International psychometric validation of an EORTC quality of life module measuring cancer related fatigue (EORTC QLQ-FA12)

Joachim Weis[1], Krzysztof A.Tomaszewski[2], Eva Hammerlid[3], Juan Ignacio Arraras[4], Thierry Conroy[5], Anne Lanceley[6], Heike Schmidt[7], Susanne Singer[8] ,Monica Pinto[9], Mohamed Alm El-Din[10], Inge Compter[11], Bernhard Holzner[12], Dirk Hofmeister[13], Wei-Chu Chie[14], Marek Czeladzki[15], Amelie Harle[16], Louise Jones[17], Sabrina Ritter[18], Markus Wirtz[19], Hans-Henning Flechtner[20], Andrew Bottomley[21]

On Behalf of the EORTC Quality of Life Group

1 Tumor Biology Centre University Clinic Centre Freiburg, Freiburg, Germany
2 Faculty of Medicine, Jagiellonian University Medical College, Krakow, Poland
3 Department of Otolaryngology Head and Neck Surgery, Sahlgrenska University Hospital, Gothenburg, Sweden
4 Oncology departments. Complejo Hospitalario de Navarra. Pamplona. Spain
5 Department of Medical Oncology, Lorraine Cancer Institute, Nancy, Nancy, France
6 Institute for Women's Health, University College, London, UK
7 Institute of Health and Nursing Science, Medical Faculty, Martin Luther University Halle, Halle, Germany
8 Division of Epidemiology and Health Services Research, Institute of Medical Biostatistics, Epidemiology and Informatics, University Medical Centre Mainz, Mainz, Germany
9 Rehabilitation Unit, Department of Health Policy, Farmacy and Quality of Life, Istituto Nazionale Tumori "Fondazione Giovanni Pascale"- IRCCS - Naples, Italy.
10 Department of Clinical Oncology, Tanta Faculty of Medicine, Egypt
11 Department of Radiation Oncology (MAASTRO), GROW (School for Oncology & Developmental Biology), University of Maastricht Medical Center, Maastricht, The Netherlands
12 Department of Psychiatry and Psychotherapy, Medical University of Innsbruck, Austria
13 Department of Medical Psychology and Medical Sociology, University of Leipzig, Germany
14 Department of Family Medicine, National Taiwan University, Taiwan
15 Research and Development Office, Southern Health NHS Foundation Trust, Southampton, UK
16 Medical Oncology Department, The Christie NHS Foundation Trust, Manchester, UK.
17 HPB Services Directorate of Surgery, University Hospital Aintree, Liverpool, UK
18 Institute of Research in Rehabilitation Medicine at Ulm University, Ulm, Germany
19 Department of Research Methods, University of Education, Freiburg, Germany
20 Department of Paediatric Psychiatry, University of Magdeburg, Magdeburg, Germany
21 Quality of Life Department, EORTC Headquarters, Brussels, Belgium

Address of correspondence:

Joachim Weis Prof. PhD.
University Clinic Centre Freiburg
Tumor Biology Clinic
Department of Psycho-Oncology
Breisacher Str. 117
79106 Freiburg
Phone 0049-(0)761-206-2220

Fax: 0049-(0)761-206-2258
weis@tumorbio.uni-freiburg.de

Abstract

Background

The European Organisation for Research and Treatment of Cancer (EORTC) Group has developed a new multidimensional instrument measuring cancer related fatigue to be used in conjunction with the quality of life core questionnaire (EORTC QLQ-C30). The paper reports on the results of an international psychometric validation of the EORTC QLQ fatigue module, which is a multidimensional tool for assessing physical, cognitive and emotional aspects of cancer related fatigue.

Methods

The methodology follows the EORTC guidelines for phase IV validation of modules. We used a complex design assessing data in four cohorts of patients with a prospective longitudinal data collection. For validation and cross-validation confirmatory factor analysis (maximum likelihood estimation) was employed. Furthermore, sensitivity to change and test-re-test reliability have been examined. The study involved an international multi-centre collaboration of eleven European and Non-European countries.

Results

A total sample of 946 patients with various tumour diagnoses and in various stages of their disease were enrolled. Based on the confirmatory factor analysis, we assigned one item from one to another sub-dimension and removed one item to improve the scale structure, resulting in the EORTC QLQ-FA12. In addition, acceptable sensitivity to change as well very good test-retest reliability has been proven.

Conclusion

The EORTC QLQ-FA12 is now available as a validated phase IV module with excellent psychometric characteristics, and can be used as a robust instrument for measuring cancer related fatigue in international clinical trials, in daily clinical routine or in quality assurance.

