Psycho-Oncology
Psycho-Oncology **25**: 275–281 (2016)
Published online 22 July 2015 in Wiley Online Library (wileyonlinelibrary.com). **DOI**: 10.1002/pon.3907

Fatigue screening in breast cancer patients: identifying likely cases of cancer-related fatigue

Martine M. Goedendorp[1], Paul B. Jacobsen[2] and Michael A. Andrykowski[3]*
[1]*Department of Health Psychology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands*
[2]*Health Outcomes and Behavior Program, Moffitt Cancer Center, Tampa, FL, USA*
[3]*Department of Behavioral Science, College of Medicine, University of Kentucky, Lexington, KY, USA*

**Correspondence to: Department of Behavioral Science, College of Medicine, University of Kentucky, Lexington, KY, USA. E-mail: mandry@uky.edu*

Abstract

***Objective*: For clinical and research purposes, efficient identification of cases of cancer-related fatigue (CRF) is important, as CRF can be persistent and interfere with usual functioning. While various fatigue-screening instruments are available, no brief screening indices have been developed using formally diagnosed CRF cases as the criterion.**

***Methods*: Breast cancer patients ($n = 385$) completed a fatigue diagnostic interview and self-report fatigue measures (Profile of Mood States-fatigue subscale, Fatigue Symptom Inventory, and SF-36 vitality subscale), after initial adjuvant therapy (post-treatment (post-Tx) 1 assessment), after completion of radiotherapy for women receiving chemotherapy + radiotherapy (post-Tx 2 assessment), and 6 months after completion of all adjuvant therapy (6-month post-Tx assessment). CRF cases were identified using specific diagnostic criteria. ROC analyses identified screening indices, which could accurately identify CRF cases after initial adjuvant therapy. Screening indices were cross-validated using post-Tx 2 and 6-month follow-up assessment data.**

***Results*: A total of 104 women (27%) met CRF criteria after initial adjuvant therapy. Six two-item screening indices were identified. For all indices, area under the curve exceeded 0.80, sensitivity exceeded 0.80, and specificity exceeded 0.57. Cross-validation suggested that, except for the index based on SF-36, all the indices continued to accurately identify CRF cases at the post-Tx 2 and 6-month post-Tx assessments. Overall, a two-item composite index based on Fatigue Symptom Inventory 'most severity' and 'work interference' items performed best.**

***Conclusions*: Breast cancer patients and survivors meeting CRF diagnostic criteria can be accurately identified using brief screening indices derived from common self-report fatigue measures.**
Copyright © 2015 John Wiley & Sons, Ltd.

Received: 7 April 2015
Revised: 16 June 2015
Accepted: 16 June 2015

Background

Fatigue is frequently reported by cancer patients and survivors regardless of tumor type or treatment [1–3]. Fatigue can be highly distressing [4,5] and can have a profound impact on daily functioning [6]. According to the National Comprehensive Cancer Network, cancer-related fatigue (CRF) is a 'distressing, persistent, subjective sense of physical, emotional, and/or cognitive tiredness or exhaustion, related to cancer or cancer treatment that is not proportional to recent activity and interferes with usual functioning' [7].

Cancer-related fatigue prevalence rates vary depending on phase of the disease, treatment, and assessment instrument used [8–11]. During cancer treatment, fatigue prevalence rates range from 25% to 95% [12,13]. Additionally, cancer survivors can experience fatigue after treatment is completed, and again, prevalence rates vary greatly, ranging from 5% to 34% [10,13].

A variety of instruments have been used to assess fatigue severity. Some instruments have been specifically designed to assess fatigue, such as the Fatigue Symptom Inventory (FSI) [14], while other instruments include a subscale to assess fatigue within a larger instrument designed to assess broader domains of functioning, such as the Medical Outcome Study Short Form-36 Health Survey (SF-36) [15] and the Profile of Mood States (POMS) [16]. Overall, there is little consensus regarding the optimal instrument for assessing fatigue in cancer patients and survivors [11,17].

