Research and Theory for Nursing Practice: An International Journal, Vol. 20, No. 1, 2006

Self-Report Instruments for Fatigue Assessment: A Systematic Review

Dálete D. C. F. Mota, MSN, RN
Cibele A. M. Pimenta, PhD, RN
University of Sao Paulo
Sao Paulo, Brazil

This systematic review analyzed 18 self-reported fatigue instruments for adults. Five databases were searched combining fatigue with instrument, questionnaire, inventory, scale, or assessment. Eighteen fatigue instruments and six definitions of fatigue were found. Six instruments apply to physical or psychiatric disease; five are cancer-specific. Nine were unidimensional; others included intensity ($n = 10$), mental ($n = 7$), and physical ($n = 7$) dimensions. Eleven instruments had 15 or fewer items; the longest had 40 items. Four instruments were discriminative; 14 were evaluative. Fifteen assessed reliability using Cronbach's alpha, which was greater than or equal to .90 seven times. Validity tests were convergent or divergent ($n = 11$), discriminant ($n = 10$), and factor analysis ($n = 9$). Because fatigue is a highly prevalent limiting symptom, this review is important for improving fatigue assessment.

Keywords: measurement; scale; psychometric; tiredness; fatigue

Fatigue is a universal symptom because it is associated not only with chronic and acute illnesses, but also with normal everyday activities (Aaronson et al., 1999). Although there is no consensus on its definition, concept analysis provides the foundation for conceptualization of fatigue as a subjective, unpleasant symptom that incorporates both physical and psychological feelings and interferes with one's ability to conduct normal activities (Trendall, 2000).

Prevalence studies showed significant occurrence of fatigue, varying in intensity, in different population groups and illnesses. For the general population, the prevalence was approximately 20% (Kaasa, Loge, Knobel, Jordhou, & Brenne, 1999); among the elderly, more than 50% exhibited this symptom (Liao & Ferrel, 2000); among patients with post-polio syndrome, it varied between 79% and 89% (Packer, Sauriol, & Brouwer, 1994); for those with sleep disturbance, 78% (Lichstein, Melanie, Noe, & Aguillard, 1997); and for those with multiple sclerosis, 77% (Schwartz, Coulthard-Morris, & Zeng, 1996). Among patients with cancer, the prevalence was different according to the stage of the illness—around 40% at the moment of diagnosis or after the first cycle of chemotherapy (Passik, Kirsh, Rosenfeld, McDonald,

& Theobald, 2001), and more than 75% for those with advanced cancer (Conill et al., 1997; Esper & Redman, 1999; Jenkins, Schulz, Hanson, & Bruera, 2000; Stone et al., 1999).

In spite of the high frequency of fatigue in different populations, its presence is rarely investigated by health professionals (Detmar, Muller, Wever, Schornagel, & Aaronson, 2001). Detmar and associates observed that discussion of fatigue was always initiated by patients, and when they referred to feeling fatigued, professionals tended to ignore the referral because they did not know how to assess or manage the symptom.

Valid methods to assess fatigue are fundamental for its treatment. The subjectivity of the symptom reinforces the importance of self-report measures in capturing the patients' point of view, rather than the professional's (Mota & Pimenta, 2002). However, different instruments can generate different results, even when fatigue is assessed in the same person at the same moment (Brunier & Graydon, 1996). Few studies have reviewed extant instruments for fatigue and most that do included only a small number of instruments and analyzed multiple aspects of them (Aaronson et al., 1999; Brunier & Graydon, 1996; Kjerulff & Langenberg, 1995; Meek et al., 2000). In a recent review by Dittner, Wessely, and Brown (2004), 30 instruments were analyzed. However, some are generic measures, which include fatigue as one dimension, and others deal with concepts such as vigor, affect, or energy instead of fatigue.

The objective of the present study was to identify and evaluate the existing self-report instruments that exclusively assess fatigue in adults regarding their general characteristics and psychometric properties using a systematic review.

METHOD

The study is a systematic review of diagnostic studies. The method followed was recommended by Devillé et al. (2003) and Attalah and Castro (2003), and the steps consisted of research question formulation; identification, selection, and critical assessment of papers; and analyses of obtained data and interpretation of results.

Criteria for Considering Studies for Review

The articles reviewed were those that described the construction and presented the psychometric performance of the self-report instruments. Selected instruments were limited to those that assessed fatigue and no other symptom, that is, no generic instruments were included. Interventions for fatigue were not an exclusion criterion, because it could be used to assess responsivity. Language was limited to Portuguese, English, and Spanish. The subjects included in the studies were adult patients suffering from any physical or psychiatric illness, in any health care setting, or healthy adults. Considering that this present study is reviewing outcome measures, it is important to emphasize that only self-report instruments were included.

Search Strategy for Identification of Studies. The literature search covered the following databases and time periods: MEDLINE (1966–2003), LILACS (1982–2003), BDENF (1988–2003), WHOLIS (1948–2003), Cochrane (1993–2003), and PAHO (1902–2003).

The key words were fatigue combined with the terms instrument, questionnaire, inventory, scale, and assessment, in the following manner: "fatigue" AND ("instrument" OR "questionnaire" OR "inventory" OR "scale" OR "assessment").

Methods of the Review. The titles and abstracts of the 4,136 citations identified were carefully inspected by the principal reviewer (Mota). All citations were re-inspected by the same author and by the collaborator a second time. Eligibility criteria were used to select the relevant studies. No disagreement occurred regarding the studies to be included.

The assessment of methodological quality was evaluated by the description of the process to elaborate the instrument and its psychometric testing. If both of these aspects were present in the paper, the study was included. The included studies were read and analyzed, and the data were extracted using a standardized sequence.

The criteria used to analyze the instruments encompassed the following aspects: country of origin, year of publication, method to construct the instrument, definition of fatigue, target population, dimensions of fatigue, number and gradation of items, scores and cut-off scores, purpose of the instrument, reliability, and validity as described by the article's authors.

RESULTS

Twenty-one studies appeared to fit the inclusion criteria, but, on closer examination, only 16 described instruments that assess only fatigue, and were self-reported measures. The 16 articles included in this review presented 18 different instruments. Table 1 presents a summary of the characteristics of the 18 instruments.

General Characteristics

Eight instruments (44.4%) were designed in the United States, three (16.7%) in The Netherlands, two (11.1%) in the United Kingdom, two (11.1%) in Canada, two (11.1%) in Australia, and one (5.6%) in Japan.

Among the instruments identified in this review, the first self-report instrument exclusively for fatigue assessment in adults and with established psychometric properties was the Fatigue Severity Scale (Krupp, La Rocca, Nuir-Nash, & Steinberg, 1989). From 1998 to 2003, 11 instruments, 61.1% of the total, were constructed and published.