Keywords

Cancer related fatigue; cancer, international field validation, quality of life; module development; confirmatory factor analysis

Background

Fatigue is one of the most distressing symptoms for cancer patients affecting their quality of life in all phases of the treatment or stages of the disease. Cancer related fatigue (CrF) is commonly defined as a self-recognised phenomenon that is subjective in nature and experienced as a feeling of tiredness or lack of energy that varies in degree, frequency and duration [1]. From a patient's perspective, fatigue is described as an unusual feeling of exhaustion, weakness or a loss of activity with sequels to emotional and cognitive functions [2, 3], which, in general, can not be reduced by sleep or rest. Fatigue is the most frequent symptom occurring in cancer patients during or after medical treatment and also as a long term late effect. Prevalence rates ranging from 59%-100%, whereas fatigue as a long term sequelae or late effect is estimated to have an average prevalence rate of approximately 30%, dependent on the type of assessment and diagnostic criteria used [4,5,6]. Due to an increased interest and research output in CrF, more detailed uni- or multidimensional instruments have been developed to assess CrF (7,8]. While many of the fatigue scales have strengths and limitations, actually there are no clear recommendations which measure is the most appropriate. Using a standardized questionnaire allows clinicians to measure CrF in the course over time and allows comparisons between various patient subgroups. The EORTC QLQ-FA13 module has been developed following the methodological guidelines of EORTC, which include four phases of development (I. generation of issues, II. construction of items list III. pre-testing IV. field testing) [9]. The strengths of the questionnaires developed by EORTC Quality of life Group lie in an international multi-center approach following high methodological standards and multicultural applicability. The pre-tested module EORTC QLQ-FA13 (phase III) is based on a multidimensional concept of fatigue including 13 items (2 global items on interference with daily activities and social sequelae of fatigue and 11 items allocated to a physical, emotional and cognitive domain) (see appendix A). It has been designed to measure fatigue in conjunction with the quality of life core questionnaire EORTC QLQ-C30 [10].

Aims and purpose

According to the EORTC Quality of Life Group guidelines [9], the purpose of phase IV of the module development is the evaluation of the psychometric characteristics and the validity of the EORTC QLQ-FA13 in an international sample of tumour patients. The psychometric validation includes the evaluation of the scale structure of the EORTC QLQ-FA13 using confirmatory analyses. The main hypotheses for the factorial structure to be tested were the following:

1. The factorial structure of the EORTC QLQ-FA13 phase III module may be replicated according to the underlying theory.

2. The global items (endogenic variables) FA12 ("did fatigue interfere with your daily activities") and FA13 ("did you have the feeling that fatigue was not understood by people close to you") are predicted adequately by the sub-scales physical, cognitive and emotional fatigue.

In addition we analysed test-retest reliability, internal consistency and the responsiveness to change.

Design and Methods

The design for the psychometric evaluation of the EORTC QLQ-FA13 followed the guidelines of EORTC for the development of modules in phase IV. Patients were enrolled in four distinct groups as following (see table 1):

- group A: Cancer patients with first-line treatment with curative intention ($t1_A$-$t3_A$)
- group B: Cancer patients with second-line treatment with palliative intention ($t1_B$-$t3_B$)
- group C: Cancer patients off treatment (≥12 and ≤ 18 months since end of treatment and no evidence of cancer disease or recurrence) ($t1_C$,$t2_C$)
- group D: Cancer patients off treatment (≥ 36 and ≤ 72 months (*survivors*) and no evidence of cancer disease or recurrence ($t1_D$,$t2_D$)

Insert table 1 here

The patient's questionnaire includes the EORTC core questionnaire (EORTC QLQ-C30 version 3.0) and the fatigue module EORTCQLQ-FA13 in all four groups. In addition, for patients of the groups A and B, a global screening of Cancer Related Fatigue (CRF) was used to determine an initial global fatigue score according to the NCCN guidelines 2015. All patients provided clinical and socio-demographic data (gender, date of birth, country of origin, marital status, education level, employment status). Time since diagnosis, tumour locations, type of treatments, metastases, ECOG Performance Status, and time since completion of treatment (only for groups C and D) were taken from the medical records at each cooperating center.

The study was carried out as an international multicentre study including 17 centres in 11 European and Non-European countries (Europe: England, France, Germany, Austria, Poland, Netherlands, Sweden, Spain, Italy, Non-Europe: Egypt and Taiwan). For phase IV the EORTC QLQ-FA13 has been translated to the languages of the cooperating countries. The translations were carried out in close cooperation with the translation team of the Quality of Life Department of EORTC following the translation guidelines of EORTC QoL group [11]. We initially calculated a sample of 135 per group (over all group a total of n=520 patients [12]. Due to higher drop out rates than expected, we increased recruitment and sample size up to the final sample of n=946 patients. Patient recruitment was from February 2011 to November 2014. The study was registered with the German Clinical Trial Studies Registry (DRKS-ID: DRKS00003091). National and local ethics approvals were obtained for the recruiting centers before commencement of this study.