For clinical and/or research purposes, it is often important to identify individuals who are experiencing 'clinically significant' CRF. Consequently, efforts have been made to identify 'clinical' cutoff scores for measures of fatigue severity that could identify cases of clinically significant CRF. A cutoff score for the FSI has been identified, using scores on another fatigue measure, the SF-36 [18], as the reference point for defining clinically significant CRF. Two different cutoff scores have been identified for the four-item SF-36 vitality subscale: a cutoff score of ≤45, corresponding to the 25th percentile for women in the US general population [18] and a cutoff score of ≤50, representing the midpoint of the subscale [19]. To our knowledge, there is no cutoff score available to

Copyright © 2015 John Wiley & Sons, Ltd.

identify cases of clinically significant CRF using the POMS-fatigue subscale.

At present, a clinician or researcher interested in identifying clinically significant cases of CRF can choose from a variety of cutoff scores on common measures of fatigue such as the SF-36 and FSI. However, most existing choices suffer from a common flaw. Specifically, rarely, these cutoff scores were developed using the ability to accurately identify actual clinically significant cases of CRF as the criterion. Rather, cutoff scores were identified in reference to scores evidenced by the general population or in reference to an individual's score on another measure of fatigue severity. Furthermore, existing measures of fatigue severity, such as the SF-36 and POMS, lack information regarding other aspects of fatigue that are critical to defining its clinical significance. These include fatigue duration and impact on functioning. What is needed are clinical cutoff scores, which have been identified through their ability to accurately identify cancer patients and/or survivors who clearly evidence clinically significant CRF.

Clinically significant cases of CRF can be identified using a fatigue diagnostic interview and a set of specific diagnostic criteria consistent with the definition of CRF provided by the National Comprehensive Cancer Network [7,20]. The strength of this approach for identifying clinically significant cases of CRF is its ability to identify individuals who differ dramatically on a range of physical and psychological outcomes from individuals without CRF [21]. The weakness of this approach is the fatigue diagnostic interview can be time consuming and thus not easily employed in a research or clinical setting with limited resources.

Clinical treatments, such as psychosocial interventions and exercise, are available to manage CRF in cancer patients and survivors [22–24]. These treatments are typically time consuming and costly to implement. Hence, it is important to identify clinically significant CRF cases as accurately as possible, in order to target fatigue treatment to the individuals most in need. What is needed is a method for accurately and efficiently screening large groups of cancer patients and survivors for individuals likely to meet diagnostic criteria for clinically significant CRF. Therefore, the aim of the current study is to identify cutoff scores on brief, commonly used fatigue measures, SF-36, FSI, and POMS, that could be used to quickly and accurately identify cases of clinically significant CRF based on a fatigue diagnostic interview.

Methods

Participants

Participants were female breast cancer patients recruited as part of a larger quality of life study between November 1999 and May 2006 at two sites: Moffitt Cancer Center at the University of South Florida and the University of Kentucky. Eligibility criteria included the following: (a) ≥18 years; (b) stages 0–II breast cancer; (c) scheduled to receive adjuvant chemotherapy (CT), radiotherapy (RT), or both (CT + RT); (d) able to speak, read, and understand English; (e) no cancer history other than basal cell skin carcinoma; and (f) no chronic disease in which fatigue is a potentially prominent symptom.

Procedure

Procedures were approved by Institutional Review Boards at both study sites. Participants were recruited and informed consent obtained after breast surgery but prior to starting adjuvant therapy. Participants completed the following assessments: (a) before beginning adjuvant therapy (baseline); (b) at conclusion of an initial course of adjuvant therapy, either RT or CT depending on the participant's treatment plan (post-treatment (post-Tx) 1); (c) at conclusion of RT for women receiving CT + RT (post-Tx 2); and (d) 6 months after conclusion of adjuvant therapy (6-month post-Tx). Assessments consisted of a clinical interview and completion of questionnaires. The questionnaire portion of each assessment was completed in-person, by telephone, or by mail, as necessary. The clinical interview portion of each assessment was completed in-person or via telephone. Information regarding disease stage, surgery, and adjuvant therapy was obtained from medical records.

