Comparison of Legacy Fatigue Measures With the PROMIS Pediatric Fatigue Short Form

Catherine F. Macpherson, RN, PhD, CPON®, Jichuan Wang, PhD, Darren A. DeWalt, MD, MPH, Emily D. Stern, RN, BSN, CPON®, CCRP, Shana Jacobs, MD, and Pamela S. Hinds, PhD, RN, FAAN

OBJECTIVES: To compare Patient-Reported Outcomes Measurement Information System (PROMIS) Pediatric Fatigue Short Form measures and legacy patient-reported outcome fatigue measures to capture cancer-related fatigue change in pediatric patients with cancer.

SAMPLE & SETTING: 96 racially diverse children and adolescents with cancer. The study occurred in a 32-bed inpatient unit and three regional outpatient clinics.

METHODS & VARIABLES: The Fatigue Scale–Child, Fatigue Scale–Adolescent, and the PROMIS Pediatric Fatigue Short Form measures were administered at three time points during chemotherapy. Descriptive, correlational, psychometric, and receiver operating characteristic (ROC) curve analyses were conducted. The variable was pediatric patient-reported fatigue.

RESULTS: All measures were positively correlated at each time point. ROC curves were not statistically different from each other at any data point.

IMPLICATIONS FOR NURSING: Nurses have psychometrically strong options for measuring cancer-related fatigue in pediatric patients with cancer, but the PROMIS Pediatric Fatigue Short Form is applicable to more age groups.

KEYWORDS pediatric oncology; PROMIS; fatigue; patient-reported outcomes

ONF, 45(1), 106–114.
DOI 10.1188/18.ONF.106-114

Children, adolescents, and their parents, as well as adult survivors of childhood cancer, report cancer-related fatigue (CRF) to be a prevalent, intense, and distressing symptom that has a significant adverse effect on their health-related quality of life during cure-directed therapy, end-of-life care, and survivorship (Hsiao et al., 2017; Kestler & LoBiondo-Wood, 2012; Rach et al., 2017; Tomlinson et al., 2016; Wolfe et al., 2015; Zhukovsky et al., 2015). Although CRF can vary by age, gender, type of cancer treatment, and point in treatment (Dobrozsi, Yan, Hoffmann, & Paneپint, 2017; Hinds, Hockenberry, Tong, et al., 2007; Sanford et al., 2008; Tomlinson et al., 2013; Vallance et al., 2010), it is a near universal and troubling symptom often associated with worse mood, pain, distress, and disrupted sleep (Hinds, Hockenberry, Gattuso, et al., 2007; Rach et al., 2017; Sanford et al., 2008).

Although no standardized treatment for CRF exists, it is often amenable to interventions, such as yoga, music, walking, other forms of activity, rest, and distraction (Baumann, Bloch, & Beulertz, 2013; Danhauer, Addington, Sohl, Chaoul, & Cohen, 2017; Lopes-Júnior et al., 2016; Tomlinson et al., 2016). Because of the prevalence and disruptiveness of CRF in children and adolescents and their responsiveness to certain interventions, being able to accurately and sensitively measure CRF by patient report is essential to identifying those who are at highest risk and to determine the interventions that will most relieve them of this unpleasant symptom.

Measuring Cancer-Related Fatigue in Children and Adolescents

Symptoms like CRF have been assessed in pediatrics with commonly used metrics referred to as legacy measures. Often, legacy measures focus on assessment of a specific symptom in the context of a specific disease, such as cancer. Pediatric legacy measures also tend to be created for a specific developmental
stage. Among the legacy pediatric CRF measures are the Fatigue Scale–Child (FS-C) and Fatigue Scale–Adolescent (FS-A). The FS-C and FS-A have been identified in two reviews of pediatric fatigue measures as two of the three most recommended CRF measures for use in clinical research given their strong psychometric properties (Chrichton, Knight, Oakley, Babi, & Anderson, 2015; Tomlinson et al., 2013). These two measures are supported by qualitative documentation of the need for child and adolescent forms; are well-tested in terms of reliability, validity, and ability to detect change over time; and have cut scores to indicate the presence of high fatigue (Hinds et al., 1999; Hinds, Hockenberry, Tong, et al., 2007; Hockenberry et al., 2003; Mandrell et al., 2011; Tomlinson et al., 2013). Despite these strengths, legacy measures have limitations. They were developed using the classical test theory rather than the item response theory. Child- and adolescent-specific forms result in a need to sample for both age groups, posing a challenge in studies of a small population of patients with rare disease. The use of different age-specific legacy measures for a given symptom also confounds comparison of scores across studies, disease states, and developmental stages (Hinds et al., 2013).

