in pulmonology and oncology departments in hospitals in Belgium, Denmark, Italy, the Netherlands, Slovenia, and the United Kingdom between June 2015 and May 2017. Patients were invited by their treating health care professional. The ACTION study protocol has been described elsewhere [13].

Eligibility requirements included a histologically verified cancer diagnosis of either lung cancer stage III or IV, or colorectal cancer stage IV, being aged at least 18 years, and being physically and mentally competent to give consent and complete the questionnaire. Patients had to have a World Health Organization/Eastern Cooperative Oncology Group performance status of 0 to 3 [14]. Written informed consent was obtained. Ethics committees of the participating countries approved the study.

Patients provided information about their age, sex, living situation, and educational level. Their health care professionals provided clinical information, such as the date of diagnosis, current disease stage, and current treatment.

Assessment of the RV

See Table 1.

We used the method of known-groups validation to evaluate the RV (i.e., statistical power to detect group differences) of the EF10 compared with those of the EF4 and the EF2 [18]. EF is a multidimensional and complex construct that is influenced by various characteristics on the patient level (e.g., sex), clinical level (e.g., type of cancer), coping, satisfaction and experience with care, and patient involvement. For each patient characteristic and for each questionnaire item (except for the EF items), participants were divided into "known groups": two groups based on a priori formulated expectations on differences with regard to EF. For example, we hypothesized that patients who felt nauseated would have a worse EF than would patients who did not feel nauseated. In case of continuous variables, such as age, the median value was used as a cutoff for the dichotomization. For example, we hypothesized that older patients would have a better EF than would younger patients [19]. If the median was similar to the highest or the lowest score, and thus no groups could be

Table 1 – Scales for the assessment of emotional functioning

The EORTC QLQ-C30 EF scale (EF4) consists of four items asking about feeling tense, worrying, feeling depressed, and being irritable [11]. The EF4 was scored using sum scoring following the EORTC scoring manual [15].

The two-item EF version in the QLQ-C15-PAL (EF2) consists of the items about feeling depressed and feeling tense [12]. The EF2 was scored using an appendix to the EORTC scoring manual [16].

A customized 10-item EF short form (EF10) was composed for the trial, including the original 4 QLQ-C30 items (EF4) and 6 additional items from the EORTC EF item bank [6,9]. The EORTC EF item bank includes 24 items and is based on the response of 1023 patients from different countries. The development of the item bank has been described elsewhere in more detail [6]. The items for this study were selected by methodological and clinical experts, based on characteristics of the target population, such as age and type of cancer [17]. The selection of items based on these population characteristics is aimed at making the measure more applicable for the specific population in this study (i.e., patients with advanced cancer). Based on results of observed [6] and simulated data (currently in press) put together, asking fewer than 5 to 6 items might give notable loss in power, whereas asking more than about 14 items would give only negligible gain in power. Hence, 10 items were chosen as a good compromise between response burden and optimal measurement precision. The EF10 was scored using the item-response theory model calibrated for the EORTC EF item bank [6]. The complete questionnaire is included in the Supplemental Materials found at https://doi.org/10.1016/j.jval.2018.07.002. All EF items concern the experiences "during the past week" and use a four-point response scale ranging from "not at all" to "very much."

EORTC, European Organisation for Research and Treatment of Cancer; Quality of Life Questionnaire Core 30 (QLQ-C30); Quality of Life Questionnaire Core 15 Palliative Care (QLQ-C15-PAL).

created on the basis of median value, we calculated the mean value and used it as the cutoff to create two groups. In total, we formed 53 pairs of known groups for the comparison. The EF10, EF4, and EF2 were then each compared within these known groups (e.g., the EF10 of patients who felt nauseated vs the EF10 of patients who did not feel nauseated).