Inclusion and exclusion criteria

Patients with cancer of all tumour sites were included if they met the following criteria: Histologically confirmed cancer, with written informed consent and the ability to understand the

language of the questionnaire. Patients had to have an absence of severe psychiatric or cognitive mental conditions potentially hampering compliance with the study protocol and follow-up schedule, and all patients had to be aged over 18 years. Patients undergoing allogeneic hematological stem cell transplantation (HSCT) or neoadjuvant therapy were excluded. Patient could not participate in other quality of life studies that might interfere with this validation study.

Statistical methods

Data entry, management and statistical evaluation were conducted from the coordination centre in Freiburg. Data analysis was done using SPSS v21 supported from a biostatistician specialized in psychometric analyses. For the confirmatory factor analysis, AMOS 21.0 (Analysis of Moment Structures, maximum likelihood method) was used. Missing values (<13) were replaced with stochastic-regression-based imputation available in AMOS 21.0 [13]. This estimation procedure uses information within the available data information, to avoid biases in the analysed variance-covariance matrix. Hence, the analysed information is not affected, if data are missing because of missing-at-random processes [14]. The assumption of normal distribution was checked by Mardia test in AMOS 21.0 and corrected in case of violation using the Bollen-Stine-Bootstrap strategy [15]. Item characteristics are described in terms of acceptance (percentage of responders), item discrimination (corrected item scale correlation) and item difficulty (mean score). Reliability was determined by score for internal consistence (Cronbach's alpha) [15]. The a priori defined factor structure was checked by the discrepancy between the empirical based variance-covariance matrix and the variance-covariance matrix based on the model. For the evaluation of the model, indices of global and local fitness were used. Chi-square test was used for the statistical testing of the model. Further indices for the goodness of fit were Root-Mean-Square-Error of Approximation (RMSEA) and Goodness of Fit Index (GFI) as absolute fit-Indices. Normed Fit Index (NFI), Tucker-Lewis Index (TLI) und Comparative Fit Index (CFI) were used as measures for incremental fit and the Adjusted Goodness of Fit Index (AGFI). RMSEA

scores <.08 indicates an acceptable fit, score <.05 a good model fit. GFI scores >.95 are indicating a good model fit, scores >.90 an acceptable model fit [13]. For the estimation of the reliability of the single items, local fit indices were calculated. For convergent validity, the indicator reliability should be > .4 factor reliability > .6 and DEV (mean variance) > .5 [16]. For sensitivity to change we used MANOVA with standardized effect sizes (ηp^2 = partial eta-squared) analysing the data of group A and B separately as differences between the two groups are to be expected. For univariate comparisons we used t-tests and for interpretation of the effect sizes we followed Cohen (1988) (d<0.1 = small effect; d<0.3 = medium effect; d> 0.5 = large effect). For test-re-test reliability we used intra-class correlation with the data of group C and D.

Results

Description of the sample

The total sample of patients recruited in all groups was n=946. The average age was 58.7 (sd 13.1 years) (range from 22-97 years). Patients were recruited in Germany (16.1%) and Poland (15.6%) followed by UK (11.1%), Sweden (10.0%), Egypt (9.9%), Spain (8.4%), Netherlands (7.2%), Italy (5.0%), Austria (4.5%) and Taiwan (3.8%). Gender distribution was balanced (female 54.1 %, male 45.9%). The sample comprised of a wide spectrum of tumour diagnoses with the highest percentages in breast cancer (24.0%) head and neck cancer (22.6%), lung cancer (11.1%) and colorectal cancer (9.5%) (for further details see table 2).

Insert table 2 here

Confirmatory factor analysis

We use confirmatory analysis (CFA) to check the three factorial model of the EORTC QLQ-FA13 including the $t1_{ABCD}$ measurement of patients of all four groups (A to D) (n=944) (see figure 1). We used the two global items as indicators of the three latent constructs, as these two

items are not underlying the latent construct but represent potential effects of fatigue on daily and social life (see appendix A).

The results of the CFA show that the underlying model of the EORTC QLQ-FA13 was not adequately reproduced by the data both for global (table 3) and local fit indices (table 4).

