Fatigue in Child Chronic Health Conditions: A Systematic Review of Assessment Instruments

Alison Crichton, DPsych, a, Sarah Knight, PhD, a, Ed Oakley, PhD, d, e, Franz E. Babl, MD, d, e, Vicki Anderson, PhD, c, e

abstract

BACKGROUND AND OBJECTIVE: Fatigue is common in chronic health conditions in childhood, associated with decreased quality of life and functioning, yet there are limited data to compare assessment instruments across conditions and childhood development. Our objective was to describe fatigue assessment instruments used in children with chronic health conditions and critically appraise the evidence for the measurement properties of identified instruments.

METHODS: Data sources included Medline, Cumulative Index to Nursing and Allied Health Literature, and PsycINFO (using the EBSCOhost platform). Study selection included quantitative assessment of fatigue in children with health conditions. Data extraction was as follows: (1) study design, participant and fatigue instruments, (2) measurement properties of fatigue instruments, (3) methodological quality of included studies, and (4) synthesis of the quality of evidence across studies for the measurement properties of fatigue instruments.

RESULTS: Twenty fatigue assessment instruments were identified (12 child reports, 7 parent reports, 1 staff report), used in 89 studies. Fatigue was assessed in over 14 health conditions, most commonly in children with cancer and chronic fatigue syndrome. Evidence for the measurement properties of instruments varied, and overall quality was low. Two fatigue instruments demonstrated strong measurement properties for use in children with diverse health conditions and children with cancer.

CONCLUSIONS: The review is limited to children younger than 18 years and results are specific to health conditions described, limiting generalizability of findings to other populations. Evidence for the measurement properties of fatigue instruments varied according to the population in which instruments were used and informant. Further evidence is required for assessment of fatigue in younger children, and children with particular health conditions.
Fatigue refers to a subjectively overwhelming sense of tiredness, lack of energy, and feeling of exhaustion. It is a common, independent, and nonspecific symptom identified in numerous chronic health conditions in childhood. Consequently, it has been described as 1 of the most universal experiences for children with chronic health conditions, although reported rates of fatigue vary within and between conditions.

Excessive fatigue leads to reduced quality of life, depression, and reduced school participation. Fatigue is also identified in healthy children with normal daily functioning. Given these characteristics, it is important to identify accurate and reliable child-friendly measures to detect excessive fatigue.

**Epidemiology**

Excess fatigue is a common presenting problem in clinical practice, occurring in children with cancer, neurologic conditions, chronic fatigue syndrome, postural tachycardia syndrome, multiple sclerosis, epilepsy, traumatic brain injury, juvenile idiopathic arthritis, diabetes, inflammatory bowel disease, fibromyalgia, and obesity. Importantly, fatigue in these conditions is distinguishable from other related symptoms including sleepiness, depression, and apathy, and is not ameliorated by sleep. Precise prevalence estimates are constrained by issues related to assessment, terminology, and measurement, which are heightened in childhood.

**Measurement**

Fatigue has been defined as an overwhelming and sustained sense of exhaustion that decreases one’s capacity for physical and mental work. Quantification of fatigue remains impeded by lack of consensus framework, terminology, and the subjective and multidimensional nature of symptoms. Three alternative frameworks for describing subjective fatigue are as follows: “central fatigue” (resulting from central as opposed to peripheral nervous system damage and incorporating mental and physical symptoms); as a “feeling state”; and as a combination of primary (disease specific) and secondary factors (such as pain, sleep disturbance, or mood changes). The subjectivity and breadth of symptoms make quantification challenging, and resultantly assessment instruments focus on the duration or severity of symptoms (physical, mental, affective) and other dimensions of the fatigue experience, including impact on daily functioning and exacerbating factors (e.g., sleep habits).

In this review, based on earlier work, the theoretical construct of fatigue will be categorized as “impact of fatigue on daily life,” “fatigue severity,” or “factors influencing fatigue.”

**Table 1: Data Extraction Items**

<table>
<thead>
<tr>
<th>Data Category</th>
<th>Description</th>
<th>Data Extracted</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Study description</td>
<td>Study aims, design, study duration, and any bias. Participant numbers, setting, diagnostic criteria, age, gender, country, comorbidity, and diagnostic criteria used.</td>
</tr>
<tr>
<td>2</td>
<td>Instrument description</td>
<td>For each study: name of instrument(s), language version, theoretical constructs, number of items, description of domains assessed, theoretical construct, age, respondents, response format, and extent of active patient involvement in instrument development.</td>
</tr>
<tr>
<td>3</td>
<td>Psychometric properties of fatigue measures</td>
<td>Criteria rating for instrument-level measurement properties (as reported in study): (1) internal consistency, (2) reliability, (3) measurement error, (4) content validity, (5) construct validity, (6) structural validity, (7) hypothesis testing, (8) cross-cultural validity, and (9) responsiveness. Rating of study methodological quality for each of the following measurement properties reported: (1) internal consistency, (2) reliability, (3) measurement error, (4) content validity, (5) construct validity, (6) structural validity, (7) hypothesis testing, (8) cross-cultural validity, (9) responsiveness.</td>
</tr>
<tr>
<td>4</td>
<td>Study quality</td>
<td>Acceptability (respondent burden, time for completion) and feasibility (time for scoring, ease of scoring).</td>
</tr>
<tr>
<td>5</td>
<td>Acceptability and feasibility</td>
<td>Acceptability (respondent burden, time for completion) and feasibility (time for scoring, ease of scoring).</td>
</tr>
<tr>
<td>6</td>
<td>Risk of bias</td>
<td>Study and outcome level data: outcome data completeness, drop-out rate description, any selective outcome reporting.</td>
</tr>
</tbody>
</table>