In contrast to the legacy measures, the measures of the pediatric Patient-Reported Outcomes Measurement Information System (PROMIS) were developed using sophisticated analytic strategies and advanced administration and scoring technology, and should facilitate age, place, and time comparisons (Hinds et al., 2013). The adult PROMIS fatigue measure has been reconciled with the adult legacy CRF measures, and they are highly correlated (Barsevick et al., 2013). Reconciling the pediatric legacy CRF measures and the PROMIS Pediatric Fatigue measure has been recommended (Barsevick et al., 2013). Reconciliation of the pediatric measures may be of even greater value than reconciliation of the adult measures because of age specificity of the legacy measures, which leads to the issue of patients “aging out” of measures in longitudinal studies (Barsevick et al., 2013).

Feasibility, acceptability, reliability, and known-group validity of the original eight PROMIS pediatric measures have been established in pediatric patients with cancer in cure-directed treatment, in survivorship, and in enrollment in phase 1 or 2 trials for incurable cancer (Dobrozs, et al., 2017; Hinds et al., 2013, 2017). Evaluating the measures’ responsiveness to capture change over time is essential to inform their use in clinical trials. Establishing the PROMIS measures’ sensitivity to change over time would enhance their use as patient-reported outcome measures of symptom experience and functional status and create a primary endpoint of child voice to augment the traditional trial endpoints of disease status and treatment-related toxicities. To this end, the authors have conducted a head-to-head comparison of legacy measures versus PROMIS pediatric measures that concurrently measure the same construct.

The purpose of this study was to directly compare the PROMIS Pediatric Fatigue Short Form with the FS-C and FS-A in regard to their ability to discriminate among pediatric patients with cancer with high and low fatigue at three time points within a single course of chemotherapy, and to compare the psychometric properties of the legacy measures (FS-C and FS-A) with those of the PROMIS Pediatric Fatigue Short Form.

Methods
Sample and Setting
The single site study was conducted in the Division of Oncology at Children’s National Health System in Washington, DC. Annually for the past four years, 154–220 newly diagnosed children and adolescents have been admitted to the 32-bed inpatient unit at the Children’s National Health System main hospital or have been seen in one of three regional outpatient clinics: Children’s National Health System main hospital, Pediatric Specialists of Virginia Center for Cancer and Blood Disorders, and the Montgomery County Outpatient Center. Data were collected during inpatient and outpatient admissions.

Procedures
Following approval by the Children’s National Health System’s Institutional Review Board, eligible patients were identified during daily review of patient admissions and clinic visits and confirmation of eligibility with the patients’ clinical care team. A parent or legal guardian of the eligible patients was approached by one of three study team members trained in research consent and assent processes. Patients were considered eligible if they were aged 8–18 years, were diagnosed with cancer, were receiving chemotherapy, and could speak and read English at the time of the study (the legacy patient-reported measures were available in more than one language, but the PROMIS measure was not). Written permission was obtained from the parents or guardians of participants aged younger than 18 years and written assent was obtained from the participants themselves, whereas written consent was obtained from participants aged 18 years or older. Demographic data collected at study enrollment included patient age, gender, ethnicity, race, grade.
in school, living arrangement, other current health concerns or history of health concerns, guardian’s gender and age, race and ethnicity, and relation to the patient, marital status, highest grade completed by the legal guardian, and occupation. Clinical data collected at each data point included the patient’s hemoglobin and white blood count. Study measures were primarily completed on a laptop or desktop computer in a private clinic area or inpatient room, allowing for measures to be randomly sequenced. In a limited number of instances, paper copies were used to complete study measures.