As shown in Table 1, a definition of fatigue was presented only in six instruments (33.3%). It is important to highlight that seven instruments (38.9%) did not use the word fatigue but a similar word instead, such as tiredness, weariness, exhaustion, weakness, lack of energy, or anergia.

The target population was identified in six instruments (33.4%) as the general population, which consists of people with any physical or mental illness, or no illness. Five instruments (27.8%) had cancer patients as the target population, and another five instruments (27.8%) did not specify the target population. One instrument (5.5%) was constructed for fatigue assessment in patients with Chronic Fatigue Syndrome, and one other (5.5%) for patients with multiple sclerosis and systemic lupus erythematosus.

TABLE 1. Main Characteristics of Fatigue Self-Report Instruments

Instrument (Abbreviation)	Country of Origin, Year of Publication	Concept	Number of Items/Scaling	Target Population	Purpose	Dimensions Assessed
Fatigue Severity Scale (FSS)	United States, 1989	N	9/7-point Likert-type scale	Multiple sclerosis, systemic lupus erythematosus	Evaluative	• Severity
Visual Analogue Scale for Fatigue (VAS-F)	United States, 1991	N	18/Visual Analogic Scale	Subjects with physical or psychiatric morbidities	Evaluative	• Severity
Fatigue Assessment Instrument (FAI)	United States, 1993	P	29/7-point Likert-type scale	Subjects with physical or psychiatric morbidities	Evaluative	• Severity • Situation-specific fatigue • Fatigue consequences • Responsiveness to rest/sleep
Brief Mental Fatigue Questionnaire (BMFQ)	United Kingdom, 1993	N	9/5-point Likert-type scale	Subjects with physical or psychiatric morbidities	Evaluative	• Mental
Fatigue Scale (FS)	United Kingdom, 1993	N	11/4-point Likert-type scale	Subjects with physical or psychiatric morbidities	Evaluative/ Discriminative	• Mental • Physical

Fatigue Impact Scale (FIS)	Canada, 1994	N	40/5-point Likert-type scale	Not specified	Evaluative	• Mental • Physical • Psychosocial
Multidimensional Fatigue Inventory (MFI)	Norway, 1995	N	20/7-point Likert-type scale	Not specified	Evaluative	• General fatigue • Mental • Physical • Reduced motivation • Reduced activity
Fatigue Symptom Inventory (FSI)	United States, 1998	N	13/Numeric Scale (0–10)	Cancer General population (afterwards)	Evaluative	• Severity • Impact • Duration
Dutch Fatigue Scale (DUFS)	The Netherlands, 1998	P	9/ Agree/ Disagree	Not specified	Discriminative	• Severity
Dutch Exertion Fatigue Scale (DEFS)	The Netherlands, 1998	P	9/5-point Likert-type scale	Not specified	Evaluative	• Severity
Revised Piper Fatigue Scale (Revised-PFS)	United States, 1998	N	22/Numeric Scale (0–10)	Cancer	Evaluative	• Severity/ Behavioral • Affective meaning • Sensory • Cognitive/Mood

(*Continues*)

TABLE 1. Continued

Instrument (Abbreviation)	Country of Origin, Year of Publication	Concept	Number of Items/Scaling	Target Population	Purpose	Dimensions Assessed
Schwartz Cancer Fatigue Scale (SCFS)	United States, 1998	P	28/5-point Likert-type scale	Cancer	Evaluative	• Physical • Emotional • Cognitive • Temporal
Brief Fatigue Inventory (BFI)	United States, 1999	N	9/Numeric Scale (0–10)	Cancer	Evaluative	• Severity
Cancer Fatigue Scale (CFS)	Japan, 2000	P	15/5-point Likert-type scale	Cancer	Evaluative	• Mental • Physical • Affective
Schedule Of Fatigue and Anergia for Chronic Fatigue Syndrome (SOFA-CFS)	Australia, 2000	N	10/5-point Likert-type scale	CFS	Discriminative	• Severity

Schedule Of Fatigue and Anergia for General Practice (SOFA-GP)	Australia, 2000	N	10/4-point Likert-type scale	Users of primary care settings	Discriminative	• Severity
Cancer-related Fatigue Distress Scale (CRFDS)	United States, 2000	N	20/Numeric Scale (0–10)	Cancer	Evaluative	• Distress
Fatigue Impact Scale for Daily Administration (FIS-D)	Canada, 2002	P	8/5-point Likert-type scale	Not specified	Evaluative	• Mental • Physical • Psychosocial

Note. P = present; N = not present.

Nine (50%) of the identified instruments were unidimensional, and most of them assessed intensity of fatigue (n = 7; 77.8%); one (11.1%) assessed intensity of distress, and one (11.1%) assessed intensity of mental fatigue. Of the multidimensional instruments, one (11.1%) assessed five dimensions, three (33.3%) assessed four dimensions, four (44.5%) assessed three dimensions, and one instrument (11.1%) assessed two dimensions. The instruments that assessed the greater number of dimensions, that is, that assessed fatigue in a more complete manner, were the Fatigue Assessment Instrument, Revised-Piper Fatigue Scale, Schwartz Cancer Fatigue Scale, and Multidimensional Fatigue Inventory. It is important to mention that the Revised-Piper Fatigue Scale is derived from the Piper Fatigue Scale, which was not included in this review because it was published in a source not captured in the search (Piper et al., 1989).

The instruments varied in length from eight to 40 items. Eleven instruments (61.1%) had 15 or fewer items, six instruments (33.3%) had between 16 and 30 items, and one instrument (5.6%) had 40 items.

The analysis of items' scaling showed that 12 instruments (66.7%) used a Likert scale, varying from 4 to 7 points. Four instruments (22.1%) had an interval numeric scale, varying from 0 to 10. One instrument (5.6%) used the Visual Analogic Scale (VAS) with two endpoints (not present/most severe fatigue), and one instrument (5.6%) had dichotomous scaling (present/not present).

Psychometric Properties: Reliability

All analyses of psychometric properties were based on the terminology proposed by Pasquali (1997), McDowell and Newell (1996), and Streiner and Norman (1995). The description of each psychometric property proposed by these authors is presented along with the results. A summary of the results of the psychometric tests applied for each instrument is presented in Table 2.

Of the 18 instruments, 14 (77.8%) had an evaluative purpose, that is, the instruments were constructed to measure the severity of the phenomena (Table 1). Three instruments (16.6%) had a discriminative purpose, that is, they were constructed to differentiate fatigued from nonfatigued individuals. One instrument (5.6%), the Fatigue Scale, had both purposes.