We used the two-sample *t* test (assuming equal variances) to assess significant differences between the known groups for the comparison. The RV was calculated only when the *t* test for at least one of the measures being compared (EF10, EF4, or EF2) showed a significant difference ($P < 0.05$) [18]. For each comparison, the *t* statistic was calculated and subsequently squared. The ratio (= RV) was calculated for each comparison by using the squared *t* statistic of the EF10 as the numerator and the squared *t* statistic of the EF4 or the EF2 as the denominator [6]. We used the EF10 as the numerator because we wanted to evaluate the potential gain (or loss) in measurement precision and power using the customized EF scale (EF10) compared with the two existing scales (EF4 and EF2). Hence, RV > 1 would confirm the expectation that EF10 is the more precise measure. The mean RV across all characteristics or variables was calculated with a bootstrap-based 95% confidence interval [20]. A mean RV above 1 indicates that the EF10 has higher RV than the EF4 or the EF2. With increasing RV, one can expect more power gained by using the EF10. On the basis of the mean RV, we estimated the potential savings in sample size requirements using the EF10 compared with using the EF4 or the EF2 [21]. See Table 2.

Known-Group Comparisons

The known-group comparisons were based on the following measures and characteristics.

Sociodemographic and clinical characteristics

Patients provided information about their age, sex, marital status, whether they had children or not, educational level, and religiosity. Their health care professionals provided clinical background information on the type and current stage of the disease and the time since diagnosis of both the primary tumor and the current stage of the disease. In addition, they indicated whether the patient received chemotherapy.

Quality of life and symptoms

The EORTC Quality of Life Questionnaire Core 15 Palliative Care (QLQ-C15-PAL), an abbreviated version of the EORTC QLQ-C30 for palliative care, was used to measure patients' quality of life and symptoms [12]. Fourteen items use a four-point response scale, ranging from "not at all" to "very much." The final and 15th item concerns a rating of the overall quality of life during the past week, with response options on a seven-point scale ranging from 1 "very poor" to 7 "excellent."

Table 2 – Example for the translation of the relative validity into a percentage of sample size reduction

Because the RV is calculated as $RV = \frac{t^2(EF10)}{t^2(EF4)}$, if RV = 1.21 for the E10 vs the EF4, it means that the *t* test for the EF10 was ($\sqrt{1.21} = 1.1$) 1.1 times that of the EF4, or equivalent with $SD(EF10) = \frac{SD(EF4)}{1.1}$. We used this SD ratio to calculate an estimate of the required sample for the EF10 compared with the EF4. As an example, if the EF4 with N = 128 had power = 80% at $\alpha = 5\%$ to detect a specific difference, then it can be calculated from the SD ratio using standard sample size calculations that the EF10 would need N = 106 to obtain the same power, or 106/128 = 83% of the sample of the EF4.

We used power = 80%, $\alpha = 5\%$, and effect size (ES) = 0.5 in the calculations. The expected savings for any combination of power and ES will be similar to those presented here, except in very extreme cases with very low power or large ES.

Coping

Patients completed three scales, that is problem-focused coping, acceptance, and avoidance (denial) of the COPE and brief COPE [22,23]. The 12 questions have a four-point response scale.

Satisfaction with care

Items from the European Organization for Research and Treatment of Cancer Inpatient Satisfaction with care (EORTC IN-PATSAT) Questionnaire were used to assess patients' satisfaction with care, their appraisal of hospital doctors and nurses, and aspects of care organization and services [24]. The questions have a five-point response scale.

Experience of cancer care

The "Assessment of Patients' Experience of Cancer Care" assesses patients' perceptions of the quality of their cancer care [25]. To assess medical decision making of the treatment and care, we selected five items measuring the quality of the medical decision making. Responses are given on a five-point scale.

Patient involvement

We developed four questions on patient involvement in treatment and care and the awareness of relatives and physicians of the patients' wishes and preferences, which can be answered on a five-point scale.

Results

Patient characteristics

Characteristics of the 1028 participants in the study are presented in Table 3. The majority of the sample was male (60%), married (70%), and living in a private household (95%). Most patients had lung cancer stage IV (45%) or colorectal cancer stage IV (29%). Most had a World Health Organization performance status of 1 (52%).