Design
In this descriptive, longitudinal design, participants completed study measures at three data points—the beginning of a course of chemotherapy (T1), at count nadir (T2), and just before the beginning of the next course of chemotherapy (T3). All participants completed the PROMIS Pediatric Fatigue Short Form and the Symptom Distress Scale (SDS). Participants aged 8–12 years completed the FS-C, and participants aged 13–18 years completed the FS-A. The average time interval was 11.1 days (SD = 3.2) between T1 and T2, 18.4 days (SD = 10.7) between T2 and T3, and 28.6 days (SD = 10.8) between T1 and T3.

Measures
Symptom Distress Scale fatigue item: The SDS is a self-report instrument that measures distress associated with 10 common symptoms, including fatigue, on a five-point Likert-type scale (McCorkle & Young, 1978). Higher scores represent greater symptom distress, and individual item scores are summed for a total scale score (McCorkle & Young, 1978). The SDS is commonly used to measure symptom distress in adults with cancer, but has also been validated for use with children and adolescents with cancer (Hinds, Schum, & Srivastava, 2002). Although reliability and validity assessment of single items are limited, use of single SDS items is recommended for measurement of distress related to a specific symptom, which is distinct from the overall burden of symptoms measured by the SDS total score (Hinds et al., 2002; McCorkle & Young, 1978). A single item score of 3 or more was identified as a cut score indicative of clinically significant symptom distress in a study of 49 children and adolescents with cancer (Hinds et al., 2002). The SDS fatigue item ranges from 1 (I am not tired at all) to 5 (could not feel more tired).

Fatigue Scale–Child: The original FS-C was a 14-item self-report instrument that measured the extent to which each of the 14 statements described how the respondent had been feeling during the past seven days on a five-point Likert-type scale ranging from 0 (not at all) to 4 (all the time). The potential score range was 0–56, with higher scores representing greater fatigue. The FS-C was developed from data derived from focus group interviews with children aged 7–12 years with cancer, and content validity was established by an expert panel of pediatric oncology clinicians (Hockenberry et al., 2003). On initial testing with 149 children aged 7–12 years receiving treatment for cancer, the internal consistency reliability estimate (Cronbach alpha) was 0.84 and item-total correlations ranged from 0.34–0.6. Three factors emerged from an

TABLE 1. Sample Characteristics (N = 96)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td></td>
</tr>
<tr>
<td>8–12</td>
<td>40</td>
</tr>
<tr>
<td>13–18</td>
<td>56</td>
</tr>
<tr>
<td>Child gender</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>52</td>
</tr>
<tr>
<td>Female</td>
<td>44</td>
</tr>
<tr>
<td>Child race</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>42</td>
</tr>
<tr>
<td>Black or African American</td>
<td>23</td>
</tr>
<tr>
<td>Hispanic</td>
<td>19</td>
</tr>
<tr>
<td>Other</td>
<td>12</td>
</tr>
<tr>
<td>Fatigue screening*</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>48</td>
</tr>
<tr>
<td>Yes</td>
<td>46</td>
</tr>
<tr>
<td>Missing</td>
<td>2</td>
</tr>
<tr>
<td>Highest education of guardian</td>
<td></td>
</tr>
<tr>
<td>Less than 8th grade</td>
<td>6</td>
</tr>
<tr>
<td>Some high school</td>
<td>3</td>
</tr>
<tr>
<td>High school diploma or GED</td>
<td>15</td>
</tr>
<tr>
<td>Some college or technical degree</td>
<td>18</td>
</tr>
<tr>
<td>College degree</td>
<td>32</td>
</tr>
<tr>
<td>Advanced degree</td>
<td>21</td>
</tr>
<tr>
<td>Missing</td>
<td>1</td>
</tr>
<tr>
<td>Other health problems</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>64</td>
</tr>
<tr>
<td>Yes</td>
<td>32</td>
</tr>
<tr>
<td>Type of tumor</td>
<td></td>
</tr>
<tr>
<td>Lymphoma or acute lymphoblastic leukemia</td>
<td>56</td>
</tr>
<tr>
<td>Solid tumor</td>
<td>35</td>
</tr>
<tr>
<td>Brain tumor</td>
<td>5</td>
</tr>
</tbody>
</table>