---

*Criteria for good measurement properties on the basis of criteria proposed by Terwee et al. The quality of each included study was evaluated by using the COSMIN checklist and rating system. Theoretical construct of fatigue rated as impact of fatigue on daily life, fatigue severity, or factors influencing fatigue. INVOLVE definitions of extent of patient involvement in instrument development (consultation, “collaboration,” or “user-led”). The quality of each included study was evaluated by using the COPE checklist and rating system.
TABLE 2 Levels of Evidence for Measurement Properties of Included Fatigue Instruments

<table>
<thead>
<tr>
<th>Level of Evidence</th>
<th>Rating</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strong</td>
<td>+++ or – – –</td>
<td>Consistent findings in multiple studies of good methodological quality or in 1 study of excellent methodological quality</td>
</tr>
<tr>
<td>Moderate</td>
<td>++ or –</td>
<td>Consistent findings in multiple studies of fair methodological quality or in 1 study of good methodological quality</td>
</tr>
<tr>
<td>Limited</td>
<td>+ or –</td>
<td>One study of fair methodological quality</td>
</tr>
<tr>
<td>Conflicting</td>
<td>+/-</td>
<td>Conflicting findings</td>
</tr>
<tr>
<td>Unknown</td>
<td>?</td>
<td>Only studies of poor methodological quality</td>
</tr>
</tbody>
</table>

*+, positive rating; ?, indeterminate rating; –, negative rating.

Exceptions are reviews of measures in children with cancer46,47 and chronic fatigue syndrome.48 It is important to identify the instruments available and to evaluate their measurement properties. The objectives of this review of fatigue in children and adolescents with chronic health conditions were to (1) describe participant and study characteristics of included studies, (2) describe and systematically review available instruments used to assess fatigue, detailing those available for child (self) and parent (proxy) report, (3) interpret instruments by using a multidimensional framework of fatigue, (4) review identified instruments within a developmental framework, and (5) summarize and synthesize the psychometric properties of the these instruments and the quality of the evidence.

METHODS

Information Sources

The search strategy was designed to retrieve literature relating to measures of subjective fatigue in children with chronic health conditions. The systematic review was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA).49 The EBSCOhost platform was used to search the electronic databases Medline, Cumulative Index to Nursing and Allied Health Literature, and PsycINFO from 1990 to August 24, 2013 (updated June 12, 2014). Nonconventional documents (gray literature) and publications, and secondary references were also checked.

Search

The searches were limited to English-language publications. Subject headings and key words used were the following: (1) fatigue, weary, weariness, exhaustion, exhausted, lacklustre or asthenia*, letharg*, tired; or (2) (lack* or loss or lost) and (energy or vigor or vigor); (3) feeling and (drained or sleepy or sluggish or weak*); (4) central nervous system diseases, neoplasms, multiple sclerosis, congenital heart defects, anemia, chronic pain, chronic disease, cerebral palsy, cardiovascular diseases, diabetes mellitus (type 1); chronic fatigue syndrome; and (5) questionnaires or psychological tests.

Eligibility Criteria

Population

Children and/or adolescents (≤18 years of age) with a health condition.

Assessment Tool

Any questionnaire designed to measure fatigue by using more than 1 item. Fatigue subscales included in quality-of-life or other measures were eligible for inclusion, provided that fatigue scores could be extracted separately (ie, total subscale scores provided). Fatigue assessment must be subjective, quantified, and based on child report or parent/other report, to be eligible for inclusion.

Outcome

Any clinical outcome.

Study Design

Interventional and observational studies published in peer-reviewed journals were eligible for inclusion.

Exclusion Criteria

Review articles, opinion pieces, and single case studies were excluded as follows: studies that only reported qualitative information as a measure of fatigue (ie, professional opinion or interview data); fatigue measures that used a single item, either stand alone (single item visual analog scales [VASs]) or as part of another questionnaire (ie, quality of life/depression/symptom checklists); and objective measures of performance decrement (ie, attention vigilance, muscle fatigue).

Study Selection

Titles and abstracts of all articles were assessed for inclusion by 2 independent reviewers (Drs Crichton and Knight) and agreement checked.

Data Collection Process

A standardized data extraction form was developed and informed by standard processes for systematic reviews,51,52 earlier reviews in the field,37,46–48 and incorporating recommendations for patient-reported outcome measure evaluation and standards for appraisal of methodological quality.53–55 The form was independently piloted in 2 randomly selected studies by Drs Crichton and Knight. Data were extracted by Dr Crichton, independently checked by Dr Knight, and disagreements were resolved by consensus with consultation of a third reviewer (Dr Anderson; Table 1).

The data extraction and assessment included information from 6 domains (Table 1). The measurement properties evaluated in data categories 3 and 4 in Table 1 incorporate a consensus-based taxonomy of measurement properties considered important in patient-reported outcome measures, defined...
by Mokkink et al\textsuperscript{56} (Table 2, p. 743). Relevant properties include the following: (1) internal consistency, (2) reliability, (3) measurement error, (4) content validity, (5) construct validity, (6) structural validity, (7) hypothesis testing, (8) cross-cultural validity, and (9) responsiveness. Criterion validity is not relevant for fatigue assessment, so it was excluded. The 9 listed measurement properties provided the domains for data extraction for 2 key data sets at an individual study level: (1) rating the measurement property of the instrument; (2) rating the quality of the study (data categories 3 and 4, Table 1; floor/ceiling effects and interpretability not rated). According to predefined criteria, each measurement property was rated as "positive," "indeterminate," "negative," or "no information available."\textsuperscript{55} If measurement properties were not checked, but the authors cited published measurement properties in another article, this was recorded and the above ratings were maintained.

