Published in final edited form as:

J Clin Epidemiol. 2009 September ; 62(9): 991–997. doi:10.1016/j.jclinepi.2008.10.007.

Classical testing theory and Item Response theory/Rasch model to assess difference between Patient-Reported Fatigue Using Seven-Day and Four-Week Recall Periods

Jin-Shei Lai, PhD, OTR/L^{1,2}, Karon Cook, PhD³, Arthur Stone, PhD⁴, Jennifer Beaumont, MS¹, and David Cella, PhD^{1,2}

¹Center on Outcomes, Research and Education (CORE), Evanston Northwestern Healthcare, Evanston, IL

²Institute for Healthcare Studies, Northwestern University School of Medicine, Chicago, IL

³University of Washington, Seattle, WA

⁴Stony Brook University, Stony Brook, NY

Keywords

Cancer-related Fatigue; Recall; Functional Assessment of Chronic Illness Therapy-Fatigue

What Is New?

1. This is one of the first studies to examine the timeframe issue by using both classical testing theory and Item Response theory/Rasch model
2. This manuscript provides empirical evidence comparing recall periods which has been a long-debated issue.
3. We conclude that substantive considerations regarding the appropriate time frame should outweigh statistical ones.

INTRODUCTION

Cancer-related fatigue (CRF) has been defined as overwhelming and sustained exhaustion that decreases capacity for physical and mental work,[1] and it is the most common unrelieved symptom experienced by cancer patients and survivors.[2–5] Numerous measurement tools have been developed to measure patient-reported fatigue: from a 0–10 screening item to a multi-dimensional fatigue inventory, and from traditional fixed length scales to full item banks that serves as the foundation for computerized adaptive testing. [6–11] However, there is still no consensus on which time frame better captures fatigue experienced by patients. For example,

© 2009 Elsevier Inc. All rights reserved.

Correspondences and reprint requests should be directed to the first author, Jin-shei Lai, Ph.D., OTR/L, Center on Outcomes Research and Education (CORE), Evanston Northwestern Healthcare and Northwestern University, 1001 University Place, Suite 100, Evanston, IL 60201. TEL: 224-364-7312, FAX: 847-570-8033, js-lai@northwestern.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

the Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-F[10]) taps fatigue experienced during the past seven days; the Brief Fatigue Inventory[7] captures it within 24 hours; the Fatigue Symptom Inventory[6] measures fatigue severity in the past seven days as well as “current” fatigue; and the SF-36 Health Survey[12] and the PedsQL™ Multidimensional Fatigue Scale[13] have both 7-days and 4-weeks versions available.

Various factors (e.g., gender, memory and personality) may impact how patients report their symptom severity.[14;15] For example, when persons are asked to report their fatigue based on longer periods of time memory processes and personality disposition are likely to influence their responses. Although little is actually known about the “extraneous” factors that influence cancer-related fatigue, we may extrapolate or infer based on what is known about other symptoms such as pain or mood. Patient evaluations of pain and mood are known to vary systematically with the length of the recall period.[15] Robinson and Clore[16] argue that different types of memory are accessed when people are recalling relatively recent affect (the last few days) versus longer-term recall (the last week or month). Episodic memory is operative for short-term recall whereas semantic memory comes into play with longer-term recall. Since semantic memory is closely aligned with one’s beliefs as opposed to actual experience, longer-term recall may be biased. Other factors such as gender also influence the accuracy of recall. Additionally, investigators noted that patients tend to endorse fatigue ratings based on selective memory of the worst fatigue experienced during a period and to downplay less severe or fatigue-absent periods (peak-end effect and duration neglect, respectively).[17]

This study aimed to compare fatigue reported based on 7-days and 4-weeks time frames and explore factors that might affect patients’ responses. We compared the impact of a 7-days versus a 4-weeks time frame as both are commonly used in fatigue assessments. Results of this study therefore offer insight into the practical importance of differences in ratings by time frame.

METHODS

Procedures

Two touch-screen computers were dedicated to this study. One was loaded with the Functional Assessment of Chronic Illness Therapy-Fatigue, FACIT-F,[10] using a 7-days time frame. The other was loaded with exactly the same questions, but with a 4-weeks time frame. These two computers were assigned to two research assistants (RAs) to recruit patients from five Chicago metropolitan clinics. Since both RAs recruited patients from all five clinics, it was expected that each clinic had similar percentage of patients completing 7-days and 4-weeks versions of FACIT-F and therefore, the selection bias among patients was minimized.