Relative validity

The results of the known-group comparisons of the three scales (EF10, EF4, EF2) are summarized in Table 4. The table presents the *t* statistics and the RV per known-group comparison as well as the mean RV for the comparisons of the EF10 with the EF4, the EF10 with the EF2, and the EF4 with the EF2, respectively.

Of the 53 pairs of known groups considering the EF10 and the EF4, 45 showed a statistically significant difference for at least one of the two measures. For these 45 pairs we calculated the RV for the EF10 compared with the EF4 and found that the RV ranged from 0.47 to 3.71. In 41 out of the 45 (91%) comparisons, the RV was above 1, indicating that the EF10 had a higher RV than did the EF4. The mean RV for the comparison of the EF10 with the EF4 was 1.41, with a bootstrap-based 95% confidence interval of 1.28 to 1.57. Transforming this confidence interval to sample size requirements resulted in a 20% to 34% reduction in sample size without loss of power.

When calculating the RV of the EF10 in comparison with that of the EF2, we found significant differences in EF in 45 of 53 known-group comparisons; 44 of these 45 comparisons were also significant when comparing the EF10 with the EF4. The RV was calculated for these 45 pairs, and the EF10 had a higher RV than the EF2 in 40 out of 45 (89%) significant comparisons. The

Table 3 – Sociodemographic and clinical characteristics of the study sample (N = 1028)

Characteristic*	Value
Age (y), mean ± SD	65.91 ± 9.85
Range	18-91
Sex, n (%)	
Male	592 (60.2)
Country of residence, n (%)	
Belgium	180 (18.1)
Denmark	134 (13.6)
Italy	147 (14.9)
The Netherlands	229 (23.2)
Slovenia	74 (7.5)
The United Kingdom	218 (22.1)
Civil status, n (%)	
Married, civil partnership	683 (70.1)
Divorced, separated	110 (11.3)
Widowed	96 (9.8)
Unmarried	86 (8.8)
Living with partner, n (%)	717 (74.1)
Living conditions	
Private household	915 (94.4)
Institutionary care	6 (0.6)
Other	48 (5.0)
Having children, n (%)	848 (86.8)
Education	
Years of education, mean ± SD	13.17 ± 4.59
Range	0-45
Interquartile range	10.0, 13.0, 16.0
Religion, n (%)	
Religious	491 (50.5)
Not religious	353 (36.3)
Prefers not to specify	128 (13.2)
Ethnicity, n (%)	
No ethnic minority	943 (99.0)
Cancer type and current stage, n (%)	
Lung cancer (stage III or IV)	534 (57.6)
Colorectal cancer (stage IV or metachronous metastases)	393 (42.4)
Time since diagnosis (y), mean ± SD	1.4 ± 1.82
Range	0–11
Time since diagnosis of current stage (y), mean ± SD	0.85 (1.16)
Range	0–9.69
Current treatment, n (%)†	
Chemotherapy	301 (29.3)
Targeted therapy	866 (84.2)
Immunotherapy	897 (87.3)
Radiotherapy	921 (89.6)
WHO performance status, n (%)	
0: fully active	318 (34.4)
1: restricted in physically strenuous activity	486 (52.6)
2: ambulatory and capable of all self-care	107 (11.6)
3: capable of only limited self-care	13 (1.4)

WHO, World Health Organization.

* Missing: Age (n = 11), sex (n = 6), country of residence (n = 7), civil status (n = 14), living with partner (n = 21), living conditions (n = 20), having children (n = 12), education (n = 130), religion (n = 17), ethnicity (n = 36), cancer type and stage (n = 62), WHO performance status (n = 65).

† Several options may apply.

mean RV for the EF10 compared with that for the EF2 was 1.74 (95% CI 1.48–2.10). This would allow for a 31% to 52% reduction in sample size when using the EF10 instead of the EF2, without loss of power.