Reliability is a psychometric property that measures whether the instrument is assessing something in a reproducible way. One manner to assess reliability is measuring the internal consistency. The internal consistency was evaluated in 16 (88.9%) instruments, and the Cronbach's alpha was established in 15 (83.3%) of them (Table 2). The alpha value varied between .7–1.0 in all subscales of the instruments. Seven instruments (38.9%) had alpha values superior to .9 in all subscales. Only the Dutch Fatigue Scale did not calculate internal consistency by Cronbach's alpha, but the value obtained by the Kuder-Richardson test was also good (K-R 20 = .79)—since the range of values in the Kuder-Richardson formula are equal to that of Cronbach's alpha test when the instrument is dichotomous (Streiner & Norman, 1995).

Reliability was also analyzed by test-retest in 6 instruments (33.3%), testing the stability of the instrument (Table 2). The time interval between the application of the

TABLE 2. Synthesis of Psychometric Properties of the Self-Report Instruments for Fatigue Assessment

			Reliability		Validity					
Instrument	Subjects	Number of items/ Dimensions	Internal Consistency	Test-Retest	Content	Criterion	Discriminant	Divergent/ Convergent	Factorial Analysis	Responsivity
Fatigue Severity Scale (FSS) Krupp et al., 1989	25 multiple sclerosis		α = .81	*r* = .84 (*p* < .01)		*Concurrent*: FSS and Visual Analogic Scale (*r* = .68, *p* < .001) group	The FSS discriminated subjects with ME and SLE from control	• Weak correlation between FSS and depression measure (*r* = .26) • Positive correlation between FSS and depression measure (*r* = .46, *p* < .05)	Done for item reduction	*t* = 2.16, *p* < .01
	29 lupus (SLE)	9 Severity	α = .89							(*N* = 8)
	20 controls		α = .88	(*N* = 11)				• Weak correlation between FSS and depression measure (*r* = .20)		

(*Continues*)

TABLE 2. Continued

			Reliability		Validity					
Instrument	Subjects	Number of items/ Dimensions	Internal Consistency	Test-Retest	Content	Criterion	Discriminant	Divergent/ Convergent	Factorial Analysis	Responsivity
	57 sleep disturbance		α = .96/.95 (fatigue subscale) evening/ morning; α = .96/.94 (energy subscale) evening/ morning			*Concurrent*: VAS-F fatigue subscale and POMS fatigue subscale (r = .70, p < .01) VAS-F energy subscale and POMS energy/ vigor subscale (r = -.73, p < .01)		• Positive correlation between VAS-F fatigue subscale and quality of sleep measure (r = .65–.81; p < .01) • Weak correlation between VAS-F and POMS depression, tension and anger subscale (r = -.23–.35) • Positive correlation between VAS-F fatigue subscale and POMS confusion subscale (r = .41; p < .01)		
Visual Analogue Scale for Fatigue (VAS-F) Lee et al., 1991	75 controls	13 Fatigue subscale 5 Energy subscale	α = .91/.96 (fatigue subscale) evening/ morning; α = .94/.95 (energy subscale) evening/ morning							t > 3.6, p < .001 (control group)

						• Negative correlation between VAS-F energy subscale and POMS confusion subscale (r = -.40; p < .01)	
	35 Lyme disease; 46 chronic	Severity subscale	α = .92				
Fatigue Assessment Instrument (FAI) Schwartz et al., 1993	fatigue syndrome (CFS); 17 Post-Lyme disease; 47 Lupus; 40 Multiple sclerosis; 13 Dysthymia; 37 Controls	Situation-specific subscale	α = .77	r = .50–.70, except for the sleep subscale (r = .29) (N = 61 multiple sclerosis)	The ill individuals were significantly more fatigued than the controls	• Negative correlation between the FAI severity and consequences subscales, and a vitality measure (r = -.72; r = -.41) • Positive correlation between the FAI sleep subscale, and energy and humor measure (r = .43)	Confirmed all 4 domains
		Fatigue consequences subscale	α = .70				
		Responsiveness to rest/ sleep subscale	α = .85				
		TOTAL 29					

(*Continues*)

TABLE 2. Continued

Instrument	Subjects	Number of items/ Dimensions	Reliability: Internal Consistency	Reliability: Test-Retest	Validity: Content	Validity: Criterion	Validity: Discriminant	Validity: Divergent/ Convergent	Validity: Factorial Analysis	Validity: Responsivity
Brief Mental Fatigue Questionnaire (BMFQ) Bentall et al., 1993	94 Clinical: muscular dystrophy, CFS, treated CFS, depression 142 Nonclinical: controls	9 Severity	$\alpha = .85$ (nonclinical) $\alpha = .93$ (entire sample)	$r = .89$ ($p < .001$)			ANOVA: discriminated subjects with muscular dystrophy, treated CFS and controls, from subjects with CFS and depression			
Fatigue Scale (FS) Chalder et al., 1993	274 patients in general hospital 100 controls	7 Physical subscale 4 Mental subscale	$\alpha = .85$ $\alpha = .82$	Split half: $r = .85$ Split half: $r = .82$	Adequate		Established cutoff score of .85: discriminates the level of fatigue's morbidity		Confirmed both dimensions	

Fatigue Impact Scale (FIS) Fisk et al., 1994	145 Investigating CFS 34 Hypertension 105 Multiple sclerosis	10 Mental subscale 10 Physical subscale 20 Psychosocial subscale	$\alpha > .87$		The FIS discriminated the 3 sample groups	• Positive correlation between FIS and general health measure ($r > .53$)		
Multi-dimensional Fatigue Inventory (MFI-20) Smets et al., 1995	111 Radiotherapy 357 CFS 481 psychology students 158 medical students 46 junior physician 156 army recruits in physical training 160 recruits in barracks	4 General subscale 4 Physical subscale 4 Mental subscale 4 Reduced motivation 4 Reduced activity	$\alpha = .65–.93$ (except for 2 domains in the soldiers group—$\alpha = .57$ and $\alpha = .53$)	*Concurrent:* MFI-20 general, physical, reduced motivation, and reduced activity subscales, and Visual Analogic Scale ($r = .77$, $p < .001$)	The MFI discriminated all the sample groups The MFI discriminated subjects from same groups (psychology students) on general and physical dimensions The MFI discriminated subjects in different situations (soldiers) on general and reduced motivation dimensions		Confirmed the five dimensions	The MFI-20 was sensible only for the general dimension in the sample of junior physicians

(*Continues*)