Measures

The FACIT-F is a 13-item questionnaire that assesses self-reported tiredness, weakness, and difficulty conducting usual activities due to fatigue.[10] A 5-point intensity type of rating scale (from “not at all” to “very much”) is used. The FACIT-F is proved to be psychometrical sound instrument and has been widely used to measure fatigue for patients with various chronic illness [18] as well as for the US general population.[19]

As part of the assessment, patients rated their current fatigue at the beginning and at the end of the assessment using a 0–10 rating (0=not tired at all; 10=as tired as you can imagine), i.e., pre-assessment of fatigue and post-assessment of fatigue, respectively. Using the National Comprehensive Cancer Network (NCCN) fatigue guideline,[20] these fatigue scores were converted into mild (ratings of 0–3), moderate (4–6), and severe (7–10). This classification has been commonly used in oncology clinics. Additionally, patients were also asked to answer the

questions “what time frame did you have in mind when answering the previous questions” (i.e., time frame in mind) and “what time frame do you think would be the best to use for these questions” (i.e., preferred time frame) at the end of the assessment.

Analysis Plan

• **Comparison of Samples who Endorsed Different Time frames**—We first compared the samples completing the two time frame versions on fatigue severity (both pre- and post-assessment of fatigue as mentioned before) using the non-parametric Mann-Whitney U statistic. This analysis was conducted to rule out the possibility that difference between two timeframe versions was a result of different fatigue severity between samples. Additionally, Cochran-Mantel-Haenszel statistics and Cochran-Armitage trend test were used to test whether Eastern Cooperative Oncology Group Performance Status Rating, ECOG PSR, reported by patients was associated with time frame (i.e., 7-days and 4-weeks). With a sample size of at least 100 patients per group, this study was sufficiently powered (power = 0.80, 2-sided alpha = 0.05) to detect, as statistically significant, differences corresponding to an effect size (mean difference / standard deviation) of 0.40.

• **Comparisons of Items between 7-days and 4-weeks versions**—We first used Cochran-Mantel-Haenszel statistics and Cochran-Armitage trend test to assess the association between time frame and item scores. We then depicted item information function curves and scale information curves[21–23] along the fatigue continuum. Unlike classical test theory that provides a single reliability estimate(e.g., Cronbach’s alpha) for all points in the scale, item response theory models (including Rasch analysis) estimate information function at both scale and item levels to allow for examining the precision levels along the measurement continuum. The information function indicates the degree of accuracy with which a patient’s fatigue level is estimated at different points along the fatigue continuum. The higher information function indicates more accuracy (i.e., less measurement error). The item information functions for all 13 items were summed to produce a scale information function. This technique has been used to describe characteristics of items and scales in measuring health-related quality of life and symptoms.[21;22] In this manuscript, we evaluated item information function curves between time frames to compare their precision levels across the fatigue continuum. The Winsteps computer program[24] was used to produce these curves.

Finally, we examined the stability of the psychometric properties at the item-level between time frames by using differential item functioning, DIF.[25] DIF was examined by using two techniques: hierarchical sequential modeling processes in ordinal logistic regression[26] (DIF criterion: significant chi-square statistic, $p < 0.01$) and comparisons of item calibrations produced by Rasch analysis on each sub-group (DIF criterion: $t > 2.58$, $p < .01$). Items that exhibited DIF with respect to time frame in both DIF detection methods approaches were considered unstable. Items that did not exhibit DIF or exhibited DIF based on only one method were considered sufficiently stable across time frames.

• **Examination of Potential Influential Factors**—We tabulated patients’ preferred time frame and the time frame they used in completing the FACIT-F based on their responses to items “time frame in mind” (*what time frame did you have in mind when answering the previous questions*) and “preferred time frame” (*what time frame do you think would be the best to use for these questions*). We then examined whether gender and fatigue severity impacted patients’ “time frame in mind” and “preferred time frame” with chi-square statistics. Further, we examined whether patients’ preferred time frame was associated with the time frame they reported using to complete the FACIT-F.