Comparison of the EF4 with the EF2 revealed higher RV of the EF4 in 36 out of 45 (80%) significant comparisons. These 45 comparisons differed from the previous comparisons (Table 4). The mean RV for the comparison of the EF4 with the EF2 was 1.16 (95% CI 1.11–1.26). Based on this confidence interval, using the EF4 instead of the EF2 would allow for a 9% to 20% reduction in sample size without loss of power.

Discussion

One of the most powerful implications of using CAT is the more precise and efficient estimate of a patient's quality of life by tailoring the items to the patient's individual characteristics. The aim of this study was to compare the RV of the customized EF10 with the original validated EF4 using the method of known-groups comparison.

The results of the study confirm our hypothesis that the EF10 has a better RV compared with those of the EF4 and the EF2. This means that the EF10 has superior power to detect differences between groups, which allows for a smaller sample size to detect differences without reducing power. Ninety-one percent of the known-group comparisons indicated higher RV of the EF10 compared with the EF4, with an average RV of 1.41. This allows for a smaller sample size of about 20% to 34% when using the EF10 instead of the EF4. This indicates that clinical trials having EF as primary outcome and using the EF10 instead of the EF4 scale of the QLQ-C30 can be carried out with considerably smaller sample sizes without loss of power.

The EF10 had a higher RV in 89% of the comparisons with the EF2. We also found that the EF4 had a higher RV than did the EF2. Although this finding was expected and hypothesized when the EF2 was developed 11 years ago [12,26], it had not been confirmed empirically.

Mapping the quality of life of patients is important for the improvement of care and to inform treatment decisions. Quality of life is often an important outcome measure in randomized clinical trials and observational studies. The savings in sample size that can be obtained by using the customized short-form EF10 are particularly important when the study concerns a vulnerable group of patients, such as patients with advanced cancer, because the reduced required sample sizes can address the often-encountered problems with reaching target numbers in studies. Many palliative care trials have failed because of problems with recruitment [27,28]. An additional benefit is that when fewer study participants are required while maintaining the same power, innovative findings may become ready for implementation sooner.

Using a customized short form has many benefits, yet it comes with the costs of the complex development (e.g., the creation of an item bank or the selection of items based on the samples' characteristics). Besides, although the items might lead to less missing values, because the questions are more applicable to the respondents' situation, the questionnaires tend to be somewhat longer, which is adding some burden to respondents. Therefore, as is generally the case in patient-reported outcome assessments, it is important to weigh the practical gain (i.e., increased power) of a longer, more precise measure against minimizing the burden to respondents. Note that we aimed here to improve measurement precision by adding relevant items, but customized short forms can also be used to form shorter measures if, for example, the minimum response burden has priority.