The methodological quality of included studies were evaluated by using the consensus-based standards for the selection of health measurement instruments (COSMIN) checklist and 4-point rating scale.\textsuperscript{57} The COSMIN checklist contains 114 data points (12 boxes, 9 assessing measurement properties), relating to design requirements and preferred statistical methods and consensus on statistical values considered adequate. Of these, 91 data points (8 boxes) were relevant to the current review: internal consistency (11), reliability (14), measurement error (11), content validity (5), structural validity (7), hypothesis testing (10), cross-cultural validity (15), and responsiveness (18). Each item was scored on a 4-point rating scale (ie, "poor," "fair," "good," or "excellent," and "not applicable" if omitted). For each measurement property, the final rating was determined by the lowest rating of any of the items within the box. If more than 1 instrument was included in a study (eg, parent and child versions), the same measurement property was assessed for each instrument.

**Data Synthesis**

Study level data were synthesized to provide summary information for each fatigue instrument. Information across studies was summarized to describe identified instruments and the populations to whom instruments were administered. Second, as guidelines for pooling of measurement properties are still under development, best evidence synthesis was performed integrating: (1) the rating for the instrument’s measurement property across studies\textsuperscript{55} and (2) the COSMIN rating for the methodological quality of the studies.\textsuperscript{57} For each fatigue
instrument, the consistency across studies was reviewed for each measurement property. The level of evidence for the measurement properties of instruments across studies was rated by using predefined criteria of positive, indeterminate, or negative, accompanied by levels of evidence, as proposed by the Cochrane Back Review Group (Table 1), detailed in Table 2.

RESULTS

Figure 1 illustrates the flow of article selection. A total of 89 studies measuring fatigue by using 20 instruments were included in data extraction and synthesis. One study was not included in data synthesis for measurement properties because insufficient information was available for data extraction.

Patient and Study Characteristics

Most reviewed studies were cross-sectional (55; 62%), with a median sample size of 69 subjects (6–3890). Instruments were used to assess fatigue in children with a number of health conditions (grouped into 14 diagnostic categories). The most frequently studied patient groups were hematologic/oncology diagnoses (37; 42%) and chronic fatigue syndrome (17; 19%; Fig 2).

Identified Instruments

Twenty fatigue instruments were identified (Fig 3). The Pediatric Quality of Life Inventory (PedsQL) Multidimensional Fatigue Scale (MFS) parent and child report was the most commonly used, cited in 26 (29%) and 20 (22%) studies, respectively. The fatigue scales (Fatigue Scale-Child [FS-C]; Fatigue Scale-Adolescent [FS-A], and the Fatigue Scale-Parent [FS-P]) were the second most common instruments to be cited in 18 (20%), 16 (18%), and 12 (13%) studies, respectively. The injury scales (Fatigue Scale-Child [FS-C]; Fatigue Scale-Adolescent [FS-A], and the Fatigue Scale-Parent [FS-P]) were the second most common instruments to be cited in 18 (20%), 16 (18%), and 12 (13%) studies, respectively.

![Figure 2](http://pediatrics.aappublications.org/)

Child health conditions of patient group in included studies, n (%).

![Figure 3](http://pediatrics.aappublications.org/)

Fatigue instruments used across identified studies, n (%).

---

**TABLE 1**

<table>
<thead>
<tr>
<th>Health condition</th>
<th>Number of studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Haematology/oncology</td>
<td>37 (42)</td>
</tr>
<tr>
<td>Chronic fatigue syndrome</td>
<td>17 (19)</td>
</tr>
<tr>
<td>Neurological disease</td>
<td>9 (11)</td>
</tr>
<tr>
<td>Other chronic health</td>
<td>7 (8)</td>
</tr>
<tr>
<td>Rheumatological disease</td>
<td>5 (4)</td>
</tr>
<tr>
<td>Musculoskeletal</td>
<td>3 (3)</td>
</tr>
<tr>
<td>Obesity</td>
<td>2 (2)</td>
</tr>
<tr>
<td>Renal disease</td>
<td>2 (2)</td>
</tr>
<tr>
<td>Diabetes, type 1</td>
<td>2 (2)</td>
</tr>
<tr>
<td>Neuromotor disease</td>
<td>2 (2)</td>
</tr>
<tr>
<td>Brain injury</td>
<td>2 (2)</td>
</tr>
<tr>
<td>Gastroenterology</td>
<td>1 (1)</td>
</tr>
<tr>
<td>Chronic pain</td>
<td>1 (1)</td>
</tr>
<tr>
<td>Heart diseases (congenital)</td>
<td>1 (1)</td>
</tr>
</tbody>
</table>