TABLE 2. Continued

Instrument	Subjects	Number of items/ Dimensions	Reliability: Internal Consistency	Reliability: Test-Retest	Validity: Content	Validity: Criterion	Validity: Discriminant	Validity: Divergent/ Convergent	Validity: Factorial Analysis	Validity: Responsivity
		4 Severity subscale 7 Interference with functioning subscale						• Negative correlation between FSI and POMS energy subscale ($r > -.50$, $p < .001$),		
Fatigue Symptom Inventory (FSI) Hann et al., 1998	107 Treatment 70 Post-treatment 94 Control	2 Duration subscale	$\alpha > .93$	Reasonable stability for the group receiving treatment Inadequate stability for the control group	Adequate	*Concurrent:* FSI and POMS fatigue subscale ($r > .50$, $p < .001$)	Significant difference between treatment group, post-treatment group, and control group	• Positive correlation between FSI and depression and anxiety measure ($r > .30$, $p < .001$; $r > .30$, $p < .001$) • No correlation between FSI and social desirability measure ($r > -.07$)		

						• Positive correlation between DUFS and loss of vitality measure (ρ = .85)	
Dutch Fatigue Scale (DUFS) Tiesinga et al., 1998	138 Chronic heart failure	9 Severity	KR-20 = .79 ρ = .80		*Concurrent*: DUFS and numeric scale (ρ = .77)	• Negative correlation between DUFS and sleep quality (-.44, $p < .001$) • No correlation between DUFS and sleep quantity (ρ = .05)	Mokken test: coefficient *H* (load) = [.30, .64]
Dutch Exertion Fatigue Scale (DEFS) Tiesinga et al., 1998	138 Chronic heart failure	9 Severity	α > .91 ρ = .91	Adequate	*Concurrent*: DEFS and numeric scale of fatigue caused by exertion (ρ = .62)	• Positive correlation between DEFS and loss of vitality measure (ρ = .52)	Excluded items with load < .30 Mokken test: coefficient *H* = [.50, .67]

(*Continues*)

TABLE 2. Continued

Instrument	Subjects	Number of items/ Dimensions	Reliability: Internal Consistency	Reliability: Test-Retest	Validity: Content	Validity: Criterion	Validity: Discriminant	Validity: Divergent/ Convergent	Validity: Factorial Analysis	Validity: Responsivity
								• Negative correlation between DEFS and sleep quality (-.31, $p < .001$) • No correlation between DEFS and sleep quantity ($r = .03$)		
Revised Piper Fatigue Scale (RPFS) Piper et al., 1998	382 Breast cancer	6 Behavioral/ severity subscale	α = .92						Identified 5 constructs	
		6 Cognitive/ mood subscale	α = .93							
		5 Affective meaning subscale	α = .95							
		5 Sensory subscale	α = .96							

Schwartz Cancer Fatigue Scale (SCFS) Schwartz, 1998	166 Cancer 20 Controls	11 Physical subscale 7 Emotional subscale 5 Cognitive subscale 5 Temporal subscale	α = .93 α = .90 α = .85 α = .82	Adequate		Visual Analogic Scale score made possible the classification of fatigued subjects at 4 levels: absent, low, moderate, and high		Confirmed all 4 proposed dimensions
Brief Fatigue Inventory (BFI) Mendoza et al., 1999	305 Cancer 290 Controls	9 Severity	α = .96		*Concurrent*: BFI and POMS fatigue subscale (r = .84, p < .001)	Significant difference between groups	• Negative correlation between BFI and hemoglobin level (r = -.36, p < .001) • Negative correlation between BFI and FACT-F functionality measure (r = -.88, p < .001)	Indicated only one construct

(*Continues*)

TABLE 2. Continued

			Reliability		Validity					
Instrument	Subjects	Number of items/ Dimensions	Internal Consistency	Test-Retest	Content	Criterion	Discriminant	Divergent/ Convergent	Factorial Analysis	Responsivity
Cancer Fatigue Scale (CFS) Okuyama et al., 2000	307 Cancer	7 Physical subscale 4 Affective subscale 4 Cognitive subscale	α = .90 α = .78 α = .79	$r = .50$ ($p < .001$) women with breast cancer on treatment $r = .49$ ($p < .001$) patients in palliative care		*Concurrent*: Positive correlation between CFS and visual analogic scale ($r = .49$, $p < .001$)		• Positive correlation between CFS and depression measure ($r > .32$, $p < .001$) • Correlation between CFS and performance status measure ($r > .16$, $p < .001$) • No correlation between CFS and cognitive function measure ($.05 \leq r \leq .12$)	Identified 3 domains	

Schedule of Fatigue and Anergia for Chronic Fatigue Syndrome (SOFA/CFS) Hadzi-Pavlovic et al., 2000	368 chronic fatigue syndrome 430 controls	10 Severity		Established the cutoff score as .50 for the diagnosis of CFS	
Schedule of Fatigue and Anergia for General Practice (SOFA/GP) Hadzi-Pavlovic et al., 2000	1593 Patients in a general hospital	10 Severity	$r = .64$	Established cutoff score ≥ 3: discriminates the level of prolonged fatigue as a state of morbidity	• Positive correlation between SOFA/GP and health measure ($r = .47$) • Positive correlation between SOFA-GP physical symptoms measure ($r = .58$)

(*Continues*)

TABLE 2. Continued

Instrument	Subjects	Number of items/ Dimensions	Reliability: Internal Consistency	Reliability: Test-Retest	Validity: Content	Validity: Criterion	Validity: Discriminant	Validity: Divergent/ Convergent	Validity: Factorial Analysis	Validity: Responsivity
Cancer-related Fatigue Distress Scale (CRFDS) Holley, 2000	221 Cancer	20 Distress	α = .98		Adequate				Identified one construct	
		4 Cognitive subscale								
Fatigue Impact Scale for Daily administration (D-FIS) Fisk & Doble, 2002	57 Influenza	1 Psychosocial subscale	α = .91			*Concurrent*: D-FIS and numeric scale (r = .50)		• Negative correlation between D-FIS and work productivity (r = -.41)		ANOVA: F = 23.73, p < .001, t > 3.5, p < .001 (N = 26)
		3 Physical subscale						• Correlation between D-FIS and general health measure (r = -.22)		

					• Negative correlation between D-FIS and sleep quality (r = -.34) • Negative correlation between D-FIS and ability to do activities (r = -.55) • Positive correlation between D-FIS and respiratory symptoms (r = .29) and constitutional symptoms (r = .26)

instrument was not described in one of the articles; and in the others, the interval varied between 2 days and 12 months, which is a very wide range of time. The number of subjects who answered the instruments for the test-retest analysis varied substantially, decreasing considerably the second or third time.