Table 4 – Relative validity of the three scales to assess emotional functioning

Patient characteristics	Scale			Relative validity		
	EF10 t statistic	EF4 t statistic	EF2 t statistic	EF10/EF4	EF10/EF2	EF4/EF2
Age*	1.43	2.09	1.54	0.47	–	1.84
Sex	3.90	3.72	4.31	1.10	0.82	0.75
Spouse†	–	–	–	–	–	–
Children: yes/no†	–	–	–	–	–	–
Education	2.06	3.01	3.02	0.47	0.46	0.99
Religious	−4.47	−4.43	−3.81	1.02	1.38	1.35
Lung cancer vs colorectal cancer	−3.04	−2.94	−2.86	1.07	1.13	1.06
Stage†	–	–	–	–	–	–
Chemotherapy yes/ no†	–	–	–	–	–	–
Time since diagnosis	2.99	2.27	2.16	1.73	1.92	1.11
Time in stage	3.43	2.88	2.51	1.41	1.86	1.31
WHO status	4.69	3.95	4.07	1.41	1.33	0.94
Quality of life (EORTC QLQ-C15-PAL)						
Physical functioning: trouble taking a short walk	−10.69	−9.07	-8.19	1.39	1.70	1.23
Physical functioning: need to stay in bed or a chair	−8.53	−6.63	-6.43	1.65	1.76	1.06
Physical functioning: help with eating, dressing, washing	−7.34	−7.00	−6.69	1.10	1.20	1.09
Dyspnea	−8.65	−7.41	−6.66	1.36	1.68	1.24
Pain	−8.21	−8.12	−7.28	1.02	1.27	1.24
Insomnia	−8.29	−7.95	−7.76	1.09	1.14	1.05
Fatigue	−13.70	−11.04	−10.54	1.54	1.69	1.10
Lack of appetite	−11.78	−9.82	−9.29	1.44	1.61	1.12
Nausea	−10.25	−9.10	−8.77	1.27	1.37	1.08
Constipation	−5.63	−5.21	−5.54	1.17	1.03	0.88
Tiredness	−14.88	−13.20	−13.46	1.27	1.22	0.96
Interference due to pain	−11.31	−10.65	−10.42	1.13	1.18	1.04
Overall quality of life	15.41	12.80	12.51	1.45	1.52	1.05
Coping (COPE)						
Active: efforts on doing something about it	3.20	2.91	3.01	1.21	1.13	0.94
Acceptance: accepting the reality	5.44	4.65	4.47	1.37	1.48	1.08
Planning: coming up with a strategy	2.66	2.99	3.72	0.79	0.51	0.65
Active: taking action	3.34	2.92	2.90	1.31	1.33	1.01
Denial: acting as though it hasn't happened†	–	–	–	–	–	–
Denial: saying "this isn't real"	−3.08	−3.19	−3.42	0.94	0.81	0.87
Denial: pretending this hasn't happened‡	−1.84	−1.88	−1.98	–	0.86	0.90
Acceptance: learning to live with it	8.15	6.40	6.13	1.62	1.77	1.09
Planning: thinking about what steps to take†	–	–	–	–	–	–
Denial: refusing to believe that it has happened	−3.36	−2.78	−2.73	1.46	1.51	1.04
Acceptance: getting used to the idea	4.94	4.28	3.51	1.33	1.98	1.48
Acceptance: accepting that it has happened	4.43	3.92	3.43	1.28	1.67	1.30
Satisfaction with care (IN-PATSAT)						
Information about illness (by doctors)	5.50	4.42	3.62	1.55	2.31	1.49
Information about medical tests (by doctors)	5.90	4.53	3.94	1.70	2.24	1.32
Information about treatment (by doctors)	5.82	4.90	3.97	1.41	2.15	1.52
Information about medical tests (by nurses)	5.79	4.88	4.68	1.41	1.53	1.09
Information about care (by nurses)	5.46	4.14	3.68	1.74	2.20	1.26
Information about treatment (by nurses)	5.94	4.83	4.27	1.51	1.93	1.28
General rating of received care	6.71	4.95	4.39	1.84	2.34	1.27
Experience of cancer care (APECC)						
Detailed discussions	6.88	5.80	5.09	1.41	1.83	1.30
Concerns/questions	6.73	5.07	4.78	1.76	1.98	1.13
Preferred option	7.31	6.32	5.58	1.34	1.72	1.28
Work out differences	6.24	5.29	4.52	1.39	1.90	1.37
Responsible for final decision	5.34	4.36	3.58	1.50	2.22	1.48
Patient involvement						
Friends are aware of wishes§	3.40	1.77	1.20	3.71	8.03	–
Doctors are aware of wishes*	–	–	–	–	–	–
Involvement as preferred	5.16	3.79	3.61	1.85	2.04	1.10
Great influence on care	3.65	2.26	1.97	2.60	3.43	1.32

continued on next page

Table 4 – *continued*

Patient characteristics	Scale			Relative validity		
	EF10 *t* statistic	EF4 *t* statistic	EF2 *t* statistic	EF10/EF4	EF10/EF2	EF4/EF2
Mean ratio				1.41	1.74	1.16
95% CI				1.28–1.57	1.48–2.10	1.11–1.26

APECC, Assessment of Patients' Experience of Cancer Care; COPE; EORTC IN-PATSAT, the European Organization for Research and Treatment of Cancer Inpatient Satisfaction with care (EORTC IN-PATSAT) Questionnaire; Quality of Life Questionnaire Core 15 Palliative Care (QLQ-C15-PAL); WHO, World Health Organization.