Questionnaire measures

Demographic information (age, race, partner status, and education) was obtained at baseline. Self-report fatigue measures (described in the succeeding text) were completed at each assessment.

The Medical Outcome Study Short Form-36 Health Survey (SF-36) [15] yields a vitality subscale score based on four items. Scores range from 0 to 100, with lower scores suggesting less vitality, which has been interpreted as fatigue [18,25].

The FSI [14] is a 14-item measure, which assesses fatigue frequency, severity, and interference. Items are rated on a 0–10 scale with higher ratings representing greater fatigue severity and interference.

The POMS-fatigue subscale (POMS-fatigue) [16] consists of seven items, rated on a 0–4 scale. Scores range from 0 to 28 with higher scores representing greater fatigue.

Finally, two items assessing fatigue severity and interference were created using the approach employed by Jacobsen *et al*. to develop items assessing nausea severity and interference [26,27]. Specifically, the word 'fatigue' was substituted for 'bodily pain' in the two items on the SF-36 bodily pain subscale. The first item was 'How much fatigue have you had during the past week'? Responses were made on a six-point scale ranging from 'none' to 'very severe'. The second item was 'During the past week, how much did

Copyright © 2015 John Wiley & Sons, Ltd.

DOI: 10.1002/pon

fatigue interfere with your normal work (including both work outside the home and housework)? Responses were made on a five-point scale ranging from 'not at all' to 'extremely'. For both items, higher scores represent greater fatigue severity or interference. These items were summed to create a composite index, fatigue severity + interference, with scores ranging from 2 to 11.

Clinical interview

Participants completed the mood, anxiety, and adjustment disorders modules from the Structured Clinical Interview for Diagnostic and Statistical Manual of Mental Disorders (4th ed.) (DSM-IV) [28] and the Diagnostic Interview Guide for CRF [20] in all assessments. The latter is a structured interview for determining whether a person meets specific criteria for a diagnosis of CRF [20]. All interviews were conducted by doctoral students in clinical psychology trained in administration and scoring of the clinical interview. Training involved practice interviews and review of audiotaped interviews. Research using similarly trained interviewers demonstrated high inter-rater agreement in CRF diagnosis [29].

Data analysis

Receiver operating curve (ROC) analyses were used to determine whether specific indices derived from self-report fatigue data collected at the post-Tx 1 assessment could distinguish accurately whether a woman met CRF diagnostic criteria at the post-Tx 1 assessment. ROC curves depict the trade-off between sensitivity (true positive rate) and specificity (true negative rate) for every possible cutoff score on a specific index. For any ROC curve, an area under the curve (AUC) can be calculated. The AUC estimates the overall discriminative value of each index with regard to the criterion, here, CRF caseness. An AUC value of 1.0 represents a test with perfect accuracy relative to the criterion. An AUC of 0.50 represents a test with accuracy no better than chance relative to the criterion. AUC values between 0.70 and 0.80 represent acceptable discrimination, while AUC values >0.80 represent excellent discrimination [30].

In addition to the two-item fatigue severity + interference composite index, one or more potential fatigue-screening indices were considered for each of the POMS, SF-36, and FSI measures. For the POMS, we considered the POMS-fatigue subscale score, and for the SF-36, we considered the SF-36 vitality subscale score. Several two-item composite indices using items from the POMS, SF-36, and FSI were also considered. These included the following: (a) sum of the ratings for two POMS-fatigue subscale items: 'fatigued' and 'exhausted' (POMS-fatigued + exhausted, range 0–8); (b) sum of the ratings for two FSI items: 'Rate your level of fatigue on the day you felt *most* fatigued in the past week' and 'Rate how much in the past week fatigue interfered with your normal work activity (includes both work outside the home and housework)' (FSI-severity + work interference, range 0–20); and (c) sum of the ratings for two SF-36 vitality subscale items: 'Did you feel worn out'? and 'Did you feel tired'? (SF-36 worn out + tired, range 2–12). Selection of specific items for these three, two-item indices was based on ROC and decision tree analyses.