Psychometric Properties: Validity

Validity was the psychometric property with the largest divergence in terminology and variety of tests. The tests applied to assess the instruments' validity were content validity, factor analysis, discriminant, criterion, and divergent or convergent validity. In the following paragraphs, a description of the tests is given.

Content validity consists of a judgment by specialists about the pertinence of items and domains in a scale in order to avoid under- or overrepresentation of a domain (Pasquali, 1997). This validity test was performed in five instruments (27.8%), and in all of them it was considered adequate.

Factor analysis was done in nine instruments (50%), and the Mokken test, which has similar purpose (Tiesinga, Dassen, & Halfens, 1998), in one instrument (5.6%). These tests are fundamental for identifying and confirming the instruments' theoretical domains (construct validity).

Discriminant validity was tested in 10 instruments (55.6%), and in all of them it was possible to discriminate fatigued from nonfatigued subjects (the difference in the scores was statistically significant). The discriminant validity should be established in instruments with discriminative purpose, and it has to have a cut-off point to operationalize the discrimination of the subjects. The authors of the Dutch Fatigue Scale did not present the criteria for discrimination of subjects or the cut-off point for the instrument.

Criterion validity was assessed by concurrent validity in nine articles (50%). This test was used to compare the constructed instrument with another instrument that measures the same phenomenon, usually one considered the "gold standard." Gold standards are considered rare in the field of measurement of subjective phenomena, and presently, no gold standard for fatigue assessment has been identified.

To establish concurrent validity three instruments (16.7%) used as the gold standard the VAS, and another three instruments (16.7%) used interval numeric scales, varying from 0 to 10. In all studies, the results were considered good or excellent.

Three instruments (16.7%) analyzed the concurrent validity by comparing the data of the constructed instrument with the fatigue subscale of the Profile of Mood States (McNair, Lorr, & Droppleman, 1971).

Divergent or convergent validity were tested in 11 instruments (61.1%), assessing correlations between fatigue and other health variables. The authors of the Fatigue Severity Scale observed weak correlations (varying from .20 to .46) between fatigue and depression (Krupp et al., 1989). This result was unanticipated, because it suggests that fatigue and depression are somewhat independent. The Visual Analogue Scale for Fatigue also showed weak correlations (varying from -.23 to .35) to a depression measure (Lee, Hicks, & Nino-Murcia, 1991). Other authors tested the relation between fatigue, depression, and anxiety, and observed positive correlation using

the Cancer Fatigue Scale and the Fatigue Symptom Inventory (Hann et al., 1998; Okuyama et al., 2000).

The Cancer Fatigue Scale presented no correlation with the Mini-Mental State Examination (Folstein, Folstein, & McHugh, 1975), that is, fatigue is not related to cognitive functions (Okuyama et al., 2000). The validation of the Visual Analogue Scale for Fatigue included the analysis of the correlation between energy and confusion (Table 2). This correlation was considered weak ($r = 0.41$, $p < .01$) (Lee et al., 1991), similar to that between the Cancer Fatigue Scale and the MMSE.

For some instruments the authors tried to determine the correlation between fatigue and quality and/or quantity of sleep. The Dutch Fatigue Scale and the Dutch Exertion Fatigue Scale showed negative correlations with quality of sleep ($r = -.44$; $r = -.31$) and no correlation ($r = .05$; $r = .03$) with the quantity (Tiesinga et al., 1998). Matching these results, the energy subscale of the Visual Analogue Scale for Fatigue also correlated positively (varying from .65 to .81) with quality of sleep (Lee et al., 1991).

Correlations between the several instruments (Fatigue Assessment Instrument, Fatigue Impact Scale, Fatigue Symptom Inventory, Dutch Fatigue Scale, Dutch Exertion Fatigue Scale, Brief Fatigue Inventory, Cancer Fatigue Scale, Schedule of Fatigue and Anergia, and Fatigue Impact Scale for Daily Administration), and functionality, vitality, and general health measures were negative and significant in all cases.

The Brief Fatigue Inventory was the only instrument that tried to test convergent validity using hemoglobin level as an objective measure. A negative correlation was observed, that is, the higher the fatigue, the lower was the hemoglobin level.

An important validity test that has to be performed in instruments with evaluative purpose is responsivity. This test, frequently mistaken with the test-retest, analyzes the sensitivity of the instrument to phenomena severity changes. Only four (28.6%) of the 14 instruments (77.8%) with evaluative purpose were tested for responsivity, namely, the Fatigue Severity Scale, Visual Analogue Scale for Fatigue, Fatigue Impact Scale for Daily Administration, and Multidimensional Fatigue Inventory. The first three presented good responsivity after therapeutic intervention, and the last, which tested responsivity comparing the data before and after a fatiguing situation, was only sensitive in one of the instrument's dimensions (general dimension) (Smets, Garssen, Bonke, & Haes, 1995).

DISCUSSION

Systematic review is considered a key strategy for the organization of information about a certain subject, being used for analysis of interventional or diagnostic studies. The development of a systematic review allows a constant update of data and the reproduction of the study, if anyone wishes to do so. Despite its importance, studies like this one are a great challenge, because the methodological procedures for developing systematic reviews of diagnostic studies are still limited compared to interventional studies (DeVillé et al., 2003).

Finding all published evidence related to fatigue assessment was difficult even though the databases and period were wide ranging. The analyses of diagnostic studies, especially those that assess subjective constructs, such as fatigue, are not easy. Some studies analyzed in this study are primary and incomplete in the description of the instruments' characteristics and construction method.

Clearly, researchers have made efforts to develop fatigue instruments to diagnose or evaluate symptom severity, although no gold standard is available. Since 1998, many instruments have been published, which suggests that only recently has the symptom begun to gain attention from researchers and health professionals. Many authors mention that the Fatigue Symptom Checklist, developed by Yoshitake (1971), was the first self-report instrument for fatigue assessment (Piper, 1997). However, it was not included in this review because it was not published in an indexed journal in the databases used in the present study.

The construction of the instruments was generally based on interviews with patients, literature review, and expert knowledge about the subject. The fact that an instrument is constructed based on patients' experiences is positive, because it sustains the principle of subjectivity of the phenomena and assesses the phenomena from the patients' own perspective. Literature review is also fundamental in terms of the theoretical foundation of the instrument, avoiding the development of similar instruments and improving the quality of new instruments when compared with the existing ones. The construction of instruments based only on expert knowledge tends to be of limited quality.

In the present study, the papers were carefully analyzed regarding the method to construct the instrument, definition of fatigue, target population, dimensions of fatigue, number and graduation of items, scores and cut-off scores, purpose of the instrument, reliability, and validity. These aspects are discussed below in two parts. The first one is about the general characteristics of the instruments and the second is about the psychometric properties.