Insert tables 3 and 4 here

The indices in table 3 show that the thresholds for acceptable model fit were not reached. The residual covariance suggest that item FA05 (“Did you have trouble getting things started?“) is indicative for the physical domain (increase of factor loading from 70 to .79) (table 3). Item FA05 is more closely associated to the dimension of physical fatigue which includes items of reduced physical functions and loss of performance. Therefore, the allocation of item FA05 from the cognitive to the physical dimension is acceptable. Furthermore, Item FA11 (“Did you have trouble completing things?”) was insufficiently associated with the cognitive fatigue construct. Additionally, elimination of item FA11 caused considerable lower correlation between the latent construct in the model: the correlation between the cognitive and emotional fatigue decreased from r = .74 to .58, as well as between cognitive and physical fatigue from r = .73 to .66. Hence, after deleting the hybrid item FA11 from the scale the discriminability of the three constructs is noticeably enhanced and all fit indices and factor loading (all ≥ .72) could be substantially improved, respectively. Especially, the Bayesian Information criterion (BIC: 424.521), taking model parsimony into account, indicates the three-factorial model structure as the most appropriate.

Figure 1 for the previous model and figure 2 for the revised model show a graphical illustration of these results.

Insert figure 1 and 2 here

To check the stability and generalisation of the structure equation model, a cross-validation of the results of the patients of all four groups at the time measurement $t2_{ABCD}$ was conducted. The

sample for cross validation included n= 643 patients. The results of the cross validation analysis show nearly identical model fit (AGFI= 0.921; RMSEA= 0.070; CFI= 0.970; indicator reliabilities from .52 to .81). Furthermore, the model estimations for the data of $t2_{ABCD}$ or $t3_{AB}$ confirmed the revised model for EORTC FA12 (see table 3). Given these results, the structure equation model could be conclusively replicated both in the cross-validation sample and based on the t2 resp. t3 data.

In terms of construct validity, patients with distant metastases had a significant higher score of overall fatigue than patients without distant metastases (M (metastases) = 2.09; M (no metastases = 1.69; t (262.4) = 6.93; p<.001; d= 0.63). Patients undergoing radiotherapy alone have a significant lower overall fatigue compared with patients with others therapies (M (radiotherapy) = 1.71; M (others) = 1.83; t (714) = -2.39; p = 0.017; d = 0.20). There was no significant difference between patients undergoing combined adjuvant therapy vs. monotherapy (t (714) = -1.53; p = 0.127). As we hypothesised, we found a significant correlation of the fatigue scores with the ECOG Score for all subscales (physical Fatigue: rs = .481; p<.001; Emotional Fatigue: rs = .322; p<.001; Cognitive Fatigue: rs = .328; p<.001). Post-hoc comparison of scores (Tukey Test) showed that patients of the group B (palliative treatment) have the highest fatigue scores compared with all other groups (MB = 2.10; MA = 1.75; MC = 1.69; MD = 1.67). In addition, there is a high negative correlation between fatigue in all subscales and the global quality of life (r = -.672; p<.001) showing the higher the fatigue the lower the global quality of life.

Sensitivity to change

For the analysis of sensitivity to change, we used data from all points of measurements (t1, t2, t3) of the subgroups A (curative treatment) and B (palliative treatment). Time between t1 and t2 was on an average of 6.1 weeks for group A and 5.6 weeks for group B. Time difference between t1 and t3 was on an average of 20.1 weeks for group A and 13.8 weeks for group B.

In group A (curative treatment), fatigue scores for all subdimensions increased up to t2 and slightly decreased at t3. We found for all subdimensions significant changes ($p<.05$) with a small to medium effect size for physical fatigue ($\eta p^2 = 0.19$) (Eta squared coefficient) and small effect sizes for the emotional ($\eta p^2 = 0.11$) and cognitive fatigue ($\eta p^2 = 0.09$) (table 5). Furthermore, there was a highly significant multivariate effect of time ($p <.001$) and highly significant interaction effect ($p<.001$). For group B (palliative treatment) we found no significant changes for the subdimensions physical fatigue and emotional fatigue with only very small effects sizes of $\eta p^2 = 0.03$ resp. 0.06) (table 6). Only for the subscale cognitive fatigue we found a significant change comparing t2 and t3 ($p<.05$) with only a small effect size ($\eta p^2 = 0.06$). Similarly, no multivariate effects (time or interaction of dimension and time) could be detected. The overall fatigue scores in all subdimensions were higher in the palliative group than in the curative group.

Insert table 5 and 6 here

Test-Retest-Reliability

Test-retest reliability was tested using the data of patients of group C and D (n=410). The time difference between $t1_{CD}$ and $t2_{CD}$ was an average of 9.3 days (SD= 7.41). As the results of both subgroups were very similar, we report the results for both groups together (see table 7). The correlations (Intra Class Correlation) for $t1_{CD}$ and $t2_{CD}$ in both groups show high scores for all three subdimensions ranging from $r = .90$ to .92 indicating a high stability of measurement. In addition, internal consistency was good with Cronbach's alpha ranging from 0.79 to 0.88 for $t1_{CD}$ and 0.82 to 0.89 for $t2_{CD}$ (table 7).