For each of the six fatigue-screening indices considered, ROC analyses were conducted to determine AUC values and 95% confidence intervals. Furthermore, sensitivity, specificity, positive predictive value, negative predictive value, the number of false negatives (true cases of CRF not detected) and false positives (true non-cases of CRF detected), percentage of misclassified (sum of false positive and false negative cases divided by the total number of cases), and positive likelihood ratio (LR+) (true positives/false positives) were calculated for several potential cutoff values. LR+ values ≥2.00 are considered clinically important [31].

Results

A total of 419 women were enrolled and completed the baseline assessment. Less than 5% of study-eligible women declined participation. A total of 385 (92%) completed the post-Tx 1 assessment and constituted the final study sample. The 385 women in the final sample were compared with 34 women enrolled in the study not completing the post-Tx 1 assessment (study dropouts) on demographic (age, education, race, and partner status) and clinical characteristics (disease stage and treatment). No significant differences were found except for education: study dropouts were more likely to have a high school education or less (46%) than women in the final sample (27%) (X^2 (1) = 4.202; $p < 0.05$). Mean age in the final sample was 54.7 years (SD = 10.2; range 21–82 years), 26% possessed a high school education or less, 73% were partnered, and 90% were White, non-Hispanic. Stage of disease was stage 0 (11%), stage I (50%), and stage II (39%). Most women underwent lumpectomy (85%) with the remainder undergoing mastectomy (15%). Adjuvant therapy regimens included CT only (10%), RT only

Table 1. AUC values for CRF-screening indices

Index	AUC	95% confidence interval
POMS-fatigue subscale	0.806	0.759–0.853
POMS-fatigued + exhausted Items	0.818	0.772–0.864
SF-36 vitality subscale	0.816	0.771–0.862
SF-36 worn out + tired items	0.814	0.769–0.860
FSI-most fatigue + work interference items	0.831	0.774–0.860
Fatigue severity + work interference items	0.815	0.769–0.861

AUC, area under the curve; CRF, cancer-related fatigue; POMS, Profile of Mood States; SF-36, Medical Outcome Study Short Form-36 Health Survey; FSI, Fatigue Symptom Inventory.

Copyright © 2015 John Wiley & Sons, Ltd.

DOI: 10.1002/pon

Table 2. Data for optimal cutoff score for CRF-screening indices

Cutoff score	Sen	Spec	PPV	NPV	Percentage of misclassified (%)	LR+	False negatives[a]	False positives[b]
POMS-fatigue subscale								
≥7	0.845	0.580	0.404	0.924	37.8	2.01	12/100	130/276
POMS-fatigued + exhausted								
≥3	0.816	0.677	0.539	0.925	31.2	2.53	14/103	104/277
SF-36 vitality subscale								
<50	0.851	0.575	0.413	0.936	37.3	2.00	10/103	132/278
SF-36 worn out + tired								
≤6	0.858	0.569	0.398	0.944	39.6	1.99	8/102	142/277
FSI-most fatigue + work interference								
≥8	0.868	0.628	0.458	0.954	31.3	2.33	8/102	111/278
Fatigue severity + work interference								
≥6	0.806	0.678	0.455	0.916	31.5	2.50	16/103	104/278

CRF, cancer-related fatigue; Sen: sensitivity; Spec, specificity; PPV, positive predictive value; NPV, negative predictive value; LR+, positive likelihood ratio; POMS, Profile of Mood States; SF-36, Medical Outcome Study Short Form-36 Health Survey; FSI, Fatigue Symptom Inventory.

[a]*n* true cases of CRF missed/ n of true cases.

[b]*n* true non-cases meeting screening criterion/*n* of true non-cases.