**TABLE 2**

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Number of studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>PedsQL MFS (parent)</td>
<td>26 (29)</td>
</tr>
<tr>
<td>PedsQL MFS (child)</td>
<td>20 (22)</td>
</tr>
<tr>
<td>FS-C PedsQL (parent)</td>
<td>16 (18)</td>
</tr>
<tr>
<td>FS-A PedsQL (parent)</td>
<td>12 (13)</td>
</tr>
<tr>
<td>FS-P PedsQL (parent)</td>
<td>8 (9)</td>
</tr>
<tr>
<td>FS-C PedsQL (child)</td>
<td>5 (6)</td>
</tr>
<tr>
<td>FS-A PedsQL (child)</td>
<td>4 (4)</td>
</tr>
<tr>
<td>FS-P PedsQL (child)</td>
<td>4 (4)</td>
</tr>
<tr>
<td>FS-C PedsQL (adolescent)</td>
<td>2 (2)</td>
</tr>
<tr>
<td>FS-A PedsQL (adolescent)</td>
<td>2 (2)</td>
</tr>
<tr>
<td>FS-P PedsQL (adolescent)</td>
<td>2 (2)</td>
</tr>
<tr>
<td>FS-C PedsQL (parent) PedsQL (child)</td>
<td>2 (2)</td>
</tr>
<tr>
<td>FS-A PedsQL (parent) PedsQL (child)</td>
<td>1 (1)</td>
</tr>
</tbody>
</table>

---

**FIGURE 2**

Child health conditions of patient group in included studies, n (%).

**FIGURE 3**

Fatigue instruments used across identified studies, n (%).
<table>
<thead>
<tr>
<th>Measure</th>
<th>Construct Assessed</th>
<th>Other Constructs</th>
<th>Participant Age Range</th>
<th>Healthy Controls Age Range (n = studies)</th>
<th>Recall Period</th>
<th>Dimension (n = items, range)</th>
<th>Response Options</th>
<th>Health Conditions Assessed</th>
<th>Administration Burden</th>
<th>Time to Complete</th>
<th>Ease of Scoring</th>
</tr>
</thead>
<tbody>
<tr>
<td>PedsQL MFS (child)</td>
<td>Severity of fatigue; impact of fatigue</td>
<td>Sleep/tiredness; cognitive function</td>
<td>5–7, 8–18</td>
<td>5–18 (3)</td>
<td>1 mo</td>
<td>General fatigue (6); sleep/rest fatigue (6); cognitive fatigue (6); total fatigue (18, 0–100)</td>
<td>5-point Likert (ages 8–18); 4-point Likert (ages 5–7)</td>
<td>Hematology/ oncology; neurologic disease; rheumatologic disease; obesity; chronic pain; brain injury; heart diseases (congenital); diabetes, type 1; gastroenterology</td>
<td>Administrator required to read and complete child report 5–7 y</td>
<td>5 min</td>
<td>Difficult</td>
</tr>
<tr>
<td>FS-C</td>
<td>Severity of fatigue; impact of fatigue</td>
<td>Cognitive and physical weariness; added effort and assistance needed to do usual activities; needing rest and feeling angry; avoiding social activities</td>
<td>7–12</td>
<td>7–12 (2)</td>
<td>24 h or 1 wk or 4 wk</td>
<td>Intensity (14; 0–70); frequency (14; 0–14); total fatigue (10 or 14; 0–70)</td>
<td>5-point Likert yes/no, then 4-point Likert</td>
<td>Hematology/ oncology</td>
<td>Able to be completed by 8-y-olds</td>
<td>3–10 min</td>
<td>Easy</td>
</tr>
<tr>
<td>FS-A</td>
<td>Severity of fatigue; impact of fatigue</td>
<td>Sleep/tiredness; mood; cognitive function; physical activity; school performance; withdrawal</td>
<td>13–18</td>
<td>13–18 (1)</td>
<td>24 h or 1 wk or 4 wk</td>
<td>Intensity (14; 0–56 or 14–70)</td>
<td>5-point or 4-point Likert</td>
<td>Hematology/ oncology</td>
<td>Not specified</td>
<td>2–4 min; 4–5 min</td>
<td>Easy</td>
</tr>
<tr>
<td>CIS-20</td>
<td>Severity of fatigue; impact of fatigue</td>
<td>Subjective experience of fatigue; concentration; motivation; physical activity</td>
<td>10–18</td>
<td>10–17 (5)</td>
<td>2 wk</td>
<td>Subjective fatigue (8); concentration (5); motivation (4); physical activity (3); total (20; 20–140)</td>
<td>7-point Likert</td>
<td>Neurologic/CFS</td>
<td>Not specified</td>
<td>Not specified</td>
<td>Moderate</td>
</tr>
<tr>
<td>Chalder Fatigue Scale</td>
<td>Severity of fatigue</td>
<td>Physical; mental fatigue</td>
<td>5–19</td>
<td>10–19 (5)</td>
<td>2 wk</td>
<td>Physical (7); mental (4); total fatigue (11 or 14)</td>
<td>3- or 4-item Likert</td>
<td>Hematology/ oncology; CFS</td>
<td>Not specified</td>
<td>Not specified</td>
<td>Easy</td>
</tr>
<tr>
<td>FSS</td>
<td>Severity of fatigue; impact of fatigue</td>
<td>Not specified</td>
<td>8–188</td>
<td>8–18 (2)</td>
<td>1 wk</td>
<td>Total fatigue (9; 1–7 average)</td>
<td>7-point Likert</td>
<td>Hematology/ oncology; neurologic disease</td>
<td>Not specified</td>
<td>Not specified</td>
<td>Easy</td>
</tr>
<tr>
<td>PedsQL 3.0 CP Module (child)</td>
<td>Not specified</td>
<td>Not specified</td>
<td>5–18</td>
<td>5–18 (1)</td>
<td>1 mo</td>
<td>Total fatigue (4; 0–100) as subdomain total of scale with 7 domains</td>
<td>5-point Likert (ages 8–18); 4-point Likert (ages 5–7)</td>
<td>Neurologic Motor</td>
<td>Administrator required to read and complete child report 5–7 y</td>
<td>30 min including consent</td>
<td>Difficult</td>
</tr>
</tbody>
</table>
## TABLE 3 Continued