Insert table 7 here

Discussion

We conducted a comprehensive psychometric validation of the EORTC fatigue module FA13. As a result of our study, we present a slightly revised phase IV module, the EORTC QLQ-FA12.

The international cross-cultural validation of this module, including a large and representative sample of cancer patients, allows generalization of the results and guarantees the cross-cultural applicability of this module, in line with the EORTC tradition. We used a confirmatory factor analysis to validate the a-priori three dimensional structure of EORTC QLQ-FA13 in conjunction with the two global items as criteria. The results show that the previous phase III module EORTC QLQ-FA13 did not reach a sufficient model fit for all items and therefore had to be modified. The changes include the elimination of a single item (FA11) and the allocation of one item (FA05) to the physical dimension instead of the cognitive dimension. The inter-correlation of FA11 within the factorial structure show that this item may be not sufficiently understood as part of the cognitive dimension. In addition, item FA11 did not allow a clear allocation to the three dimensions and shows low factor loading in all three factors ≤.35. As there is only a minor loss of information we decided to eliminate this item. The allocation of item FA05 to the physical dimension improved the model fit substantially. In total, by these changes the model could be improved and we attained very good scores for the model fit. A cross-validation of the data and the replication of the model for the t2 and t3 measurements confirmed the results of the revised model. In terms of convergent and divergent validity, all coefficients for the model fit showed very good to excellent fit.

The two global items (FA12 and FA13) were used as criteria variables. Item FA12 was predicted by the physical dimension, FA13 by the cognitive dimension. These results are in line with the model assumptions, although we expected the prediction of these two items by all three dimensions. As a result, acceptable to very good scores for the internal reliability (Cronbachs α from .79 to .90) were found.

High correlations between the subscales of EORTC FA13 with sociodemographic or medical parameters confirmed the convergent validity and are line with the research literature [2,3,6,17].

Analysis of test-retest reliability was conducted in two groups of patients both off treatment. The results show a high correlation for all fatigue scores between $t1_{CD}$ and $t2_{CD}$ with an average time difference of nine days which indicates a stable measurement of fatigue by the EORTC QLQ-FA12 over a time where no changes of fatigue are to expect.

The analyses of the sensitivity to change detected different results for the patients in curative treatment (group A) compared with patients under palliative treatment (group B). In group A, we found small effects in all subdimensions of EORTC QLQ-FA12 for the pairwise comparisons of all three points of measurement (before treatment up to the end of treatment). In this group we covered a time distance from t1 to t3 of more than 5 months. In group B there was only a very small effect in the subdimension of cognitive fatigue. In all other dimensions we found no significant change over time. From a clinical point of view, all these results make sense, especially in patients in palliative care, who are suffering mostly from long lasting fatigue which may be not changed during ongoing treatment [18]. In addition, the time distance of the palliative care group was with 3.5 months shorter than in group A.

Our study had some limitations. First, we could not achieve equal sample sizes for all cooperating countries, as the patient recruitment proceeded at different levels of recruitment in the various countries. Also, the recruitment in the palliative care group (group B) was more difficult to both recruit and when included we had drop out rates higher than initially planned. We were able to compensate losses in statistical power this, to some extent, by increasing the initial patient numbers and therefore we do not think this had any major impact on the findings of our analysis. As drop out is associated with disease related aspects, potential biases resulting from not-completely at random (MCAR) or randomly (MAR) missing data processes [15] may affect estimates or sensitivity for change and retest-reliability.

In conclusion, the EORTC QLQ-FA12 is now available as an internationally validated phase IV module to be used for measuring cancer related fatigue in conjunction with the EORTC QLQ-C30, but there is still a need for further analysis of the sensitivity to change especially in palliative care patients. EORTC QLQ-FA12 may be also used to assess fatigue symptoms in clinical routine or quality assurance to assess care needs. The module is currently available in the following languages: English, Dutch, German, Polish, Italian, French, Spanish, Swedish, Norwegian, Arabic and Mandarin and is available from the EORTC Quality of Life Department.

Acknowledgement

The phase IV has been funded by the EORTC QoL Group. We thank all collaborators and the patients helping us to develop this module.

Conflict of Interest Statement

All authors declare that there are no conflicts of interest in terms of the contents of this publication.

Role of the Funding Source

This work was sponsored by the European Organisation for Research and Treatment of Cancer Quality of Life Group (EORTC QoL Group) (grant number: v1 04/09). The sponsor has no influence or other type of involvement neither in the study design nor in the collection of data, the analysis nor the interpretation of the data. The EORTC QoL executive committee as the scientific sponsor has reviewed the manuscript before submission.