General Characteristics

The definition of the phenomenon was presented in only six instruments (Table 1), despite being a key aspect for construction of an instrument, for the definition represents the operationalization of the concept. An author may or may not state an explicit definition, but one can always be inferred from the instrument. Because fatigue does not yet have a refined conceptual definition, the inclusion of items that are related to other concepts frequently occurs. It is necessary to clearly define the phenomenon and its characteristics to provide for better comprehension and to improve construct validity. Definitions for fatigue in all instruments included in this review were carefully analyzed. Detailed analysis of the instruments reinforces the subjective, multidimensional, and distressing nature of fatigue, as described in the introduction of this study. Multidimensionality itself is highlighted beyond physical and psychological feelings, including affective meaning and social dimensions. In the face of the concept's analysis, we can observe three more aspects that are important to fatigue assessment: duration (acute or chronic), intensity (weak,

moderate, or severe), and fatigue impact level on daily activities. We now propose the following definition of fatigue:

> an unpleasant physical, cognitive, and emotional symptom described as a tiredness not relieved by common strategies that restore energy. Fatigue varies in duration and intensity, and it reduces, to different degrees, the ability to perform the usual daily activities.

The target population was generic for most of the instruments ($n = 6$; Table 1), yet a considerable number of instruments were specific for fatigue assessment in patients with cancer ($n = 5$; Table 1). Patients with cancer have been the focus of many research centers because fatigue is known to be the most prevalent symptom, as well as an extremely debilitating one. Other population groups in which fatigue is very prevalent are those with multiple sclerosis and lupus. The Fatigue Severity Scale was developed to assess fatigue specifically in these groups, but it is possible to use it in other population groups because its items are generic.

Among the analyzed instruments, 11 were considered short (< 15 items). Many authors mention the importance of a brief assessment of fatigue so that the instrument itself will not prove fatiguing, thus "interfering" with the results of fatigue assessment. Instruments short in length may be adequate for clinical use, but it must be remembered that instruments with fewer items may be less sensitive in identifying and quantifying.

Half of the instruments were multidimensional, varying from two to five dimensions (Table 1). The dimensions of fatigue assessed by the instrument are of particular interest considering the multidimensionality of the symptom. If an instrument assesses only physical or mental aspects of fatigue, it is assessing just a part of the phenomena, which increases possible mistakes in judgments about the severity or the domains of fatigue and treatment decisions.

Another aspect analyzed was the vocabulary used in the instruments' items and their scaling. Complex words or difficult phrases can be more fatiguing than many items in simple writing. Twelve instruments used Likert-type scale (varying from 4 to 7 points), four used numeric scales (varying from 0 to 10), and, of course, one used VAS. Perspectives vary regarding the use of Likert scales, numeric scales, or a VAS, but each has its advantages. Brunier and Graydon (1996) compared the VAS and the Likert scale, and the results showed significant relation between the two measures. Despite the similarity between both scales, the assessment by VAS was suggested because this scale was capable of better capturing the level of patients' fatigue because it has a greater number of categories (Brunier & Graydon, 1996). Aaronson and associates (1999) argued that the use of a Likert scale is justified by the fact that it demands less mental abstraction and manual ability, requires less time for answering, and is easier to analyze. To ensure the scale's reliability, it is important that each item have an adequate number of categories.

Psychometric Properties

The terminology of the psychometric properties varied greatly, and mistakes in the concepts and the use of the words are frequent, for example, divergent validity and

convergent validity as synonymous, or responsivity and test-retest as the same test. These facts result in some confusion and make the comprehension of the information difficult.

The number of subjects included in the testing influences the results of validity and reliability tests. Kerlinger (1986) recommended the inclusion of 10 subjects per item of the instrument, and this criterion was followed by 12 (66.7%) of the instruments analyzed in the present study. On the others (n = 6), at least 6 subjects per item were included, as shown in Table 2.

The reliability of the all instruments ranged from good to excellent due to the high values obtained by the measurement of internal consistency, either by calculating Cronbach's alpha or Kuder-Richardson, as shown in Table 2. The high values show that the items are closely related one to the other, measuring the same phenomena.

The test-retest was done in six studies, and the correlation was moderate (varying from .49 to .89). The time interval between the instrument's application, and the number of subjects showed great variability and might have negatively influenced the results of this test because feelings of fatigue are influenced by circadian rhythm (Piper et al., 1998) and by diseases and their stages (Mota & Pimenta, 2002). The Fatigue Severity Scale performed the test-retest with 11 subjects and obtained adequate values for a time interval of from 5 to 33 weeks. The Brief Mental Fatigue Questionnaire also obtained adequate values for the time interval of 2 days, but bear in mind that the fatigued individual might have remembered the answers given previously.

Although many validity tests are available, not all studies applied adequate testing. The establishment of concurrent validity using better-suited and more sophisticated instruments with simpler scales, such as the VAS or numeric scales, is debatable. Now that there are many instruments for fatigue assessment, the VAS and numeric scales may be substituted. Also debatable is the use of the Profile of Mood States for calculating concurrent validity, because its fatigue subscale assesses fatigue related to humor, and this can limit its use for validity tests when analyzing fatigue on physical or impact dimensions.

Regarding divergent and convergent validity, the results were interesting, but since there is insufficient knowledge about the correlation between fatigue and these variables, the validity results are difficult to analyze. For example, depression showed no positive correlation to a fatigue measure. Cognitive function also did not correlate to fatigue, which was unanticipated and raises questions regarding the conditions of the study and the adequacy of mental fatigue and cognitive impairment measurement. Variables such as functionality, vitality, and general health measures showed no confusion regarding relation to fatigue. These results indicate the close connection between these variables and confirm the present concept that fatigue results in a decrease of functional capacity. The use of objective measures to assess convergent validity happened once. This type of investigation is valuable, but it can only be used as a validity measure if there is previous information about the relation between the variables in a specific group of patients.

Tests of responsivity are important when assessing instruments with evaluative purpose. It was analyzed by four of 14 evaluative instruments, and the results were not adequate for all of these (Table 2). The responsivity results might have been influenced by the number of subjects. All four instruments that were tested for responsivity had a small sample (Fisk & Doble, 2002; Krupp et al., 1989; Lee et al., 1991; Smets et al., 1995) and, sometimes, such as in the Muldimensional Fatigue Inventory, only one subgroup of the sample participated, making generalization of the results impossible (Smets et al., 1995). Authors of the other instruments might not have tested responsivity because there are no effective interventions recognized for treatment or control of fatigue, especially in patients with chronic illness.