*Derived from the response to item 3 on the Symptom Distress Scale. Scores of 3 or greater indicated a positive response.
exploratory factor analysis using a principal components factor extraction and oblique rotation: lack of energy, not able to function, and altered mood (Hockenberry et al., 2003). Construct validity estimates involving the FS-C and the Fatigue Scale–Parent were higher than those involving the Fatigue Scale–Staff. Evaluation of convergent validity of the FS-C with the Birleson Depression Self-Rating Scale showed significant correlation of the total scales ($r = 0.537, p < 0.00$) and all subscales (Hockenberry et al., 2003). The FS-C also demonstrated a trend toward ability to measure change over time ($t[14] = 1.967, p = 0.069$) (Hockenberry et al., 2003).

Using Rasch modeling of data drawn from three studies involving 221 children receiving treatment for cancer, Hinds et al. (2010) reduced the FS-C from 14 to 10 items that identify those children with high CRF (Hinds et al., 2010). Cronbach alpha for the revised FS-C was 0.76, and concurrent and construct validity (same three domains confirmed) remained acceptable (Hinds et al., 2010). Receiving operating characteristic (ROC) curve analysis was then used to identify a cut score of 12 on the revised FS-C, with 75% sensitivity and 73.5% specificity (Hinds et al., 2010).

Fatigue Scale–Adolescent: The original FS-A was a 14-item self-report instrument that measured how the respondent had been feeling during the past seven days on a five-point Likert-type scale ranging from 1 (not at all) to 5 (all the time) (Hinds, Hockenberry, Tong, et al., 2007). The potential score range was 14–70, and higher scores represented greater fatigue (Hinds, Hockenberry, Tong, et al., 2007). The FS-A was developed from data derived from individual and focus group interviews with adolescents with cancer aged 13–18 years (Hinds, Hockenberry, Tong, et al., 2007). Content validity was established by an expert panel consisting of adolescents with cancer and healthcare professionals (Hinds, Hockenberry, Tong, et al., 2007). On initial testing with 64 adolescents receiving treatment for cancer who completed the FS-A at two to four key points in treatment in four studies, internal consistency reliability estimates ranged from 0.87–0.95 (Hinds, Hockenberry, Tong, et al., 2007). Item-total correlations ranged from 0.24–0.92, except for item 10, which had a low correlation at one time point in one study. No evidence of multicollinearity was found among items, and no item-to-item correlation coefficient exceeded 0.8 (Hinds, Hockenberry, Tong, et al., 2007). Four factors emerged from an exploratory factor analysis using a principal components factor extraction with varimax orthogonal factor rotation (Hinds, Hockenberry, Tong, et al., 2007). These were cognitive and physical weariness, added effort and assistance needed to do usual activities, needing rest and feeling angry, and avoiding social activities (Hinds, Hockenberry, Tong, et al., 2007). As with the FS-C, construct validity estimates involving the FS-A and the Fatigue Scale–Parent were higher than those involving the Fatigue Scale–Staff (Hinds, Hockenberry, Tong, et al., 2007). Construct validity estimates involving the FS-A and the Reynolds Depression Scale were 0.87 (Hinds, Hockenberry, Tong, et al., 2007). The FS-A demonstrated known groups validity in discriminating between anemic and nonanemic