References:

1. NCCN (National Comprehensive Cancer Network). *Clinical practice guidelines in oncology: cancer related fatigue. Version 2.2015*. 2015 http://www.nccn.org/professionals/physician_gls/pdf/fatigue.pdf

2. Henry DH, Viswanathan HN, Elkin EP, et al. Symptoms and treatment burden associated with cancertreatment: results from a cross-sectional survey in the U.S. *Support Care Cancer, 2008,16:*791-801.

3. Wagner LI, Cella D. Fatigue and cancer: causes, prevalence and treatment approaches. *Br J Cancer, 2004, 91:*822-828.

4. Bower JE, Bak K, Berger A et al. Screening, Assessment, and Management of Fatigue in Adult Survivors of Cancer: An American Society of Clinical Oncology Clinical Practice Guideline Adaptation. *JCO*, 2014, 32, DOI: 10.1200/JCO.2013.53.4495

5. Servaes P, Gielissen MF, Verhagen S, Bleijenberg G. The course of severe fatigue in disease-free breast cancer patients: a longitudinal study. *Psychooncology, 2007,* 16:787-795.

6. Arndt V, Stegmaier C, Ziegler H, Brenner H. A population-based study of the impact of specific Symptoms on Quality of Life in Woman with Breast Cancer 1 Year After Diagnosis. *Cancer,* 2006, 107(10):2496-2503.

7. Minton O, Stone P. A systematic review of the scales used for the measurement of cancer-related fatigue (CRF). *Annals of Oncology, 2009,* 20:17–25.

8. Ahlberg K, Ekman T, Gaston-Johansson F, Mock, V. Assessment and management of cancer-related fatigue in adults. *Lancet, 2003, 362:* 640–650.

9. Johnson, C. D., Aaronson, A., Blazeby, J. M. et al. Guidelines for Developing Questionnaire Modules (4th Edition). Brussels EORTC 2011. http://groups.eortc.be/qol/sites/default/files/archives/guidelines_for_developing_questionnaire-_final.pdf

10. Weis J, Arraras JI, Conroy T, et al. Development of an EORTC quality of life phase III module measuring cancer-related fatigue (EORTC QLQ-FA13). *Psychooncology.* 2013, 22:1002–7.

11. de Wolf L, Koller M, Velikova G. et al. *EORTC translating procedures. Brussels: Quality of Life Study Group Publications.* EORTC Publication .ISBN 978-2-930064-40-6, 2009.

12. Tabachnik BG, Fidell LS. *Using multivariate statistics*. Harper & Row, London, 2001.

13. Arbuckle, J. *IBM SPSS Amos™ 21 User's Guide* (Amos Development Corporation, Ed.) 2012.

14. Graham, J. W., Cumsille, P. E., Elek-Fisk, E. Methods for Handling Missing Data. In I. B. Weiner (Ed.), *Handbook of Psychology.* Hoboken, NJ, USA: John Wiley & Sons, 2003.

15. Kline, R. B. *Principles and practice of structural equation modeling* (Methodology in the social sciences 3rd ed). New York: Guilford Press 2011.

16. Hair, J. F. *Multivariate data analysis* (7th ed). Upper Saddle River, NJ: Prentice Hall 2010.

17. Stone, P., Richards, M., A'Hern, R. & Hardy, J. (2000). A study to investigate the prevalence, severity and correlates of fatigue among patients with cancer in comparison with a control group of volunteers without cancer. *Annals of Oncology,*2000*,11* (5), 561–567.

18. Radbruch L, Strasser F, Elsner F et al. Fatigue in palliative care patients - an EAPC approach. *Palliat Med*, 2008, 22:13-32.

Table 1 Overview: Design and analyses of the four groups

	Group A	Group B	Group C	Group D
Points of measurement				
t1	$T1_A$ ±7 days before or at the 1st day of treatment (adjuvant chemo-/radiotherapy) (n = 311)	$T1_B$ ±7 days before or at the 1st day of treatment (adjuvant chemo-/radiotherapy) (n = 222)	$T1_C$ after completion of any treatment for at least 12 respectively (n= 212)	$T1_C$ after completion of any treatment for at least 36 months (n=199)
t2	$T2_A$ ChTh: at the end of 2nd cycle or at the beginning of the 3rd cycle; Radioth.: at the end of the 4th week of radiotherapy (n= 279) (drop out: 10.3%)	$T2_B$ ChTh: at the end of 2nd cycle or at the beginning of the 3rd cycle; Radioth.: at the end of the 4th week of radiotherapy (n = 181) (drop out: 18.5%)	$T2_C$ re-test one week after first assessment (n=201) (drop-out: 5.1%)	$T2_C$ re-test one week after first assessment (n = 187) (drop-out: 6%)
t3	$T3_A$ at three months (range: 12-15 weeks) after treatment (n=243) (drop-out: 12.9%)	$T3_B$ at one month (range: 4-6 weeks) after treatment. (n =141) (drop-out: 22.1%)	-	-
Analyses				
	Confirmatory Analyses	Confirmatory Analyses	Confirmatory Analyses	Confirmatory Analyses
	Sensitivity to change	Sensitivity to change	Test Re-Test Reliability	Test- Re-Test Reliability