Table 3. Cross-validation of optimal cutoff scores

Measure	Cutoff score	True positives	False positives	Percentage of misclassified
Post-Tx 2 assessment				
POMS-fatigue subscale	≥7	25/31 (81%)	42/108 (39%)	48/139 (35%)
POMS-fatigued + exhausted items	≥3	24/31 (77%)	39/108 (36%)	46/139 (33%)
SF-36 vitality subscale	<50	24/31 (77%)	32/107 (30%)	39/138 (28%)
SF-36 worn out + tired items	≤6	12/31 (39%)	15/107 (14%)	34/138 (25%)
FSI-most fatigue + work interference items	≥8	28/31 (90%)	31/108 (29%)	34/139 (24%)
Fatigue severity + work interference items	≥6	26/31 (84%)	35/108 (32%)	40/139 (29%)
6-month follow-up assessment				
POMS-fatigue subscale	≥7	23/27 (85%)	70/267 (26%)	74/294 (25%)
POMS-fatigued + exhausted items	≥3	23/27 (85%)	50/267 (19%)	54/294 (18%)
SF-36 vitality subscale	<50	20/26 (77%)	59/269 (22%)	65/295 (22%)
SF-36 worn out + tired items	≤6	14/25 (56%)	21/267 (8%)	32/292 (11%)
FSI-most fatigue + work interference items	≥8	26/26 (100%)	51/267 (19%)	51/293 (17%)
Fatigue severity + work interference items	≥6	22/27 (81%)	47/267 (18%)	52/294 (18%)

Tx, treatment; POMS, Profile of Mood States; SF-36, Medical Outcome Study Short Form-36 Health Survey; FSI, Fatigue Symptom Inventory.

(48%), and CT + RT (42%). Of women in the final sample, 139 received RT in addition to CT and thus completed a post-Tx 2 assessment (36%), and 295 women completed the 6-month post-Tx assessment (77%).

Determination of optimal cutoff scores for cancer-related fatigue-screening indices

One hundred four women (27%) met CRF criteria at the post-Tx 1 assessment. AUC values for the six CRF-screening indices are shown in Table 1. All AUC values were significantly greater than 0.750 and ranged from 0.806 (POMS-fatigue subscale) to 0.831 (FSI-most + interference).

There is no consensus approach to identifying an optimal cutoff score. Therefore, we selected an optimal cutoff score for each screening index based on consideration of the following: (a) proportion of false negative (i.e., 'true' CRF cases not identified) screening test results ≤15%, (b) proportion of women misclassified <40%, and (c) LR+ ≥2.00. Using these criteria, optimal cutoff scores were identified for each screening index: POMS-fatigue subscale (≥7), POMS-fatigued + exhausted composite index (≥3), SF-36 vitality subscale (<50), SF-36 worn out + tired composite index (≤6), FSI-most + work interference composite index (≥8), and fatigue severity + interference composite index (≥6). Table 2 displays results for the optimal cutoff value for the six screening indices.

Cross-validation of cutoff scores for cancer-related fatigue-screening indices at post-Tx 2 and 6-month follow-up assessments

The optimal cutoff score for each screening index was cross-validated using data from the post-Tx 2 and 6-month follow-up assessments. At the post-Tx 2 assessment, 31 of

Copyright © 2015 John Wiley & Sons, Ltd.

DOI: 10.1002/pon

the 139 women (22.3%) met criteria for CRF diagnosis. At the 6-month follow-up assessment, 27 of the 295 women (9.2%) met criteria for CRF diagnosis. Results of the cross-validation are shown in Table 3. Five of the six indices continued to perform well at both the post-Tx 2 and 6-month follow-up assessments. The exception was the SF-36 index, SF-36 worn out + tired. This index accurately identified only 39% of 'true' CRF cases at the post-Tx 2 assessment and only 56% of 'true' CRF cases at the 6-month follow-up assessment. For the remaining five indices at the post-Tx 2 assessment, the proportion of true positives ranged from 77% to 90%, while the proportion of false positives ranged from 29% to 39%. These five indices performed even better at the 6-month follow-up assessment. The proportion of true positives ranged from 77% to 100%, while the proportion of false positives ranged from 18% to 26%.

Conclusions

For clinical and research purposes, there is a need for an efficient and accurate method for identifying cases of clinically significant CRF. CRF-screening indices based on their ability to identify individuals meeting diagnostic criteria for CRF [20] have been developed before [32]. However, this approach has not been used to identify screening indices for commonly used instruments such as the SF-36, FSI, or POMS. Furthermore, prior research has generally identified screening indices based on lengthier, multi-item fatigue scales or subscales, while we have identified more efficient two-item screening indices and cross-validated these.