### TABLE 2. Descriptive Findings for Each Fatigue Scale by Time Point

<table>
<thead>
<tr>
<th>Measure</th>
<th>Time Point 1 (N = 96)</th>
<th>Time Point 2 (N = 84)</th>
<th>Time Point 3 (N = 86)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\bar{X}$</td>
<td>SD</td>
<td>Range</td>
</tr>
<tr>
<td>FS-C</td>
<td>23.3</td>
<td>9.1</td>
<td>11–49</td>
</tr>
<tr>
<td>FS-A</td>
<td>31.2</td>
<td>13.2</td>
<td>13–59</td>
</tr>
<tr>
<td>PROMIS A</td>
<td>53</td>
<td>14.2</td>
<td>30–80</td>
</tr>
<tr>
<td>PROMIS C</td>
<td>51.3</td>
<td>13</td>
<td>30–85</td>
</tr>
</tbody>
</table>

FS-A—Fatigue Scale–Adolescent; FS-C—Fatigue Scale–Child; PROMIS A—Patient-Reported Outcomes Measurement Information System Pediatric Fatigue Short Form completed by adolescents; PROMIS C—PROMIS Pediatric Fatigue Short Form completed by children

**Note.** The FS-A measures the extent to which 14 statements describe how the respondent has been during the past seven days on a five-point Likert-type scale ranging from 1 (not at all) to 5 (all the time). The total range is 14–70, with higher scores representing greater fatigue. The FS-C measures the extent to which 14 statements describe how the respondent has been feeling during the past seven days on a five-point Likert-type scale ranging from 0 (not at all) to 4 (all the time). The total range is 0–56, with higher scores representing greater fatigue. The PROMIS Pediatric Fatigue Short Form consists of 10 items scored on a Likert-type scale ranging from 1 (never) to 5 (almost always). The total possible range is 10–50, with higher scores representing greater fatigue.
patients (p = 0.042), and demonstrated the ability to measure change over time (t = 2.55, p < 0.01) (Hinds et al., 2013). In a study using Rasch modeling of data drawn from nine studies involving 138 adolescents receiving treatment for cancer, the FS-A was reduced from 14 to 13 items that identify adolescents with high CRF. Cronbach alpha for the revised FS-A was 0.87, and concurrent and construct validity remained acceptable (Mandrell et al., 2011). ROC curve analysis was then used to identify a cut score of 31 on the revised FS-A, with 66.6% sensitivity and 82.6% specificity (Mandrell et al., 2011).

PROMIS Pediatric Fatigue Short Form: This measure consists of 10 items. The recall period is seven days. Five response options ranging from “never” to “almost always” are used. Higher scores represent greater fatigue. Scales are given a t score, with the mean of the reference sample being 50 and standard deviation being 10 (Hinds et al., 2013). Low scores and very low scores are 40 and 30, respectively, whereas high scores and very high scores are 60 and 70, respectively (Hinds et al., 2013). Hinds et al. (2013) has provided psychometric properties. The measure has been tested in a large and diverse group of children and adolescents in cure-directed cancer treatment and in survivors of childhood cancer, in a group of children and adolescents aged 8–18 years enrolled in phase 1 and 2 clinical trials for incurable cancer; and in a group of 40 pediatric patients with cancer during the first six months of their treatment (Dobrozsi et al., 2017; Hinds et al., 2013, 2017). The findings indicate that the PROMIS Pediatric Fatigue Short Form measure is unidimensional, internally consistent and has criterion validity (Dobrozsi et al., 2017; Hinds et al., 2013, 2017). Using additional items increased reliability, but short forms consistently demonstrate reliability of 0.85 over 2 to 4 standard deviations (Hinds et al., 2013).

Demographic Form

Items on the demographic form included age, gender, ethnicity, race, site of recruitment (inpatient, clinic 1, clinic 2), child’s highest grade completed in school, living arrangements, other health problems or concerns, type of cancer, age of child when first diagnosed with cancer, if a recurrence had been diagnosed, number of medications taken each day, time since most recent treatment, parent/guardian gender and age, ethnicity and race, relationship to the ill child, marital status, highest grade completed, and occupational status. Six additional items asked parents to identify if, in the past seven days, the ill child has had low blood counts, not eaten well, not slept well, been more tired than usual or felt nauseated or vomited, and the number of hospitalizations in the previous six months.