Table 2 Sociodemographic and medical data

Total n=946 (at t1 $_{ABCD}$)	**Total Sample** n = 946 (100%)	**Group A Curative treatment** n=311 (32.9%)	**Group B Palliative Treatment** n=222 (23.5%)	**Group C off treatment <12 months** n=212 (22.4%)	**Group D off treatment >36 months** n=199 (21.1%)
Age Mean (SD) (N=943)	58.8 (13.1)	59.3 (14.0)	62.7 (11.6)	58.7 (12.9)	59.0 (13.5)
Range	22-97	26-87	31-97	28-90	25-95
Sex					
Female	512 (54.1%)	157 (50.5 %)	118 (53.2)	125 (59%)	110 (55.3%)
Male	434 (45.9%)	154 (49.5%)	104 (46.8%)	87 (41.0%)	89 44.7%)
ECOG Score (N= 861; 91.0 %)					
0 Fully active	399 (46.3%)	151 (48.6%)	40 (18.0%)	99 (46.7%)	108 (54.3%)
I Restricted	303 (35.2%)	103 (33.1%)	94 (42.3%)	54 25.5%)	52 (26.1%)
II Self care	121 (14.0%)	35 (11.3%)	58 (26.1%)	15 (6.6%)	14 (7.0%)
III Limited self care	29 (3.4%)	8 (2.6%)	18 (8.1%)	1(0.5%)	2 (1.0%)
IV Completely disabled	9 (1.1%)	1 80.3%)	5 (2.3%)	1 (0.5%)	2 (1.0%)
Metastases (N= 903; 95.6 %)					
No	671 (74.3%)	275 (88.4%)	35 (15.8%)	190 (89.6%)	169 (84.9 %)
Yes	231 (25.7 %)	31 (10.0%)	178 (80.2%)	11 (5.2%)[1]	12 (6.0%)[1]
Location of tumour					
Breast	227 (24.0%)	60 (19.3%)	41 (18.5%)	75 (35.4%)	49 (24.6%)
Head/Neck	214 (22.6%)	90 (28.9%)	32 (14.4%)	48 (22.6%)	44 (22.1%)
Lung	105 (11.1%)	42 (13.5%)	50 (22.5%)	8 (3.8%)	5 (2.5%)
Colorectal	90 (9.5%)	33 (10.6%)	23 10.4%)	17 (8.0%)	17 (8.5%)
Prostate	61 (6.4%)	16 (5.1%)	5 (2.3%)	14 (6.6%)	26 (13.1%)
Gynaecological	61 (6.4%)	18 (5.8%)	16 (7.2%)	13 (6.1%)	14 (7.0%)
Haematological	49 (5.2%)	20 (6.4%)	7 (3.2%)	12 (5.7%)	10 (5.0%)
Testicular	14 (1.5%)	3 (1.0%)	1 (0.5%)	4 (1.9%)	6 (3.0%)
Pancreatic	13 (1.4%)	1 (0.3%)	12 (5.4%)	0 (0.0%)	0 (0.0%)
Others	118 (12.5%)	30 (9.6%)	48 (21.8%)	15 (7.1%)	25 (12.6%)
Treatment (multiple choice)					
Surgery	599 (63.3%)	159 (51.1%)	112 (72.5%)	176 (83.0%)	150 (75.4%)
Chemotherapy	528 (55.8%)	178 (57.2%)	161 (50.5%)	103 (48.6%)	85 (42.7%)
Radiotherapy	528 (55.8%)	141 (45.3%)	136 (61.3%)	129 (60.8%)	120 (60.3%)
Hormone	83 (8.8%)	8 (2.6%)	17 (7.7%)	32 (15.1%)	25 (12.6%)
Others	59 (6.2%)	11 (3.5%)	22 (9.9%)	11 (5.2%)	15 (7.5%)

[1] Due to the definition of groups C and D patients with metastases were excluded from the test-re-test analyses

Table 3: Gobal indices of convergent and divergent validity of the revised model (n=944)