[*] Not significant for the comparison of E10 with E2 ($P > 0.05$).

[†] Not significant for any comparison ($P > 0.05$).

[‡] Not significant for the comparison of E10 with E4 ($P > 0.05$).

[§] Not significant for the comparison of E4 with E2 ($P > 0.05$).

The EORTC CAT has been designed for international use in patients with cancer, and we tested it for the first time in a large sample of patients with advanced cancer in various European countries. This makes generalizability of the results possible. Another strength of this study was the high number of known-group comparisons (53 pairs, of which 45 were significant), which makes the findings robust and reduces the risk that conclusions are influenced by chance findings. In addition, we calculated the RV only when the *t* test for at least one of the measures being compared (EF10, EF4, or EF2) showed a significant difference.

Although this study used cross-sectional data, extending and replicating our approach based on longitudinal ACTION data will be possible in due course. We used the EORTC item bank on EF, which has the structure of the original QLQ-C30 scale with respect to the number of response options and phrasing of items. One might consider this a limitation in comparison to, for example, the PROMIS item banks that were developed from scratch, while aiming for the optimal way to address concepts. However, because of its design (being a new measurement system), the PROMIS item banks do not allow comparison to a pre-existing, validated instrument measuring exactly the same construct, using the same wording. Instead they need to compare the RV to alternative instruments measuring a similar construct. In our study, we were able to assess the RV of the EF10 to a validated instrument measuring exactly the same construct.

Conclusions

We found that the customized EF10 based on the EORTC CAT item bank performs better than the EF4 in detecting differences in EF between groups of patients with advanced cancer. Compared with the EF4, the EF10 showed superior power, allowing a 20% to 34% smaller sample size without reducing power, when used as a primary outcome measure.

Acknowledgments

This publication is based on the ACTION project (Advance care planning – an innovative palliative care intervention to improve quality of life in oncology), conducted by a collaboration of research teams from the Netherlands, Denmark, Belgium, Slovenia, the United Kingdom, and Italy. We thank all participating patients and relatives, facilitators, trainers, hospital staff, and the Advisory Board for their valuable contribution to this project. We thank Bud Hammes and Linda Briggs for their advice throughout the project.

Source of financial support: The ACTION trial and thus this study were supported by the European Union's 7th Framework Programme for Research and Technological Development (FP7) (Proposal No. 602541-2).

Supplementary Materials

Supplementary data associated with this article can be found in the online version at https://doi.org/10.1016/j.jval.2018.07.002.

REFERENCES

[1] Zebrack B, Burg MA, Vaitones V. Distress screening: an opportunity for enhancing quality cancer care and promoting the oncology social work profession. J Psychosoc Oncol 2012;30:615–24.

[2] Turk DC, Dworkin RH, Allen RR, et al. Core outcome domains for chronic pain clinical trials: IMMPACT recommendations. Pain 2003;106:337–45.

[3] Smith AB, Wright EP, Rush R, et al. Rasch analysis of the dimensional structure of the Hospital Anxiety and Depression Scale. Psychooncology 2006;15:817–27.

[4] Cull A, Gould A, House A, et al. Validating automated screening for psychological distress by means of computer touchscreens for use in routine oncology practice. Br J Cancer 2001;85:1842–9.

[5] Wainer H. Computerized Adaptive Testing: A Primer. 2nd ed. Mahwah, NJ: Erlbaum; 2000.

[6] Petersen MA, Gamper EM, Costantini A, et al. An emotional functioning item bank of 24 items for computerized adaptive testing (CAT) was established. J Clin Epidemiol 2016;70:90–100.

[7] Hambleton RK, Swaminathan H, Rogers HK. Fundamentals of Item Response Theory. Sage: Newbury Park; 1991.