RESULTS

Sample

Two hundred and sixteen patients were recruited, 116 completed the 4-weeks version and 100 completed the 7-days version of the questionnaires. The Institutional Review Board of each study site approved the study before patients were approached, and all participants provided written informed consent. Sample demographic and clinical information, grouped by the time frame they were assigned, is shown in Table 1. In brief, the majority of patients in both groups (4-weeks; 7-days) was female (64%; 63%), white (73%; 88%), received chemotherapy (62%; 86%) and/or surgery (50%; 54%), and months since diagnosis were 76.4 (SD=88.3) and 91.1 (SD=112.9), respectively. Of the patients who completed the 4-weeks version of the FACIT-F (mean age= 54.7 years), the most frequent diagnosis was breast cancer (40.4%), followed by colorectal cancer (15.8%), non-Hodgkin's disease (12.3) and ovarian cancer (4.4%). The sample that completed the 7-days version of the FACIT-F were older on average (mean age=60.4 years) and included more patients with primary cancer types of ovarian (8.2%), lung (9.2%), Hodgkin's (4.1%), bladder (5.1%) and prostate (6.1%). The comparison results of these variables were also reported on Table 1.

Using the NCCN guideline as mentioned earlier, patients who completed the 4-weeks version of the FACIT-F, 65.5% rated themselves as having mild fatigue, 21.6% moderate, and 12.9% severe fatigue in the pre-assessment of fatigue, and similar groupings for the post-assessment of fatigue. For those who responded to the 7-days version of the FACIT-F, 56% reported mild fatigue, 28% moderate and 16% severe fatigue in the pre-assessment of fatigue. Slightly more patients reported experiencing greater post-assessment of fatigue, in which, 50%, 33% and 17% reported mild, moderate and severe fatigue, respectively.

Sample Comparison

Mann-Whitney tests showed no significantly different fatigue severity (reported by using 0–10 ratings) between 7-days and 4-weeks versions, $p = .09$ and $p = 0.06$ for pre- and post-assessment of fatigue, respectively. FACIT-F scores were calculated using a standardized procedure that is consistent with all other scales of the FACIT measurement system. Averaged FACIT-F score was 36.6 (SD=10.8; range: 8–52) and 35.9 (SD=12.2; range: 3–52) for 4-weeks and 7-days versions, respectively. Mann-Whitney U statistic results showed no statistically significant differences between scores from these two versions, $z = -0.13$, $p = 0.45$. Results from both Cochran-Mantel-Haenszel statistics ($p = 0.17$) and Cochran-Armitage trend test ($p = 0.08$) indicated no significant association between time frame and ECOG performance rating. It was therefore concluded that patient samples were comparable in fatigue severity and functional status. We further concluded that data from these samples could be used for the consequent analyses regardless the significant difference in age, race (i.e., White), months since diagnosis, and receiving chemotherapy were found.

Item Comparisons

Item descriptive statistics are shown on Table 2. Both Cochran-Mantel-Haenszel statistics and Cochran-Armitage trend test failed to reject the null hypothesis on all comparisons, indicating that the time frame did not influence patients' endorsements of any of the items.

Comparisons of item information function curves between time frames are shown in Figure 1. Some inconsistent results were noted. For example, "I need to sleep during the day" was more precise (higher information function) for patients who reported more fatigue by using a 7-days time frame. The same question was more precise for patients with less fatigue when the time frame was 4 weeks. However, the opposite was observed for the item: "I have trouble starting things because I am tired." A seven-day time frame was superior to a 4-weeks when measuring

“I am too tired to eat,” as its information function was consistently higher than when the 4-weeks time frame was used. The 4-weeks was superior for item “I need help doing my usual activities.” Despite these small inconsistencies at the item level, little difference was found in the summary scale information function (Figure 2) although 7-days time frame was slightly more precise overall than the 4-weeks time frame.

No item demonstrated DIF in ordinal logistic regression; while “I need to sleep during the day” exhibited significant DIF ($p=0.0047$) when comparing item calibrations between time frames. Therefore this item did not meet both criteria for DIF.