<table>
<thead>
<tr>
<th>Measure</th>
<th>Construct Assessed</th>
<th>Other Constructs</th>
<th>Participant Age Range (n = studies)</th>
<th>Recalls Period</th>
<th>Dimension (n = items, range)</th>
<th>Response Options</th>
<th>Health Conditions Assessed</th>
<th>Administration Burden</th>
<th>Time to Complete</th>
<th>Ease of Scoring</th>
</tr>
</thead>
<tbody>
<tr>
<td>VAS</td>
<td>Severity of fatigue</td>
<td>Not specified</td>
<td>8–18</td>
<td>None</td>
<td>Present state</td>
<td>Total fatigue (1 or 6; 0–4 or 0–100 mm)</td>
<td>Visual analog</td>
<td>Hematology/oncology, musculoskeletal disease</td>
<td>Not specified</td>
<td>Not specified</td>
</tr>
<tr>
<td>POMS</td>
<td>Severity of fatigue</td>
<td>Not specified</td>
<td>12–17</td>
<td>None</td>
<td>1 wk or 4 wk</td>
<td>Total fatigue (5 or 15; not specified) as subdomain total of scale with 6 domains</td>
<td>5-point Likert</td>
<td>Hematology/oncology, neurologic disease</td>
<td>Not specified</td>
<td>Not specified</td>
</tr>
<tr>
<td>PROMIS (child)</td>
<td>Severity of fatigue; impact of fatigue</td>
<td>Tiredness, lack of energy, impact on activities</td>
<td>8–17 (1)</td>
<td>1 wk</td>
<td>Fatigue (37, factor array --2 to +2) as subdomain total of scale with 6 domains; fatigue item bank: tiredness (39); energy (14); total (39; mean of 50 and SD of 10)</td>
<td>5-point Likert</td>
<td>Obesity, diabetes, type 1; other chronic health</td>
<td>Computer literacy, computer, screen, and mouse access</td>
<td>25 min younger children, 15 or 32 min adolescents</td>
<td>Difficult</td>
</tr>
<tr>
<td>Q-sort task</td>
<td>Severity of fatigue; impact of fatigue</td>
<td>Energy and related capacity for physical functioning; fatigue related to psychosocial functioning; fatigue uniquely associated with anemia and/or treatment</td>
<td>12–18</td>
<td>None</td>
<td>1 wk</td>
<td>Total fatigue (57, factor array --4 to +4)</td>
<td>Rank order --4 to +4</td>
<td>Hematology/oncology</td>
<td>Not specified</td>
<td>Not specified</td>
</tr>
<tr>
<td>PedsQL 3.0 ESRD Module (child)</td>
<td>Severity of fatigue; impact of fatigue</td>
<td>Not specified</td>
<td>5–18 (1)</td>
<td>1 mo</td>
<td>General fatigue (4; 0–100)</td>
<td>5-point Likert</td>
<td>Renal disease</td>
<td>Not specified</td>
<td>Not specified</td>
<td>Difficult</td>
</tr>
</tbody>
</table>

CAT, computerized adaptive testing; CFS, chronic fatigue syndrome.

a Based on earlier work by Elbers et al,37 the theoretical construct of fatigue was categorized as impact of fatigue on daily life, fatigue severity, or factors influencing fatigue.