[8] van der Linden WK, Hambleton RK. Handbook of Modern Item Response Theory. Berlin: Springer-Verlag; 1997.

[9] Gamper EM, Groenvold M, Petersen MA, et al. The EORTC emotional functioning computerized adaptive test: phases I-III of a cross-cultural item bank development. Psychooncology 2014;23:397–403.

[10] Fayers P, Bottomley AEORTC Quality of Life Group. Quality of Life Unit. Quality of life research within the EORTC - the EORTC QLQ-C30. Eur J Cancer 2002;38:S125–33.

[11] 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.

[12] Groenvold M, Petersen MA, Aaronson NK, et al. The development of the EORTC QLQ-C15-PAL: a shortened questionnaire for cancer patients in palliative care. Eur J Cancer 2006;42:55–64.

[13] Rietjens JA, Korfage IJ, Dunleavy L, et al. Advance care planning–a multi-centre cluster randomised clinical trial: the research protocol of the ACTION study. BMC Cancer 2016;16:264.

[14] Oken MM, Creech RH, Tormey DC, et al. Toxicity and response criteria of the Eastern Cooperative Oncology Group. Am J Clin Oncol 1982;5:649–55.
[15] Fayers PM, Aaronson NK, Bjordal K, et al. The EORTC QLQ-C30 Scoring Manual. 3rd ed. Brussels: European Organisation for Research and Treatment of Cancer; 2001.
[16] Groenvold M, Petersen MAon behalf of the EORTC Quality of Life Group. Addendum to the EORTC QLQ-C30 Scoring Manual: Scoring of the EORTC QLQ-C15-PAL 2006.
[17] Scott NW, Fayers PM, Aaronson NK, et alon behalf of the EORTC Quality of Life Group. EORTC QLQ-C30 Reference Values 2008.
[18] Fayers PM, Machin D. Quality of Life—The Assessment, Analysis and Reporting of Patient-Reported Outcomes. 3rd ed. Chichester, West Sussex: John Wiley & Sons, Ltd.; 2016.
[19] Wenzel LB, Fairclough DL, Brady MJ, et al. Age-related differences in the quality of life of breast carcinoma patients after treatment. Cancer 1999;86:1768–74.
[20] Deng N, Allison JJ, Fang HJ, et al. Using the bootstrap to establish statistical significance for relative validity comparisons among patient-reported outcome measures. Health Qual Life Outcomes 2013;11. https://doi.org/10.1186/1477-7525-11-89.
[21] Petersen MA, Aaronson NK, Arraras JI, et al. The EORTC computer-adaptive tests measuring physical functioning and fatigue exhibited high levels of measurement precision and efficiency. J Clin Epidemiol 2013;66:330–9.
[22] Carver CS, Scheier MF, Weintraub JK. Assessing coping strategies: a theoretically based approach. J Pers Soc Psychol 1989;56:267–83.
[23] Carver CS. You want to measure coping but your protocol's too long: consider the Brief COPE. Int J Behav Med 1997;4:92–100.
[24] Brédart A, Bottomley A, Blazeby JM, et al. An international prospective study of the EORTC cancer in-patient satisfaction with care measure (EORTC IN-PATSAT32). Eur J Cancer 2005;41:2120–31.
[25] Arora NK, Reeve BB, Hays RD, et al. Assessment of quality of cancer-related follow-up care from the cancer survivor's perspective. J Clin Oncol 2011;29:1280–9.
[26] Bjorner JB, Petersen MA, Groenvold M, et al. Use of item response theory to develop a shortened version of the EORTC QLQ-C30 emotional functioning scale. Qual Life Res 2004;13:1683–97.
[27] Buss MK, Arnold RM. Challenges in palliative care research: one experience. J Palliat Med 2004;7:405–7.
[28] Ewing G, Rogers M, Barclay S, et al. Recruiting patients into a primary care based study of palliative care: why is it so difficult? Palliat Med 2004;18:452–9.