Influential Factors

Most patients reported using the time frame they were asked to use in completing questions (i.e., response to the item “time frame in mind”), 64.6% of patients who received the 7-days version and 44.8% who received the 4-weeks version. However, many reported that they endorsed questions using “today” as the reference regardless which time frame they were assigned, 17.2% and 25.9% for 7-days and 4-weeks questionnaires, respectively (Table 3). Surprisingly, when they were asked the best time frame to measure fatigue (i.e., response to the item “preferred time frame”), far fewer patients chose “today”: 6.2% and 11.2% for patients who completed 7-days and 4-weeks versions, respectively. Patients who completed the 4-weeks version tended to consider four weeks the best time frame, Kappa=0.61 (95% confident interval: 0.50–0.72); while less agreement was found among patients who completed the 7-days version, Kappa=0.29 (95% confident interval: 0.17–0.41).

Results of chi-square statistics showed that both gender and fatigue severity were not significantly associated with the time frame patients reported using to endorse items, $p = 0.48$ and $p = 0.33$ (pre-assessment of fatigue; $p = 1.00$ for post-assessment of fatigue), respectively. There was no significant difference between pre- and post-assessment of fatigue for patients using either time frame, Student's $t = -0.61$ ($p = 0.54$) and Student's $t = 0.63$ ($p = 0.53$) for patients who completed 4-weeks and 7-days questionnaire, respectively.

We then examined whether patients who endorsed “today” to the item “time frame in mind” ($N=47$) experienced more fatigue than patients who endorsed other options to this item. No significant differences in fatigue scores (0–10 single item) were found between groups, $p = 0.23$ and $p = 0.34$ pre-assessment of fatigue and post-assessment of fatigue, respectively.

DISCUSSION

At the item and scale levels, self-reported fatigue was not significantly different between 7-days and 4-weeks time frames. Neither patient gender nor severity of fatigue had an impact on this result. When we examined the amount of information each item provided across the fatigue continuum, inconsistent results for some items were identified. While some items were equally precise across the fatigue continuum, others (2 of 13) showed considerable differences in precision. The 7-days time frame generally provides more precise estimation (i.e., higher information function value) on “I am too tired to eat” while the 4-weeks time frame had greater information for the item “I need to help doing my usual activities” regardless the fatigue severity. A few items demonstrated smaller differences in precision depending on the locations on the fatigue continuum. However, despite some variations at the item level, when we examined information function at the scale level, the 7-days time frame showed a slightly higher information function than the 4-weeks time frame. Accordingly, we conclude that scores from 7-days and 4-weeks time frames are comparable when the summation scores of the FACIT-F are used.

This study has some limitations. First, we randomly assigned computers, not individuals, and yet the level of our analysis is on the individual data. Second, our choice of recall periods (“7 days” versus “4 weeks”) does not enable us to evaluate shorter recall periods (e.g., 24 hours) which may produce even more informative responses. With the exception of asking about needing help with one’s usual activities, results support indifference to either of these two recall periods or a preference for the 7-days over the 4-weeks time frame. Future study could compare current (“right now”) fatigue and fatigue in a shorter time window (e.g., 24 hours). Nevertheless, from an applied perspective, it is useful to know that there are not substantial differences associated with reporting periods in the 1–4 weeks reporting window. Though the accuracy of fatigue recall has not been assessed, accuracy of recall of pain at selected moments has been studied and inconsistent results have been reported. Some studies found that the recall of pain is fairly accurate and others found that post-hoc ratings differ from ratings at the time of experience.[27;28] A few studies showed that inaccuracies are related to pain and emotion at the time of the elicitation.[29;30] Stone and his colleagues[15] did not find evidence to support that recall of pain was influenced by momentary reporting in the week prior to weekly recall while the associations between immediate pain at the time of recall assessment and levels of recall assessment.

Though this study was not designed to test the accuracy of recall, we compared fatigue severity for patients who endorsed “today” to the item “*what time frame did you have in mind when answering the previous questions*” and those who endorsed other time frame options (i.e., “past few days”, “past one week”, “past two weeks”, “past four weeks”, and “more than four weeks ago”). Regardless of the time frame version of the FACIT-F assigned, 22% (Table 3; 17% from 7-days version and 26% from 4-weeks version) of patients endorsed “today” for this item. We found no statistically significant difference in fatigue severity between patients who endorsed “today” and those who reported using other recall periods. Surprisingly, few patients considered “today” the best time frame for these questions (6% and 11% for 7-days and 4-weeks, respectively). In fact the majority (i.e., greater than 50%) of patients in both conditions (7-days and 4-weeks) preferred recall period equal to or longer than 1 week (see Table 3). In general, these longer periods tend to increase recall bias. Selecting an optimal recall period can therefore be an exercise in compromising between a length of time preferred by patients and one relatively less subject to recall bias. These results therefore suggest a choice of 7-days recall, leaving open the question of whether a recall period shorter than 7 days would differ in precision or patient preference.