b Goretti et al4 included children younger than 17 y, 11 mo. Age range was not specified.
<table>
<thead>
<tr>
<th>Measure</th>
<th>Construct Assessed</th>
<th>Other Constructs</th>
<th>Participant Age Range</th>
<th>Healthy Controls Age Range ( (n = \text{studies}) )</th>
<th>Recall Period</th>
<th>Dimension ( (n = \text{items}, \text{Range}) )</th>
<th>Response Options</th>
<th>Health Conditions Assessed</th>
<th>Administration Burden</th>
<th>Time to Complete</th>
<th>Ease of Scoring</th>
</tr>
</thead>
<tbody>
<tr>
<td>PedsQL MFS (parent)</td>
<td>Severity of fatigue; impact of fatigue</td>
<td>Sleep and tiredness; cognitive function</td>
<td>2–18</td>
<td>2–18 (11)</td>
<td>1 mo</td>
<td>General fatigue (6); sleep/rest fatigue (8); cognitive fatigue (6) (total = 18; 0–100)</td>
<td>5-point Likert</td>
<td>Hematology/oncology; neurologic disease; rheumatological disease; obesity; chronic pain; brain injury; heart diseases (congenital); diabetes, type 1; gastroenterology</td>
<td>Not specified</td>
<td>5 min</td>
<td>Difficult</td>
</tr>
<tr>
<td>FS-P</td>
<td>Severity of fatigue; impact of fatigue</td>
<td>Cognitive and physical weariness; effort and assistance needed to do usual activities; needing rest and feeling angry; avoiding social activities</td>
<td>7–18</td>
<td>7–18 (2)</td>
<td>24 h or 1 wk</td>
<td>Intensity (14, 17, or 18 items; 14–17 or 18–90); total (18, 17–65 or 0–86)</td>
<td>5-point Likert</td>
<td>Hematology/oncology</td>
<td>Not specified</td>
<td>Not specified</td>
<td>Easy</td>
</tr>
<tr>
<td>PedsQL 3.0 CP Module (parent)</td>
<td>Not specified</td>
<td>Not specified</td>
<td>2–18</td>
<td>2–18 (1)</td>
<td>1 mo</td>
<td>Total fatigue (4; 0–100) as subdomain total of scale with 7 domains</td>
<td>5-point Likert</td>
<td>Neumotor</td>
<td>Not specified</td>
<td>30 min including consent</td>
<td>Difficult</td>
</tr>
<tr>
<td>PROMIS (parent)</td>
<td>Severity of fatigue; impact of fatigue</td>
<td>Tiredness; lack of energy; impact on activities</td>
<td>5–18</td>
<td>1 wk</td>
<td>5-point Likert or 4-point Likert (physical health)</td>
<td>Lack of energy (11); tiredness (23); total fatigue (34) (mean of 50, SD of 10)</td>
<td>Other chronic health</td>
<td></td>
<td>Computer literacy; computer screen; mouse access</td>
<td>12–22 min (PROMIS I); 30–40 min (PROMIS II)</td>
<td>Difficult</td>
</tr>
<tr>
<td>EOSQ</td>
<td>Severity of fatigue</td>
<td>Lack of energy</td>
<td>1–12</td>
<td>4 wk</td>
<td>5-point Likert</td>
<td>Total fatigue (2; not specified) subdomain total of scale with 13 domains</td>
<td>Musculoskeletal</td>
<td>Not specified</td>
<td>Not specified</td>
<td>Difficult</td>
<td></td>
</tr>
<tr>
<td>FACIT-F (modified parent)</td>
<td>Severity of fatigue; impact of fatigue</td>
<td>Not specified</td>
<td>2–18</td>
<td>1 wk</td>
<td>5-point Likert</td>
<td>Total fatigue (15; 0–52) as subdomain total of scale with 5 domains</td>
<td>Hematology/oncology</td>
<td>Not specified</td>
<td>Not specified</td>
<td>Easy</td>
<td></td>
</tr>
<tr>
<td>PedsQL ESRD (parent)</td>
<td>Severity of fatigue; impact of fatigue</td>
<td>Not specified</td>
<td>2–18</td>
<td>1 mo</td>
<td>5-point Likert</td>
<td>General fatigue (4; 0–100)</td>
<td>Renal disease</td>
<td>Not specified</td>
<td>Not specified</td>
<td>Difficult</td>
<td></td>
</tr>
<tr>
<td>FS-S</td>
<td>Severity of fatigue; impact of fatigue</td>
<td>Not specified</td>
<td>7–18</td>
<td>7–18 (1)</td>
<td>24 h or 1 wk</td>
<td>Fatigue intensity (9, 10 or 18; 9–36)</td>
<td>4-point Likert</td>
<td>Hematology/oncology</td>
<td>Not specified</td>
<td>Not specified</td>
<td>Easy</td>
</tr>
</tbody>
</table>

* Based on earlier work by Elbers et al., the theoretical construct of fatigue was categorized as impact of fatigue on daily life, fatigue severity, or factors influencing fatigue.
unique in using staff-rated instruments (Fig 3).

**Dimensions of Fatigue Assessed**

Most instruments assessed impact of fatigue on daily life and fatigue severity (14; 70%). Few (4; 20%) assessed fatigue severity symptoms only, and no instruments assessed impact of fatigue on daily life in isolation. Although item content related to these dimensions, scoring rarely reflected the dimensionality of the constructs assessed, and items for most instruments (14; 70%) were summed into a total fatigue score. A minority (6; 30%) consisted of scores for multiple fatigue dimensions: 5 child-report measures summed items into 2 to 4 dimensions, and 1 parent report measure summed items into fatigue dimensions (3). There was insufficient information on item content for 2 instruments (10%).

**Developmental Framework**

Overall, participant age was skewed toward children 12 years and older (Fig 4). Healthy control participants were included in 23 of 81 studies for which age range could be extracted (29%). Self-report of fatigue is challenging in younger children, and only 4 (33%) child report instruments were identified in use between 5 and 7 years of age. Greater administration burden was commonly noted because items need to be read aloud to children or administrators need to be available to note children's verbal responses. In contrast, 4 (57%) parent report instruments were identified to assess fatigue symptoms in children as young as 2 years of age.

**Synthesis of Results**

**Psychometric Properties**

A synthesis summary of the investigated methodological properties of measures and methodological quality of the included studies is provided in Table 6. Full details of the resultant COSMIN ratings for each study (per measurement property and questionnaire) are provided in Supplemental Tables 7–26. The quality of evidence available across questionnaires varied by measurement properties as shown, and no single instrument demonstrated positive findings across all properties. Content validity was rated for all instruments and revealed the most promising evidence. Although assessed in several instruments, poor results were found for internal consistency. This was primarily because studies did not test for unidimensionality (factor analysis), or because despite completing factor analysis and evidence of more than 1 factor, items were summarized into a single score. Cultural validity was not considered, and COSMIN items were used to rate the quality of the translation across all studies with cultural validity data. Few studies revealed information about measurement error, reliability, or responsiveness. There was limited evidence of the floor and ceiling effects across studies, with <5% of included studies providing information.

There was robust evidence for the measurement properties of the PedsQL MFS and the FS-C/FS-P. The PedsQL MFS (child and parent versions) was the only included instrument to have strong evidence of reliability, and the only instrument to have strong evidence across a number of dimensions of assessment. There were positive ratings for reliability, content validity, and hypothesis testing for the PedsQL MFS. Despite the overall rating for reliability, internal consistency coefficient (ICC) values varied across different populations assessed from low to high (eg, 0.27–0.93 for total fatigue).23,125 Additionally, variability with age has been noted.64 There was limited evidence for the structural validity of the child version fitting a bifactor model (3 subscales: general fatigue, sleep–rest fatigue, and cognitive fatigue and a total fatigue score),122 lending new support to previous research finding good to excellent internal consistency statistics (Cronbach α values between 0.70 and 0.95). Strong evidence for hypothesis testing validity for the PedsQL MFS was evident, although internal consistency and measurement error of the PedsQL MFS require further research. The data for these measurement properties were gathered in children with numerous health conditions, noted in Tables 3 and 4.