First, we determined cutoff scores for identifying likely cases of CRF using the POMS-fatigue and SF-36 vitality subscales and identified cutoff scores of ≥ 7 and < 50 as the most optimal for identifying cases of CRF. The cutoff score on the SF-36 vitality subscale has been used before [19] but until now has not been based on empirical evidence. Next, we determined cutoff scores for identifying CRF cases based on two-item indices derived from the POMS, SF-36, and FSI. Use of these brief indices may be advantageous when minimizing respondent burden is a concern.

Our results showed that each of our six screening indices demonstrated excellent ability to identify cases of CRF [30]. These indices identified successfully at least 85% of true cases of CRF. On the other hand, the proportion of false positives, that is, individuals who did not actually have CRF but who screened positive for CRF, ranged from 38% to 51%. In a clinical setting, these individuals would need to undergo further evaluation to determine whether CRF was indeed present. This is also suggested by Alexander *et al.* [32], who noted that neither the fatigue subscale of the Functional Assessment of Cancer Therapy nor the Bidimensional (Chalder) Fatigue Scale could be used as a case identifier for CRF, although the Bidimensional (Chalder) Fatigue Scale may be useful for screening. Our results indicate that each of the six screening indices would correctly eliminate approximately 50–65% of individuals screened from needing to undergo additional evaluation while at the same time missing less than 15% of true cases of CRF. We feel this is a reasonable balance. Of the screening indices identified, the two-item index based on FSI 'most fatigue' and 'work interference' items appeared to be most accurate. Sensitivity of this two-item FSI index was 0.87, the highest of all six indices. Overall, only 31% of individuals were misclassified with a false negative rate of 8% and a false positive rate of 40%. The accuracy of this two-item FSI index is not surprising, given it incorporates aspects of fatigue severity and impact on functioning, two cardinal attributes of CRF based on diagnostic criteria [20].

Importantly, we cross-validated the six screening indices we identified and found identified cutoff scores continued to perform well. The lone exception was the index based on the SF-36 items 'worn out' and 'tired'. Using this index, only 39% of true positives were identified at the post-Tx 2 assessment, and only 56% of true positives were identified at the 6-month follow-up assessment. In contrast, true positive rates for the other five indices exceeded 77% at both the post-Tx 2 and 6-month follow-up assessments. The FSI index based on the 'most fatigue' and 'work interference' items again appeared to be the most accurate approach to identifying CRF cases. This screening index identified 90% of true positives at the post-Tx 2 assessment with an overall 25% misclassification rate and identified 100% of true positives at the 6-month follow-up assessment with a 17% misclassification rate. Granted, cross-validating identified cutoff scores using data from the same sample but from different points in time from the derivation data is not an optimal approach. However, this represents the only attempt to cross-validate proposed CRF-screening indices we could find in the literature.

This study has other limitations we should acknowledge. The sample was fairly homogeneous as it was comprised primarily of White breast cancer survivors and women with household incomes in the middle to upper class range. Consequently, our findings may not be generalizable to male cancer survivors, individuals treated for cancers other than breast cancer, racial or ethnic minority cancer survivors, or survivors with lower income or low literacy.

Four of the six screening indices we identified were two-item indices. In addition to the obvious advantage of brevity, we expect that the accuracy of identifying cases of CRF might be enhanced when screening is based on a limited number of items rather than on responses to an entire questionnaire or a more lengthy set of items. When people complete questionnaires, they can make unintended mistakes, due to inattention or confusion. These mistakes happen more often in questionnaires with reverse-worded items than with questionnaires containing items posed in the same direction [33]. While the SF-36 vitality subscale does contain reverse-worded items, each

Copyright © 2015 John Wiley & Sons, Ltd.

DOI: 10.1002/pon

of the other five screening indices we identified consists only of items posed in the same direction.