The results have implications for clinical and survey studies. There is no statistically significant difference between scores obtained from the 7-days and 4-weeks versions of the FACIT-F. Patients’ gender and fatigue severity do not influence this conclusion. There is no statistically significant difference between reported fatigue scores (0–10 rating) obtained at the beginning of the study and those obtained at the end of the study. When taking fatigue severity into account, the 7-days version demonstrated a slightly higher information function (“precision”) compared to the 4-weeks version. In studies in which 7-days and 4-weeks self-reports of fatigue are equally appropriate, preference might be given to asking questions using the more informative 7-days time frame. However, substantive considerations regarding the appropriate time frame should outweigh statistical ones.

Acknowledgments

Funding Source: This study was supported by grants from the National Cancer Institute (#CA60068, PI: David Cella), and National Institutes of Health (U01 AR 052177-01; PI: David Cella).

References

1. North American Nursing Diagnosis Association. Nursing diagnoses: Definition and Classification, 1997–1998. Philadelphia, PA: McGraw-Hill; 1996.
2. Cella D. Factors influencing quality of life in cancer patients: Anemia and fatigue. *Semin Oncol* 1998;25(3 Suppl 7):43–46. [PubMed: 9671330]
3. Cella D, Davis K, Breitbart W, Curt G. Cancer-related fatigue: Prevalence of proposed diagnostic criteria in a United States sample of cancer survivors. *J Clin Oncol* 2001;19(14):3385–3391. [PubMed: 11454886]
4. Mooney K, Ferrell BR, Nail LM, et al. Oncology Nursing Society research priorities survey. *Oncology Nursing Forum* 1991;18:1381–1388. [PubMed: 1762979]
5. Stone P, Richardson A, Ream E, Smith AG, Kerr DJ, Kearney N. Cancer-related fatigue: Inevitable, unimportant and untreatable? Results of a multi-centre patient survey. *Cancer Fatigue Forum. Ann Oncol* 2000;11(8):971–975. [PubMed: 11038033]
6. Hann DM, Jacobsen PB, Azzarello LM, Martin SC, Curran SL, Fields KK, Greenberg H, Lyman G. Measurement of fatigue in cancer patients: Development and validation of the Fatigue Symptom Inventory. *Qual Life Res* 1998;7(4):301–310. [PubMed: 9610214]
7. Mendoza TR, Wang XS, Cleeland CS, Morrissey M, Johnson BA, Wendt JK, Huber SL. The rapid assessment of fatigue severity in cancer patients: Use of the Brief Fatigue Inventory. *Cancer* 1999;85(5):1186–1196. [PubMed: 10091805]
8. Piper BF, Dibble SL, Dodd MJ, Weiss MC, Slaughter RE, Paul SM. The revised Piper Fatigue Scale: Psychometric Evaluation in Women with Breast Cancer. *Oncology Nursing Forum* 1998;25(4):677–684. [PubMed: 9599351]
9. Smets EMA, Garssen B, Bonke B, DeHaes JCJM. The Multidimensional Fatigue Inventory (MFI) psychometric qualities of an instrument to assess fatigue. *Journal of Psychometric Research* 1995;39:315–325.
10. Yellen SB, Cella DF, Webster K, Blendowski C, Kaplan E. Measuring fatigue and other anemia-related symptoms with the Functional Assessment of Cancer Therapy (FACT) measurement system. *J Pain Symptom Manage* 1997;13(2):63–74. [PubMed: 9095563]
11. Lai J-S, Cella D, Dineen K, Von Roenn J, Gershon R. An item bank was created to improve the measurement of cancer-related fatigue. *J Clin Epidemiol* 2005;58(2):190–197. [PubMed: 15680754]
12. Ware, JE.; Snow, KK.; Kosinski, M. SF-36 Health Survey: Manual and Interpretation Guide. Lincoln, RI: QualityMetric Incorporated; 2000.
13. Varni JW, Burwinkle TM, Katz ER, Meeske K, Dickinson P. The PedsQL in pediatric cancer: Reliability and validity of the Pediatric Quality of Life Inventory Generic Core Scales, Multidimensional Fatigue Scale, and Cancer Module. *Cancer* 2002;94(7):2090–2106. [PubMed: 11932914]
14. Lai J-S, Velozo C. Accuracy of recall in measuring the outcomes of ambulatory orthopedic surgery. *Journal of Rehabilitation Outcomes* 1997;1(6):14–21.
15. Stone AA, Broderick JE, Schwartz JE, Shiffman S, Litcher-Kelly L, Calvanese P. Intensive momentary reporting of pain with an electronic diary: Reactivity, compliance, and patient satisfaction. *Pain* 2003;104(1–2):343–351. [PubMed: 12855344]
16. Robinson MD, Clore GL. Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological Bulletin* 2002;128(6):934–960. [PubMed: 12405138]
17. Redelmeier DA, Kahneman D. Patients' memories of painful medical treatments: Realtime and retrospective evaluations of two minimally invasive procedures. *Pain* 1996;66(1):3–8. [PubMed: 8857625]
18. Cella D, Yount S, Sorensen M, Chartash E, Sengupta N, Grober J. Validation of the Functional Assessment of Chronic Illness Therapy Fatigue Scale relative to other instrumentation in patients with rheumatoid arthritis. *J Rheumatol* 2005;32:811–819. [PubMed: 15868614]
19. Cella D, Lai J-S, Chang CH, Peterman A, Slavin M. Fatigue in cancer patients compared with fatigue in the general United States population. *Cancer* 2002;94(2):528–538. [PubMed: 11900238]
20. National Comprehensive Cancer Network. NCCN Practice Guidelines in Oncology: Cancer-Related Fatigue. 2004. www.nccn.org