The FS-C, FS-A, and FS-P (standard and 24-hour formats) demonstrated robust measurement properties, notably content validity. There was moderate evidence for the structural validity of the child and parent forms, but limited evidence for the structural validity of the adolescent form. Internal consistency and reliability data were mixed; although internal consistency statistics were good (0.70 or greater),79–81,107,113 these statistics were calculated based on the total fatigue scores, despite findings of a 3-factor model for the FS-C and a 4-factor model for the FS-A.79,80 One study revealed adequate evidence for internal consistency and content validity and limited evidence for structural validity.38,60,75,79 The Fatigue Scale-Staff (FS-S) demonstrated moderate evidence for internal consistency and content validity and limited evidence for structural validity.38,60,75,79 The assessed evidence for FS-C, FS-A, and FS-P was gathered exclusively in studies of children with hematology/oncology diagnoses, limiting generalizability of findings to other health conditions.

The Patient-Reported Outcomes Measurement Information System
(PROMIS) is a newly developed outcome measure (fatigue subscale). Among identified studies, those that used PROMIS were the only ones to employ item response theory for statistical analysis. More published information was available for the child version, with positive ratings evident for content validity and hypothesis testing. However, there were also some positive findings for the content and validity of the PROMIS parent report. Overall, properties were given largely indeterminate ratings because the measurement properties have not yet been evaluated. Evidence was gathered in study populations of children with mixed health conditions.

There was limited evidence for hypothesis testing validity of the Checklist Individual Strength-20 (CIS-20) in children with neurologic disorders and chronic fatigue syndrome. Overall findings for the scale were equivocal.

Some disease-specific instruments demonstrated positive ratings for content validity in these specific populations (eg, the PedsQL 3.0 Cerebral Palsy Module [PedsQL 3.0 CP Module] and the Early Onset

---

**TABLE 6** Levels of Evidence for the Overall Quality of the Measurement Property per Instrument

<table>
<thead>
<tr>
<th>Instrument Name</th>
<th>Internal Consistency</th>
<th>Reliability</th>
<th>Measurement Error</th>
<th>Content Validity</th>
<th>Structural Validity</th>
<th>Hypothesis Testing</th>
<th>Cross Cultural Validity</th>
<th>Responsiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td>PedsQL MFS (parent)</td>
<td>?</td>
<td>+++</td>
<td>?</td>
<td>+++</td>
<td>?</td>
<td>+++</td>
<td>+</td>
<td>?</td>
</tr>
<tr>
<td>PedsQL MFS (child)</td>
<td>?</td>
<td>+++</td>
<td>?</td>
<td>+++</td>
<td>+</td>
<td>+++</td>
<td>+</td>
<td>?</td>
</tr>
<tr>
<td>FS-C or FS-C 24 h</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>+++</td>
<td>++</td>
<td>+</td>
<td>+</td>
<td>?</td>
</tr>
<tr>
<td>FS-A or FS-A 24 h</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>+++</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>FS-P or FS-P 24 h</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>+++</td>
<td>++</td>
<td>+</td>
<td>+</td>
<td>?</td>
</tr>
<tr>
<td>FS-S</td>
<td>++</td>
<td>?</td>
<td>?</td>
<td>++</td>
<td>+</td>
<td>?</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>PROMIS (child)</td>
<td>?</td>
<td>+</td>
<td>?</td>
<td>+++</td>
<td>?</td>
<td>++</td>
<td>?</td>
<td>?</td>
</tr>
</tbody>
</table>

*a Strong +++ or — — —, Consistent findings in multiple studies of good methodological quality or in 1 study of excellent methodological quality. Moderate ++ or — —, Consistent findings in multiple studies of fair methodological quality or in 1 study of good methodological quality. Limited + or —, 1 study of fair methodological quality. Conflicting +/-, Conflicting findings; Unknown ? ?, Only studies of poor methodological quality.
Scoliosis Questionnaire [EOSQ]), although lacked other evidence for the quality of their measurement properties. However, the majority of included instruments demonstrated low ratings overall. There was limited evidence for the content validity of the Pediatric Functional Assessment of Chronic Illness Therapy-Fatigue (PACT-F) and the Q-sort instruments for use in hematology/oncology populations, despite no published evidence for other assessed measurement properties. There was no positive evidence for measurement properties of the Chronic Illness Therapy-Fatigue Syndrome, or the Fatigue Scale, or the Pediatric Functional Assessment of Chronic Illness Therapy-Fatigue (PACT-F) and the Q-sort instruments for use in hematology/oncology populations, despite no published evidence for other assessed measurement properties.

Risk of Bias

Risk of bias was identified across studies for several instruments, due to reanalysis of participant data across studies, including PedsQL MFS, the Fatigue Scales (child, parent, and adolescent), and the CIS-20. Repeated analysis of group data possibly skews findings and limits the generalizability of findings to other populations. Further, although declared, a primary researcher holds the copyright for the PedsQL including PedsQL MFS, and the PedsQL ESRD Module.

DISCUSSION

This study is the first to systematically review and critically appraise fatigue instruments across the range of childhood chronic health conditions and ages. Twenty instruments (12 child reports, 7 parent reports, and 1 staff report) were described with reference to patient characteristics (age and diagnosis), dimensionality of fatigue assessed, and child development. This review provides a thorough and unique review of childhood fatigue instruments across conditions, using standardized published measures to assess the measurement properties of these instruments and to assess the quality of the evidence for these properties.