This study made possible the elaboration of a list of important points when choosing instruments for the assessment of fatigue: clear definition of the phenomena; moderate length and simple language; assessment of multiple dimensions; the purpose of measurement meets clinical or research targets (evaluative, discriminative, or predictive); reliability (at least internal consistency); validity (factor analysis, concurrent validity, and divergent or convergent validity), and if the instrument was constructed in a different country or expressly for a certain population, validation and adaptation of the instrument to the culture or population. The majority of points stated here agree with the issues mentioned elsewhere (Dittner et al., 2004). However, in order to better assist fatigued individuals, it is more important that the assessment of fatigue by instruments give an outcome related exclusively to this symptom because, if a generic instrument is used, its score will not provide a clear estimate of the fatigue experienced, and results of intervention may not be identified.

Three instruments analyzed in this study should be pointed out, because of good construction and validation methods: Fatigue Assessment Instrument, Schwartz Cancer Fatigue Scale, and Cancer Fatigue Scale. However, knowing that validation of instruments is a process, other instruments may be shown to be as good measures of fatigue as these. By recognizing the quality of many instruments analyzed in this paper, and knowing the importance of assessing fatigue in order to better assist patients, clinicians and researchers should validate and adapt a self-report instrument for their own culture.

This study analyzed papers published in English, Portuguese, and Spanish only, and a broader review of the literature is recommended. All instruments were developed in the English language, and validation of a fatigue instrument in other cultures should be a priority to improve assistance to fatigued individuals, allowing comparison of research outcomes among different cultures.

CONCLUSION

Characterizing the severity and the components of symptoms such as fatigue and pain is difficult, and the diagnostic instruments available differ in many ways. Analysis of validity and reliability of tests and establishing the best tests are fundamental to

diagnose fatigue and are as important as knowing the best interventions to deal with the symptom. This study contributes to knowledge development about fatigue, helping clinicians and researchers who assist fatigued patients, because it has identified and analyzed the existing instruments for fatigue assessment and pointed out the more adequate ones.

This review aimed at the identification and analysis of self-report instruments for fatigue assessment of adults. Eighteen instruments were identified. Several instruments did not explicitly define the phenomena, that is, the theoretical basis for the construction of the instrument; did not provide important information regarding the tests used for validation; did not include the scoring method; and did not state the purpose of the instrument. In addition, it was frequently observed that not all validity and reliability tests indicated for the instrument with a determined purpose were conducted. These findings show how important are systematic review for diagnostic studies.

Considering the relevance of fatigue as a limiting symptom and its prevalence, it is necessary to improve the quality of fatigue assessment, which will lead to a better understanding and management of the phenomena.

REFERENCES

Aaronson, L. S., Teel, C. S., Cassmeyer, V., Neuberger, G. B., Pallikkathail, L., Pierce, J., et al. (1999). Defining and measuring fatigue. *Image: Journal of Nursing Scholarship, 31*, 45–50.

Atallah, A. N., & Castro, A. A. (2003). *Revisão sistemática da literatura e metanálise: a melhor forma de evidência para tomada de decisão em saúde e a maneira mais rápida de atualização terapêutica* [Systematic review of literature and meta-analysis: The best evidence for decision-making in health and a fast way for therapeutic update]. Retrieved October 4, 2003, from http://www.epm.br/cochrane

Bentall, R. P., Wood, G. C., Marrinan, T., Deans, C., & Edwards, R. H. T. (1993). A brief mental fatigue questionnaire. *British Journal of Clinical Psychology, 32*, 375–379.

Brunier, G., & Graydon, J. A comparison of two methods of measuring fatigue in patients on chronic haemodialysis: Visual analogue vs. Likert scale. *International Journal of Nursing Studies, 33*, 338–348.

Chalder, T., Berelowitz, G., Pawlikowska, T., Watts, L., Wessely, S., Wright, D., et al. (1993). Development of a fatigue scale. *Journal of Psychosomatic Research, 37*, 147–153.

Conill, C., Verger, E., Henríquez, I., Saiz, N., Espier, M., Ferran, L., et al. (1997). Symptom prevalence in the last week of life. *Journal of Pain and Symptom Management, 14*, 328–331.

Detmar, S. B., Muller, M. J., Wever, L. D. V., Schornagel, J. H., & Aaronson, N. K. (2001). Patient-physician communication during outpatient palliative treatment visits: An observational study. *Journal of the American Medical Association, 285*, 1351–1357.

Devillé, W. L., Buntinx, F., Bouter, L. M., Montori, V. M., de Vet, H. C. W., van der Windt, D. A. W. M., et al. (2003) *Conducting systematic reviews of diagnostic studies: Didactic guidelines*. Retrieved October 4, 2003, from http://www.biomedcentral.com/1471-2288/2/9

Dittner, A. J., Wessely, S. C, & Brown, R. G. (2004). The assessment of fatigue: A practical guide for clinicians and researchers. *Journal of Psychosomatic Research, 56*, 157–170.

Esper, P., & Redman, B. G. (1999). Supportive care, pain management, and quality of life in advanced prostate cancer. *Urologic Clinics of North America, 26*, 375–389.

Fisk, J. D., & Doble, S. E. (2002). Construction and validation of a fatigue impact scale for daily administration (D-FIS). *Quality of Life Research, 11*, 263–272.

Fisk, J. D., Ritvo, P. G., Ross, L., Haase, D. A., Marrie, T. J., & Schlech, W. F. (1994). Measuring the functional impact of fatigue: Initial validation of the fatigue impact scale. *Clinical Infectious Diseases, 18*, S79–S83.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental State": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*, 189–198.

Hadzi-Pavlovic, D., Hickie, I. B., Wilson, A. J., Davenport, T. A., Lloyd, A. R., & Wakefield, D. (2000). Screening for prolonged fatigue syndromes: Validation of the SOFA scale. *Social Psychiatry and Psychiatric Epidemiology, 35*, 471–479.

Hann, D. M., Jacobsen, P. B., Azzarello, L. M., Martin, S. C., Curran, S. L., Fields, K. K., et al. (1998). Measurement of fatigue in cancer patients: Development and validation of the Fatigue Symptom Inventory. *Quality of Life Research, 7*, 301–310.

Holley, S. K. (2000). Evaluating patient distress from cancer-related fatigue: An instrument development study. *Oncology Nursing Forum, 27*, 1425–1431.

Jenkins, C. A., Schulz, M., Hanson, J., & Bruera, E. (2000). Demographic, symptom, and medication profiles of cancer patients seen by a palliative care consult team in a tertiary referral hospital. *Journal of Pain and Symptom Management, 19*, 174–1784.