In conclusion, previous research has identified a variety of screening indices for identifying clinically significant fatigue in cancer patients and survivors [18,19,32]. One study [32] has based identification of screening cutoff scores on a clinical diagnostic interview and specific diagnostic criteria for CRF, which have been adopted for inclusion in the International Statistical Classification of Diseases and Related Health Problems, Tenth Revision, Clinical Modification [34]. Thus, most previous research has largely ignored critical aspects of CRF such as impact on functioning and the presence of physical and psychological comorbidities in developing fatigue-screening indices. Our results show that it is possible to screen for cases of clinically significant CRF with good accuracy by using any of six, brief screening indices, four of which consist of only two items. In particular, an index using two items from the FSI appears to be particularly accurate. Using only two items to screen for CRF cases has advantages in clinical and research settings, as it is less time consuming for care providers, researchers, and patients.

Acknowledgement

This research was supported by an NIH grant R01CA82822 from the National Cancer Institute.

Conflict of interest

The authors have declared no conflicts of interest.

References

1. Prue G, Rankin J, Allen J, Gracey J, Cramp F. Cancer-related fatigue: a critical appraisal. *Eur J Cancer* 2006;**42**:846–863.
2. Hofman M, Ryan JL, Figueroa-Moseley CD, Jean-Pierre P, Morrow GR. Cancer-related fatigue: the scale of the problem. *Oncologist* 2007;**12**(Suppl 1):4–10.
3. Ahlberg K, Ekman T, Gaston-Johansson F, Mock V. Assessment and management of cancer-related fatigue in adults. *Lancet* 2003;**362**:640–650.
4. Sitzia J, Huggins L. Side effects of cyclophosphamide, methotrexate, and 5-fluorouracil (CMF) chemotherapy for breast cancer. *Cancer Pract* 1998;**6**:13–21.
5. Mulders M, Vingerhoets A, Breed W. The impact of cancer and chemotherapy: perceptual similarities and differences between cancer patients, nurses and physicians. *Eur J Oncol Nurs* 2008;**12**:97–102.
6. Minton O, Berger A, Barsevick A, *et al*. Cancer-related fatigue and its impact on functioning. *Cancer* 2013;**119**(Suppl 11):2124–2130.
7. National Comprehensive Cancer Network. Clinical practice guidelines in oncology. Cancer-related fatigue version 1.2014. http://www.nccn.org [Accessed June, 2014].
8. Seyidova-Khoshknabi D, Davis MP, Walsh D. Review article: a systematic review of cancer-related fatigue measurement questionnaires. *Am J Hosp Palliat Care* 2011;**28**:119–129.
9. Jean-Pierre P, Figueroa-Moseley CD, Kohli S, Fiscella K, Palesh OG, Morrow GR. Assessment of cancer-related fatigue: implications for clinical diagnosis and treatment. *Oncologist* 2007;**12**(Suppl 1):11–21.
10. Minton O, Stone P. How common is fatigue in disease-free breast cancer survivors? A systematic review of the literature. *Breast Cancer Res Treat* 2008;**112**:5–13.
11. Donovan KA, McGinty HL, Jacobsen PB. A systematic review of research using the diagnostic criteria for cancer-related fatigue. *Psycho-Oncology* 2013;**22**:737–744.
12. de Jong N, Courtens AM, Abu-Saad HH, Schouten HC. Fatigue in patients with breast cancer receiving adjuvant chemotherapy: a review of the literature. *Cancer Nurs* 2002;**25**:283–297.
13. Servaes P, Verhagen C, Bleijenberg G. Fatigue in cancer patients during and after treatment: prevalence, correlates and interventions. *Eur J Cancer* 2002;**38**:27–43.