Analysis

Descriptive statistics were applied to the fatigue scales and to the demographic items at each of the three data collection points. Pearson product-moment correlations were calculated to assess the relationships between the legacy fatigue measures and the PROMIS Pediatric Fatigue Short Form at each data collection point. ROC curve analyses were conducted to test the ability of the scales to discriminate between participants with high and low fatigue. The SDS single-item fatigue scores reported by participants were dichotomized to represent participants with no-to-low fatigue (score of less than 3) and participants with high fatigue (3 or higher). This dichotomous measure was used as the outcome in the ROC curve analyses. The ROC curve derived from the PROMIS Pediatric Fatigue Short Form was compared to those derived from FS-C and FS-A at each data collection point. When comparing the curves, the correlation between the areas under the ROC curve induced by the paired nature of the data was taken into consideration by using a nonparametric approach to the Mann-Whitney U test (DeLong, DeLong, & Clarke-Pearson, 1988).

Results

Sample

Of the 96 participants, 56 were aged 13–18 years, and 52 were male. The patients were racially diverse, and most guardians reported having more than a high school education. More than half the patients had...
acute leukemia or lymphoma, and about two-thirds reported no other health problems (see Table 1).

**Descriptive Fatigue Findings**

Descriptively, the mean score for the FS-C was highest at T2 and lowest at T3. The mean score for the FS-A was very similar at T1 and T2 but 6 points lower at T3. Similarly, the PROMIS Pediatric Fatigue Short Form mean score was highest at T2 and lowest at T3 (see Table 2). Based on the SDS item score of 3 or greater, almost half of the sample reported high fatigue at both T1 (n = 46) and T2 (n = 41); fewer participants (n = 36) reported high fatigue at T3.

**Reliability**

The reliability of each of the fatigue scales at each time point as measured by Cronbach alpha exceeded the accepted level of internal consistency for mature scales (see Table 3).

**Validity**

The correlations between the legacy fatigue measures and the PROMIS Pediatric Fatigue Short Form were consistently strong (0.85–0.9) across the three data collection points between the FS-A and the PROMIS measure and moderate to strong between the FS-C and the PROMIS measure (see Table 4). The values of the area under the curve (AUC), which is commonly used as a summary measure of diagnostic accuracy, are shown in Table 5 by fatigue measure and time point. When applied to adolescents, the AUC values for the PROMIS Pediatric Fatigue Short Form all were greater than 0.8 at all three time points, indicating that this measure has a good ability to discriminate between pediatric patients with cancer with high and low levels of fatigue. When applied to children, the AUC values were also good (greater than 0.8) at T2 and T3 but fair (0.7–0.8) at T1. Although the AUC values for the FS-A and FS-C were slightly better than those for PROMIS Pediatric Fatigue Short Form at T1, the differences were not statistically significant. In addition, the AUC values at T2 and T3 for the PROMIS Pediatric Fatigue Short Form and the legacy fatigue measures were almost identical.

**Discussion**

In the study sample, which was measured across three time points where change in fatigue intensity was anticipated, the authors documented important and even striking similarities in fatigue patterns, internal consistency, and concurrent validity between the legacy fatigue measures and the PROMIS Pediatric Fatigue Short Form. The similarity findings were revealed among the current study sample (N = 96), which had about an equal gender distribution and age distribution and notable diversity in race. However, type of tumor did not have such distribution, and the distribution does not reflect the typical prevalence of these diagnoses in pediatrics, indicating a need to purposefully enroll a greater proportion of pediatric patients with brain tumors in future research.

Importantly, based on the SDS score, about half of the study sample screened positive for fatigue at the first two time points (T1 = 48.9%, T2 = 48.8%) and fewer at the final data point (T3 = 42.4%), making the patient-reported scores from this sample useful for reliability and validity analyses. Additional descriptive results indicate that the two legacy measures and the PROMIS Pediatric Fatigue Short Form measure had the same pattern of scores across the three time points: All measures had quite similar mean scores at the first (T1) and second (T2) data points and a decrease in the mean scores at the final data point (T3). The change in the proportion of fatigued participants and the change in the mean scores represent clinically meaningful changes, and the high similarity in score patterns is a strong confirmation of the similar validity of the legacy fatigue measures with the PROMIS Pediatric Fatigue Short Form.