21. Lai J-S, Dineen K, Reeve B, Von Roenn J, Shevrin D, McGuire M, Bode R, Paice J, Cella D. An item response theory based pain item bank can enhance measurement precision. *J Pain Symptom Manage* 2005;30(3):278–288. [PubMed: 16183012]
22. Lai JS, Cella D, Peterman A, Barocas J, Goldman S. Anorexia/cachexia related quality of life for children with cancer: Testing the psychometric properties of the Pediatric Functional Assessment of Anorexia/Cachexia Therapy (peds-FAACT). *Cancer* 2005;104(7):1531–1539. [PubMed: 16088963]
23. Baker, FB. The basics of Item Response Theory. College Park, MD: ERIC Clearinghouse on Assessment and Evaluation; 2001.
24. Linacre, JM. WINSTEPS: Rasch analysis for all two-facet models. Chicago, IL: MESA Press; 2002.
25. Lai J-S, Teresi JA, Gershon R. Procedures for the analysis of Differential Item Functioning (DIF) for small sample sizes. *Evaluation and the Health Professions* 2005;28:283–294.
26. Zumbo, BD. A handbook on the theory and methods of differential item functioning (DIF): Logistic regression modeling as a unitary framework for binary and likert-type (ordinal) item scores. Ottawa, ON: Directorate of Human Resources Research and Evaluation, Department of National Defense; 1999.
27. Beese A, Morley S. Memory for acute pain experience is specifically inaccurate but generally reliable. *Pain* 1993;53(2):183–189. [PubMed: 8336988]
28. Bryant RA. Memory for pain and affect in chronic pain patients. *Pain* 1993;54(3):347–351. [PubMed: 8233551]
29. Pearce SA, Isherwood S, Hrouda D, Richardson PH, Erskine A, Skinner J. Memory and pain: Tests of mood congruity and state dependent learning in experimentally induced and clinical pain. *Pain* 1990;43(2):187–193. [PubMed: 2087330]
30. Smith WB, Safer MA. Effects of present pain level on recall of chronic pain and medication use. *Pain* 1993;55(3):355–361. [PubMed: 8121697]

Figure 1.
Comparisons of item information function curves between timeframes

Figure 2.
Scale information function curve between 7-day and 4-week timeframes.