14. Hann DM, Jacobsen PB, Azzarello LM, *et al*. Measurement of fatigue in cancer patients: development and validation of the Fatigue Symptom Inventory. *Qual Life Res* 1998;**7**:301–310.
15. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992;**30**:473–483.
16. McNair PM, Lorr M, Droppelman L. *POMS Manual*, Educational and Industrial Testing Service: San Diego, 1981.
17. Jacobsen PB. Assessment of fatigue in cancer patients. *J Natl Cancer Inst Monogr* 2004;**32**:93–97.
18. Donovan KA, Jacobsen PB, Small BJ, Munster PN, Andrykowski MA. Identifying clinically meaningful fatigue with the Fatigue Symptom Inventory. *J Pain Symptom Manage* 2008;**36**:480–487.
19. Bower JE, Ganz PA, Desmond KA, Rowland JH, Meyerowitz BE, Belin TR. Fatigue in breast cancer survivors: occurrence, correlates, and impact on quality of life. *J Clin Oncol* 2000;**18**:743–753.
20. Cella D, Peterman A, Passik S, Jacobsen P, Breitbart W. Progress toward guidelines for the management of fatigue. *Oncology (Williston Park)* 1998;**12**:369–377.
21. Andrykowski MA, Schmidt JE, Salsman JM, Beacham AO, Jacobsen PB. Use of a case definition approach to identify cancer-related fatigue in women undergoing adjuvant therapy for breast cancer. *J Clin Oncol* 2005;**23**:6613–6622.
22. Goedendorp MM, Gielissen MF, Verhagen CA, Bleijenberg G. Psychosocial interventions for reducing fatigue during cancer treatment in adults. *Cochrane Database Syst Rev* 2009; **1**: Art. No.: CD006953. DOI: 10.1002/14651858.CD006953.pub2
23. 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. *J Clin Oncol* 2014;**32**:1840–1850.
24. Cramp F, Byron-Daniel J. Exercise for the management of cancer-related fatigue in adults. *Cochrane Database Syst Rev* 2012; **11**: Art. No.: CD006145. DOI: 10.1002/14651858.CD006145.pub3
25. Brown LF, Kroenke K, Theobald DE, Wu J. Comparison of SF-36 vitality scale and Fatigue Symptom Inventory in assessing cancer-related fatigue. *Support Care Cancer* 2011;**19**:1255–1259.
26. Jacobsen PB, Le-Rademacher J, Jim H, *et al*. Exercise and stress management training prior to hematopoietic cell transplantation: blood and marrow transplant clinical trials network (BMT CTN) 0902. *Biol Blood Marrow Transplant* 2014;**20**:1530–1536.
27. Jacobsen PB, Meade CD, Stein KD, Chirikos TN, Small BJ, Ruckdeschel JC. Efficacy and costs of two forms of stress management training for cancer patients undergoing chemotherapy. *J Clin Oncol* 2002;**20**:2851–2862.
28. First MB, Gibbons M, Spitzer RL. *Users Guide for the Structured Clinical Interview for DSM-IV Axis I Disorders: Research Version*, Biometrics Research: New York, 1996.
29. Sadler IJ, Jacobsen PB, Booth-Jones M, Belanger H, Weitzner MA, Fields KK. Preliminary evaluation of a clinical syndrome approach to assessing cancer-related fatigue. *J Pain Symptom Manage* 2002;**23**:406–416.
30. Hosmer DW, Lemeshow S. *Applied Logistic Regression*, Wiley: New York, 1989.
31. Soreide K, Korner H, Soreide JA. Diagnostic accuracy and receiver-operating characteristics curve analysis in surgical research and decision making. *Ann Surg* 2011;**253**:27–34.

Copyright © 2015 John Wiley & Sons, Ltd.

32. Alexander S, Minton O, Stone PC. Evaluation of screening instruments for cancer-related fatigue syndrome in breast cancer survivors. *J Clin Oncol* 2009;**23**:6613–6622.

33. van Sonderen E, Sanderman R, Coyne JC. Ineffectiveness of reverse wording of questionnaire items: let's learn from cows in the rain. *PLoS One* 2013;**8**:1–7.

34. National Cancer Institute. Fatigue. 2014. http://www.cancer.gov/cancertopics/pdq/supportivecare/fatigue/HealthProfessional/page4#section_4.19 [Accessed 6 April 2015].

Copyright © 2015 John Wiley & Sons, Ltd.