Kaasa, S., Loge, J. H., Knobel, H., Jordhoy, M. S., & Brenne, E. (1999) Fatigue: Measures and relation to pain. *Acta Anaesthesiologica Scandinavica, 43*, 939–947.

Kerlinger, F. N. (1986). *Foundations of behavioral research* (3rd ed.). New York: Holt, Rinehart & Winston.

Kjerulff, K. H., & Langenberg, P. W. (1995). A comparison of alternative ways of measuring fatigue among patients having hysterectomy. *Medical Care, 33*, S156–S163.

Krupp, L. B., La Rocca, N. G., Nuir-Nash, J., & Steinberg, A. D. (1989). The fatigue severity scale: Application to patients with multiple sclerosis and systemic lupus erythematosus. *Archives of Neurology, 46*, 1121–1123.

Lee, K. A., Hicks, G., & Nino-Murcia, G. (1991). Validity and reliability of a scale to assess fatigue. *Psychiatry Research, 36*, 291–298.

Liao, S., & Ferrell, B. A. (2000). Fatigue in an older population. *Journal of the American Geriatrics Society, 48*, 426–430.

Lichstein, K. L., Melanie, K. M., Noe, S. L., & Aguillard, R. N. (1997). Fatigue and sleep disorders. *Behavioral Research Therapy, 35*, 733–740.

McDowell, I., & Newell, C. (1996). *Measuring health: A guide to rating scales and questionnaires* (2nd ed.). New York: Oxford University Press.

McNair, D. M., Lorr, M., & Droppleman, L. F. (1971). *Manual for the Profile of Mood States*. San Diego: Educational and Industrial Testing Services.

Meek, P. M., Nail, L. M., Barsevick, A., Schwartz, A. L., Stephen, S., Whitmer, K., et al. (2000). Psychometric testing of fatigue instruments for use with cancer patients. *Nursing Research, 49*, 181–190.

Mendoza, T. R., Wang, X. S., Cleeland, C. S., Morrissey, M., Johnson, B. A., Wendt, J. K., et al. (1999). The rapid assessment of fatigue severity in cancer patients: Use of the Brief Fatigue Inventory. *Cancer, 85*, 1187–1196.

Mota, D. D. C. F., & Pimenta, C. A. M. (2002). Fadiga em pacientes com câncer avançado: conceito, avaliação e intervenção [Fatigue in patients with advanced cancer: Concept, assessment, and management]. *Revista Brasileira de Cancerologia, 48*, 577–583.

Okuyama, T., Akechi, T., Kuhaya, A., Okamura, H., Shima, Y., Maruguchi, M., et al. (2000). Development and validation of the Cancer Fatigue Scale: A brief, three-dimensional, self-rating scale for assessment of fatigue in cancer patients. *Journal of Pain and Symptom Management, 19*, 5–14.

Packer, T. L., Sauriol, A., & Brouwer, B. (1994). Fatigue secondary to chronic illness: Postpolio syndrome, chronic fatigue syndrome, and multiple sclerosis. *Archives of Physical Medicine and Rehabilitation, 75*, 1122–1126.

Pasquali, L. (1997). *Psicometria: teoria e aplicações* [Psychometry: Theory and applications]. Brasilia: Editora Universidade de Brasília.

Passik, S. D., Kirsh, K. L., Rosenfeld, B., McDonald, M. V., & Theobald, D. E. (2001). The changeable nature of patients' fears regarding chemotherapy: Implications for palliative care. *Journal of Pain and Symptom Management, 21*, 113–120.

Piper, B. F., Dibble, S. L., Dodd, M. J., Weiss, M. C., Slaughter, R. E., & Paul, S. M. (1998). The revised Piper Fatigue Scale: Psychometric evaluation in women with breast cancer. *Oncology Nursing Forum, 25*, 677–684.

Piper, B. F. (1997). Measuring fatigue. In M. Frank-Stromborg & S. J. Olsen (Eds.), *Instruments for clinical health-care research* (pp. 482–496). Boston: Jones & Bartlett.

Piper, B. F., Lindsey, A. M., Dodd, M. J., Ferketich, S., Paul, S. M., & Weller, S. (1989). In S. G. Funk, E. M. Tornquist, M. T. Champagne, L. A. Copp, & R. A. Wiese (Eds.), *Key aspects of comfort: Management of pain, fatigue, and nausea* (pp. 199–208). New York: Springer.

Schwartz, A. (1998). The Schwartz Cancer Fatigue Scale: Testing reliability and validity. *Oncology Nursing Forum, 25*, 711–717.

Schwartz, C. E., Coulthard-Morris, L., & Zeng, Q. (1996). Psychosocial correlates of fatigue in multiple sclerosis. *Archives of Physical Medicine and Rehabilitation, 77*, 165–170.

Schwartz, J. E., Jandorf, L., & Krupp, L. B. (1993). The measurement of fatigue: A new instrument. *Journal of Psychosomatic Research, 37*, 753–762.

Smets, E. M. A., Garssen, B., Bonke, B., & Haes, J. C. J. M. (1995). The Multidimensional Fatigue Inventory (MFI) psychometric qualities of an instrument to assess fatigue. *Journal of Psychosomatic Research, 39*, 315–325.

Stone, P., Hardy, J., Broadley, K., Tookman, A. J., Kurowska, A., & A'Hern, R. (1999). Fatigue in advanced cancer: A prospective controlled cross-sectional study. *British Journal of Cancer, 79*, 1479–1486.

Streiner, D. L., & Norman, G. R. (1995). *Heath measurement scales: A practical guide to their development and use* (2nd ed.). New York: Oxford Medical Publications.

Tiesinga, L., Dassen, T. W. N., & Halfens, R. J. G. (1998). DUFS and DEFS: Development, reliability and validity of the Dutch Fatigue Scale and the Dutch Exertion Fatigue Scale. *International Journal of Nursing Studies, 35*, 115–1123.

Trendall, J. (2000). Concept analysis: Chronic fatigue. *Journal of Advanced Nursing, 32*, 1126–1131.

Yoshitake, H. (1971). Relations between the symptoms and the feeling of fatigue. *Ergonomics, 14*, 175–186.

Acknowledgments. This work was supported by a grant from CAPES (Coordenação de Pessoal de Nível Superior—Brazilian Ministry of Education and Culture) and from the National Counsel of Technological and Scientific Development (CNPq), Brazil.

Offprints. Requests for offprints should be directed to Cibele A. M. Pimenta, PhD, RN, School of Nursing, University of São Paulo, Av. Dr. Enéas de Carvalho Aguiar 419, Cerqueira César, São Paulo SP, 05403-000, Brazil. E-mail: parpca@usp.br