Table 1

Patient demographic and clinical information (by timeframe of questionnaire)

		Time Frame	
		4-week (N=116)	7-day (N=100)p-value ^d
Gender	Male	35.6%	37.4%.794
	Female	64.4%	62.6%
Cancer Type	Breast	40.4%	31.6% ^e
	Colo-rectal	15.8%	12.2%
	Non-Hodgkin	12.3%	6.1%
	Ovarian	4.4%	8.2%
	Lung	3.5%	9.2%
	Hodgkin	2.6%	4.1%
	Bladder	0.9%	5.1%
	Prostate	0.9%	6.1%
Ethnicity ^a	Hispanic origin	4.4%	4.0%.911 ^a
	White	73.0%	87.9%.007
	African American	16.2%	6.1%.021
	Asian American	5.4%	5.1%.908
Treatment ^a	Chemotherapy	62.0%	85.9%<.001 ^a
	Radiation	7.8%	5.1%.413
	Surgery	50.0%	53.5%.607
Extent of disease	No evidence of disease	19.8%	5.6%.033
	Local	18.9%	17.2%
	Regional	4.5%	5.8%
	Metastasis	41.4%	58.6%
	NA	15.3%	12.6%
ECOG PSR ^b	Normal activity	35.7%	25.3%.049
	Some symptom	42.6%	53.5%
	Bed rest < 50%	20.0%	14.1%
	Bed rest > 50%	1.7%	7.1%
Age		54.7 (SD=13.9)	60.4 (SD=12.0).004
Months since diagnosis		76.4 (SD=88.3)	91.0 (SD=112.9)<.001
NCCN Fatigue (prior) ^c	Mild (0–3)	65.2%	56.0%.209
	Moderate (4–6)	21.6%	28.0%
	Severe (7–10)	12.9%	16.0%
NCCN Fatigue (post) ^c	Mild (0–3)	63.8%	50.0%.074
	Moderate (4–6)	23.3%	33.0%
	Severe (7–10)	12.9%	17.0%
FACIT-F score, Mean		36.6 (SD=10.8)	35.9 (SD=12.2).892

^aPatients could endorse more than one category; p-value was calculated for each category.^bECOG PSR = Eastern Cooperative Oncology Group Performance Status Rating^cNCCN Fatigue: patients were asked to rate their fatigue by using a 0–10 single item at the beginning of the surveying (i.e., pre-fatigue) and at the end of the surveying (i.e., post-fatigue). Patients with scores of 0–3 were considered “mild”, 4–6 were considered “moderate”, and 7–10 were considered “severe”.^dp values were calculated to estimate significantly different frequency between 7-days and 4-weeks^e“cancer type” was not calculated due to small sample size of each group

Table 2

Item descriptive statistics

	4-Week		7-Day		Trend ^a p-value		CMH ^b p-value	
	Mean	SD	Mean	SD				
I have energy	2.21	1.06	2.11	1.19	0.262		0.526	
I am able to do my usual activities	2.4	1.15	2.4	1.22	0.491		0.983	
I feel fatigued	2.44	1.13	2.53	1.23	0.286		0.573	
I feel weak all over	3.02	1.07	2.98	1.26	0.407		0.814	
I feel listless	2.87	1.11	2.76	1.3	0.249		0.498	
I feel tired	2.56	1.07	2.52	1.17	0.395		0.791	
I have trouble starting things because I am tired	2.87	1.08	2.76	1.22	0.234		0.470	
I have trouble finishing things because I am tired	2.81	1.09	2.76	1.16	0.371		0.742	
I need to sleep during the day	2.87	1	3.06	0.97	0.080		0.161	
I am too tired to eat	3.67	0.62	3.56	0.77	0.116		0.234	
I need help doing my usual activities	3.29	1.04	3.12	1.08	0.119		0.239	
I am frustrated by being too tired to do things I want to do	2.81	1.23	2.68	1.26	0.221		0.442	
I have to limit my social activities because I am tired	2.75	1.22	2.68	1.28	0.334		0.668	

^a Cochran-Armitage trend test.

^b Cochran-Mantel-Haenszel statistics

TABLE 3

Time frame used to endorse items and patient preferred time frame

Timeframe options	Timeframe used to answer questions (%)		Best timeframe for these questions (%)	
	7-day	4-week	7-day	4-week
Today	17.2	25.9	6.2	11.2
Past few day	5.1	1.7	4.1	1.7
Past 1 week	64.6	6.9	35.1	11.2
Past 2 weeks	5.1	6.9	16.5	18.1
Past 4 weeks	3.0	44.8	24.7	41.4
more than 4 weeks Ago	5.1	13.8	13.4	16.4