(data of $t1_{ABCD}$)

	c^2	p	df	c^2/df	GFI	AGFI	NFI	RMSEA	CFI	TLI
Thresholds for acceptable fit		>0.05		<3	≥ 0.90	≥ 0.90	≥ 0.90	≤ 0.08	≥ 0.90	≥ 0.90
Thresholds for good fit		>0.05		<3	≥ 0.95	≥ 0.95	≥ 0.95	≤ 0.05	≥ 0.95	≥ 0.95
EORTC QLQ-FA12										
Original model t1 (EORTC-FA13)	788.8	<0.001	58	13.60	0.888	0.824	0.894	0.116	**0.901**	0.866
Modified model t1 (EORTC-FA12)	205.1	<0.001	47	4.36	**0.965**	**0.942**	**0.970**	**0.060**	**0.976**	**0.967**
Modified model t2	257.5	<.001	47	5.48	.957	.929	.966	.069	.972	.961
Modified model t3	342.4	<.001	47	7.28	.946	.911	.960	.82	.965	.951

GFI= Goodness of fit index, AGFI= adjusted goodness of fit index, NFI= normed fit index, RMSEA= root mean square error of approximation, CFI= comparative fit index, TLI= Tucker-Lewis index

Table 4 Local indices of convergent and divergent validity of the revised model (n=944) (data of $t1_{ABCD}$)

Factor	Item	Indicator-reliability	t-Value of factor loading	Factor-reliability	Average variance extracted	Cronbach's α
Thresholds for acceptable fit[1]		≥ 0.4		≥ 0.6	≥ 0.5	>.70
Physical	FA01 FA02 FA03 FA04 FA05[2]	.85 .87 .81 .73 .77	34.00 *** 30.57 *** 25.77 *** 28.03***	.90	.66	.90
Emotional	FA06 FA07 FA08[2]	.82 .83 .79	26.89 *** 25.44***	.85	.66	.85
Cognitive	FA09 FA10[2]	.87 .78	20.96 ***	.82	.70	.81

[1] following [16]; [2] = no t-values, fixed reference parameters to standardize the variance of the construct
*** = p <.001

Table 5 Sensitivity to change in Group A (curative treatment) (MANOVA, data of $t1_A$ to $t3_A$)

	t1		t2		t3				
Dimension	**M**	**SD**	**M**	**SD**	**M**	**SD**	**Pillai -Bartlett's V (df)**	**ηp²** (partial eta-squared)	**Pair wise comparison**
Physical Fatigue n=225	1.99	0.79	2.33	0.79	2.05	0.76	26.59* (2,230)	0.19	t1 & t2* t2 & t3*
Emotional Fatigue n =230	1.78	0.76	1.89	0.70	1.68	0.74	13.87* (2,230)	0.11	t1 & t2* t2 & t3*
Cognitive Fatigue n = 228	1.45	0.56	1.58	0.59	1.43	0.56	11.38* (2,230)	0.09	t1 & t2* t2 & t3*
Multivariate effect of time n = 228							21.28*** (2,229)	0.16	
Multivariate interaction effect of time and dimension n =228							9.72*** (4,227)	0.15	

* = p <.05, *** = p <.001

Table 6 Sensitivity to change in Group B (palliative treatment) (MANOVA, data of $t1_B$ to $t3_B$)

	t1		t2		t3				
Dimension	**M**	**SD**	**M**	**SD**	**M**	**SD**	**Pillai -Bartlett's V (df)**	**ηp²** (eta-squared)	**Pair wise comparison**
Physical. Fatigue n = 135	2.30	0.79	2.36	0.77	2.44	0.79	1.73 (2,230)	0.03	n.s.
Emotional Fatigue n = 132	2.09	0.79	2.03	0.74	2.15	0.84	1.99 (2,230)	0.03	n.s.
Cognitive Fatigue n = 134	1.65	0.58	1.63	0.49	1.77	0.65	3.99* (2,230)	0.06	t2 & t3*
Multivariate effect of time n=132							2.89 (2,229)	0.04	
Multivariate interaction effect of time and dimension n =132							1.4 (4,227)	0.04	

* = p <.05

Table 7 Internal reliability and intraclass correlation (data of $t1_{CD}$, $t2_{CD}$)

Dimension	Cronbachs α t1 (group C & D)		Cronbachs α t2 (group C & D)		Intraclass Corr (group C and D) mean score $R_{t1\text{-}t2}$
Physical Fatigue	(n = 410)	.88	(n = 386)	.90	.921***
Emotional Fatigue	(n = 402)	.87	(n = 388)	.88	.905***
Cognitive Fatigue	(n = 408)	.79	(n = 389)	.82	.907***

*** = p <.001