The examined internal consistency by fatigue measure and by age (8–12 years, 13–18 years) is impressive.

---

**TABLE 4. Correlation of Legacy and Fatigue Measures and PROMIS Pediatric Fatigue Short Form**

<table>
<thead>
<tr>
<th>Measure</th>
<th>Fatigue Scale–Child</th>
<th>Fatigue Scale–Adolescent</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Correlation</td>
</tr>
<tr>
<td>PROMIS T1</td>
<td>28</td>
<td>0.65</td>
</tr>
<tr>
<td>PROMIS T2</td>
<td>24</td>
<td>0.88</td>
</tr>
<tr>
<td>PROMIS T3</td>
<td>26</td>
<td>0.75</td>
</tr>
</tbody>
</table>

PROMIS—Patient-Reported Outcomes Measurement Information System; T—time point
for the high coefficients (high reliability) for each measure at each data point. Each of the measures performed at a similarly high reliability level across time, indicating that these measures are similarly internally consistent in this study sample.

The outcomes of the diverse approaches to estimating validity of each of the three fatigue measures produced statistically similar and significant findings. The correlations between each legacy measure and the PROMIS Pediatric Fatigue Short Form measure at each time point were strongly and statistically significantly positive. The FS-A scale was highly correlated with the PROMIS Pediatric Fatigue Short Form at all three data points (0.85–0.9), and the FS-C was moderately to strongly correlated at the three data points (0.65–0.88). In the AUC analyses by fatigue measures and data point, the AUC values for the PROMIS Pediatric Fatigue Short Form were not significantly different from that of the FS-C or FS-A at each of the three time points, indicating similar sensitivity and specificity to the construct of pediatric fatigue across the compared measures over time.

The ability of the PROMIS Pediatric Fatigue Short Form to discriminate among pediatric patients with cancer with high and low fatigue over three time points within a single course of chemotherapy supports its use by researchers and clinicians seeking to accurately assess change in CRF over time in relation to disease status, treatment-related toxicities, or symptom management intervention directed at fatigue amelioration.

In addition, the comparable psychometric properties and ability of the PROMIS Pediatric Fatigue Short Form and the legacy fatigue measures (FS-C and FS-A) to discriminate between pediatric patients with cancer over time helps to establish the validity of the newer PROMIS measure in relation to the older gold standard legacy measures. Comparability of the measures facilitates comparison of results across studies, which is particularly important because few studies of CRF in children and adolescents are completed and they tend to have small sample sizes. To further enhance comparability across studies with different instruments, a formal instrument-linking study could be performed, but the sample size of the current study is not sufficient (Reeve et al., 2016).

### Limitations
This study contains important limitations that warrant discussion. The authors used the SDS single-item fatigue as the criterion standard in their ROC curve analyses. The dichotomized single item is an imperfect gold standard but served the authors well to facilitate comparison of the FS-C, FS-A, and PROMIS Pediatric Fatigue Short-Form questionnaires. For the purposes of this study, the criterion standard did not need to be a perfect standard of fatigue, only a standard for instrument comparisons. The sample was relatively small, but the similar results across instruments and statistically significant findings provide reassurance about the similarities among the instruments. The authors did not have an adequate sample

<table>
<thead>
<tr>
<th>Measure</th>
<th>Time Point 1</th>
<th>Time Point 2</th>
<th>Time Point 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUC</td>
<td>p</td>
<td>AUC</td>
<td>p</td>
</tr>
<tr>
<td>FS-A</td>
<td>0.93</td>
<td>0.0615</td>
<td>0.85</td>
</tr>
<tr>
<td>PROMIS A</td>
<td>0.87</td>
<td>0.0615</td>
<td>0.86</td>
</tr>
<tr>
<td>FS-C</td>
<td>0.84</td>
<td>0.3031</td>
<td>0.84</td>
</tr>
<tr>
<td>PROMIS C</td>
<td>0.72</td>
<td>0.3031</td>
<td>0.83</td>
</tr>
</tbody>
</table>