In applying COSMIN criteria and assessing the quality of the study, evidence for the measurement properties of the 20 instruments can be compared. This review did not identify any child or parent report instrument that met all of the criteria. This appraisal highlights key issues in the assessment of fatigue in childhood health conditions. It is widely acknowledged that as a subjective symptom, fatigue should be assessed by using self report, although developmental considerations necessarily restrict the feasibility of this in younger children. However, it is important to recognize the noted lack of robust evidence for reliability. Where ICC values were recorded, they were often low, indicating lack of agreement between child and parent report on assessment instruments, meaning child and parent report instruments cannot be directly compared. It is recommended that parent and child perspectives be carefully considered when interpreting findings relating to fatigue symptoms in children.

Over 14 health conditions were identified as experiencing symptoms of fatigue; however, evidence for instruments varied across these conditions. Findings from this review suggest that evidence is limited to instruments that assess fatigue in children with hematology/oncology, and other health conditions. Specifically, there was support for the measurement properties of the PedsQL MFS, and limited evidence for the PROMIS instrument in children with a range of chronic health conditions, and for the FS-C, FS-P, and the FS-A in children with cancer. Evidence for instrument properties in children with chronic fatigue syndrome was lacking, despite the high proportion of studies with this population. Further, although literature suggests the low threshold of fatigue symptoms distinguishes this clinical condition, no reviewed instruments assessed this. Practicing clinicians should be aware that fatigue can arise from many causes and conditions, and awareness of the underlying disorder and relevant clinical features is critical to selecting the appropriate valid assessment tool.

Whether fatigue should be conceptualized as unidimensional or multidimensional has been a matter of considerable debate, and evidently measurement of fatigue is a complex issue. Descriptive findings suggest that this issue is poorly managed by currently available fatigue instruments. Notably, studies failed to confirm the number of fatigue dimensions (by factor analysis or other methods), or assessment of fatigue was conducted by using instruments in which the number of scales did not correspond to the number of dimensions identified in factor analysis. For example, of the questionnaires with the most robust psychometric properties and study quality, scores are summed to 3 dimensions for the PedsQL MFS, and a total score for the fatigue scales (FS-C, FS-A, and FS-P). However, there is evidence from factor analysis that both instruments are explained by 3 to 4 factors. Studies beyond the scope of this review further suggest multiple factors in fatigue assessment instruments, with recent evidence confirming a 3-factor structure of the PedsQL MFS.
a 2-factor structure of the PROMIS scale.\textsuperscript{2} The field of research is yet to develop a coherent way of relating this emerging evidence regarding fatigue dimensionality with quantification of symptoms. In clinical practice, it is important to recognize the complexity of fatigue, and to move beyond simplistic assessment of the symptom to consider a multidimensional approach to assessment. Clinicians should be aware of the different information provided by the range of assessment instruments.

This review considered child-specific issues relating to fatigue assessment. However, the findings for reviewed fatigue instruments cannot be generalized across age ranges because participants were mainly older than 12 years of age, and gaps in the quality of evidence were identified at key developmental stages. It is important to identify the ages in which studies have been performed because measurement properties may vary in different age ranges. Few instruments are validated for use in younger children, with robust evidence provided only by the PedsQL MFS (age range, 2–18 years), and limited evidence provided by the PedsQL 3.0 CP Module and the EOSQ. There was limited reference to healthy control participants in included studies. It is crucial to the determination of excessive fatigue that normal developmental differences are taken into account because fatigue is a symptom that occurs in healthy populations, and reveals developmental variation.\textsuperscript{85,128,135}

Further studies are required to examine fatigue expression across age groups, including notably younger children, and make reference to normal development, to understand how age affects the experience of fatigue.

**LIMITATIONS**

Limitations of this review include a focus on children 18 years or younger with different health conditions. Evidence of measurement properties in use in other populations was not considered. For example, despite several of these measures being described in older or healthy childhood populations, the inclusion of these findings was beyond the scope of this review. Care should be taken when considering the generalizability of the findings to other patient groups. Further, the review was restricted to subjective measures of fatigue and did not include studies that have attempted to measure fatigue objectively.

**CONCLUSIONS**

There are a large number of instruments available to measure fatigue across a range of childhood health conditions, yet there is limited evidence for their utility. There was best evidence for the use of the PedsQL MFS (parent and child report) in a range of pediatric chronic health conditions across childhood, and the FS-C, FS-A, and FS-P in children with cancer of school age. Use of other instruments should be limited to research into the validity of the instrument, not patient outcomes. Further data are required to support the measurement properties of instruments used to assess fatigue in children with chronic fatigue syndrome, to identify appropriate tools to assess fatigue in children younger than school age, and to clarify the nature of agreement between child and proxy report across developmental stages.

**ACKNOWLEDGMENT**

We thank Poh Chua, Deputy Librarian, RCH, Melbourne, for her help in use of the EBSCOhost platform and search terms.

---

**REFERENCES**


and psychosocial features of childhood and juvenile MS. Neurology. 2008;70(20):1891–1897


69. Gibson F, Garnett M, Richardson A, Edwards J, Sepion B. Heavy to carry: a survey of parents’ and healthcare professionals’ perceptions of...


129. Yeh CH, Wang CH, Chiang YC, Lin L, Chien LC. Assessment of symptoms reported by 10- to 18-year-old cancer patients in...