**Knowledge Translation**

- Legacy pediatric patient-reported fatigue scales have been reconciled with the Patient-Reported Outcomes Measurement Information System (PROMIS) Pediatric Fatigue Short Form.
- The Fatigue Scale–Child, Fatigue Scale–Adolescent, and the PROMIS Pediatric Fatigue Short Form are equally psychometrically sound pediatric patient-reported forms for measuring fatigue in patients aged 7–18 years with cancer.
- Because of the wider respondent age, the PROMIS Pediatric Fatigue Short Form offers a more inclusive approach to studying children and adolescents with cancer.
size for subgroup comparisons by type of cancer. The effects of one type of cancer or treatment could lead to different qualities of fatigue that are assessed by questionnaires differently. The authors do not think that the legacy or PROMIS measures would be differentially valid by type of cancer.

**Implications for Nursing**

Nurses caring for patients aged 8–18 years receiving treatment for cancer have reliable and valid measures for assessing CRF. The FS-C, FS-A, and the PROMIS Pediatric Fatigue Short Form are psychometrically strong, providing nurses with options for measuring what appears to be a nearly universal symptom experience in this age group during treatment. However, when studying a patient group aged 8–18 years, nurses should use the PROMIS Pediatric Fatigue Short Form, which is more age inclusive.

Additional study of fatigue patterns is warranted to ascertain if any demographic or clinical covariates predict those who have high fatigue at multiple points across their treatment trajectory because these patients may be most likely to benefit from interventions for CRF. Examining differences in fatigue as measured with the PROMIS Pediatric Fatigue Short Form by developmental level (e.g., child versus adolescent) could also be of interest given the original conceptual foundations of the FS-C and FS-A.

**Conclusion**

The current study involved the first longitudinal comparison of the three fatigue measures. The results provide evidence that the PROMIS Pediatric Fatigue Short Form has nearly the same ability to discriminate child and adolescent patients with high and low fatigue at different time points across a single course of chemotherapy as the FS-C and FS-A measures. The PROMIS measure applies to child and adolescent participants, whereas the legacy measures apply to either child or adolescent participants. Therefore, the PROMIS measure is a more broadly applicable measure and produces reliable and valid results from children and adolescents during a course of chemotherapy.

Catherine F. Macpherson, RN, PhD, CPON®, is a staff nurse on the cancer care unit at Seattle Children’s Hospital in Washington; Jichuan Wang, PhD, is a biostatistician at the Children’s National Medical Center in Washington, DC; Darren A. DeWalt, MD, MPH, is an associate professor in the School of Medicine at the University of North Carolina in Chapel Hill; and Emily D. Stern, RN, BSN, CPON®, CCRP, is an RN, Shana Jacobs, MD, is a physician, and Pamela S. Hinds, PhD, RN, FAAN, is the director of the Department of Nursing Research, all at the Children’s National Medical Center. Macpherson can be reached at catherine.macpherson@seattlechildrens.org, with copy to ONFEditor@ons.org. (Submitted May 2017. Accepted August 2, 2017.)

This research was funded by a grant (8U01AR052181) from the National Institutes of Health, National Institute of Arthritis and Musculoskeletal and Skin Diseases. DeWalt developed and owns the copyright for the Patient-Reported Outcomes Measurement Information System (PROMIS) items under evaluation in this study and does not receive compensation for use of the items. Unlimited free license has been granted by the PROMIS Health Organization. During the writing of this article, DeWalt was supported through funding from the Agency for Healthcare Research and Quality and the Patient-Centered Outcomes Research Institute.

Macpherson, Wang, DeWalt, Jacobs, and Hinds contributed to the conceptualization and design. DeWalt, Jacobs, and Hinds completed the data collection. Wang and Hinds provided statistical support. Macpherson and Hinds provided the analysis. All authors contributed to the manuscript preparation.

**REFERENCES**


