Agreement Between Child Self-report and Caregiver-Proxy Report for Symptoms and Functioning of Children Undergoing Cancer Treatment

Jennifer W. Mack, MD, MPH; Molly McFatrich, MPH; Janice S. Withycombe, PhD, RN, MN; Scott H. Maurer, MD; Shana S. Jacobs, MD; Li Lin, MS; Nicole R. Lucas, BS; Justin N. Baker, MD; Courtney M. Mann, MA, CCRP; Lillian Sung, MD, PhD; Deborah Tomlinson, RN, MN; Pamela S. Hinds, PhD, RN; Bryce B. Reeve, PhD

**IMPORTANCE** Adult patients are considered the best reporters of their own health-related quality of life (HRQOL). Self-report in pediatrics has been challenged by a limited array of valid measures. Caregiver report is therefore often used as a proxy for child report.

**OBJECTIVES** To examine the degree of alignment between child and caregiver proxy report for Patient-Reported Outcomes Measurement Information System (PROMIS) HRQOL domains among children with cancer and to identify factors associated with better child and caregiver-proxy congruence.

**DESIGN, SETTING, AND PARTICIPANTS** In this multicenter cohort study, children with a first cancer diagnosis and their caregivers completed surveys at 2 time points: within 72 hours preceding treatment initiation (T1) and during follow-up (T2), when symptom burden was expected to be higher (eg, 7-17 days later for chemotherapy). Data were collected from October 26, 2016, to October 5, 2018, at 9 pediatric oncology hospitals. Five hundred eighty children (aged 7-18 years) and their caregivers were approached; 482 child-caregiver dyads completed surveys at T1 (response rate 83%), and 403 completed surveys at T2 (84% of T1 participants). Data were analyzed from July 1, 2019, to April 22, 2020.

**EXPOSURES** Participants received up-front cancer treatment, including chemotherapy and radiotherapy.

**MAIN OUTCOMES AND MEASURES** Congruence between child self-report and caregiver-proxy report of PROMIS pediatric domains of mobility (physical functioning), pain interference, fatigue, depressive symptoms, anxiety, and psychological stress.

**RESULTS** Of the 482 dyads included in the analysis, 262 children (54%) were male (mean [SD] age, 12.9 [3.4] years), 80 (17%) were Black, and 71 (15%) were Hispanic. Intraclass correlations between child self-report and caregiver proxy report showed moderate agreement for mobility (0.57 [95% CI, 0.50-0.63]) and poor agreement for symptoms (range, 0.32 [95% CI, 0.24-0.41] for fatigue to 0.42 [95% CI, 0.34-0.50] for psychological stress). Children reported lower symptom burden and higher mobility than caregivers reported. In a multivariable model adjusted for child and parent sociodemographic factors and the caregiver’s own self-reported HRQOL, caregivers reported the child’s mobility score 6.00 points worse than the child’s self-report at T2 (95% CI, −7.45 to −4.51), exceeding the PROMIS minimally important difference of 3 points. Caregivers overestimated the child’s self-reported symptom levels, ranging from 5.79 (95% CI, 3.99-7.60) points for psychological stress to 13.69 (95% CI, 11.60-15.78) points for fatigue. The caregiver’s own self-reported HRQOL was associated with the magnitude of difference between child and caregiver scores for all domains except mobility; for example, for fatigue, the magnitude of difference between child and caregiver-proxy scores increased by 0.21 (95% CI, 0.13-0.30) points for each 1-point increase in the caregiver’s own fatigue score.

**CONCLUSIONS AND RELEVANCE** This study found that caregivers consistently overestimated symptoms and underestimated mobility relative to the children themselves. These results suggest that elicitation of the child’s own report should be pursued whenever possible.

Published online August 24, 2020.

© 2020 American Medical Association. All rights reserved.
Patients are increasingly recognized as the best reporters of their own experiences, including symptoms related to cancer and its treatment. As a result, adult patient reports are being integrated into symptom assessment and adverse event reporting across the cancer care continuum. Integration of the patient voice has been more challenging in pediatric oncology owing to a limited array of developmentally appropriate, valid tools for child self-report. Caregiver-proxy reports are often used as an alternative. The caregiver’s perception of the child’s health and well-being typically influences use of health care services and care quality. Non-alignment of child and caregiver perceptions could result in undertreatment of symptoms if they are not recognized or overtreatment if the caregiver perceives a higher symptom burden than the child is experiencing.

Existing evidence suggests that caregivers are not always optimal reporters of children’s experiences. Correlations between parent and child reports for health-related quality of life (HRQOL) tend to be fair to moderate, with parents of children with cancer tending to rate their children’s HRQOL lower than children themselves do. The reasons for discrepancies are multifactorial. Prior work has identified that parents with symptoms of anxiety, depression, and parenting stress give lower proxy ratings of HRQOL for children with cancer. Congruence between caregiver and child report also improves over time from the diagnosis. Finally, others have found that the level of agreement between child and parent report for certain domains differs depending on the child’s age, with the biggest differences for adolescents in emotional, social, and school functioning and for younger children (aged 5-7 years) in physical functioning.

Most of this work has focused on global HRQOL in children capable of self-report. Circumstances such as very young age, cognitive disability, or critical illness necessitate the report of a caregiver proxy. Caregivers and children may have better alignment on specific symptom domains, especially highly visible physical symptoms, or when completing multiple measures of the same domain during short periods. In this study, we examined alignment between child and caregiver proxy reports in specific symptom domains using the National Institutes of Health Patient-Reported Outcomes Measurement Information System (PROMIS), which was designed to assess symptoms and functioning across a broad range of diseases and conditions. The PROMIS pediatric and parent proxy report measures were developed and validated in a diverse population of children, beginning with 2 single-site cohorts; follow-up work demonstrated construct validity and responsiveness in a large 9-center cohort of children undergoing cancer treatment. Using data from this same large multicenter cohort, this article aims to (1) examine the degree of alignment between child and caregiver proxy report for PROMIS pediatric domains of mobility, pain interference, fatigue, depression, anxiety, and psychological stress and (2) identify factors associated with better child and proxy congruence. Both aims provide a unique perspective of how the HRQOL of a child undergoing cancer treatment is viewed from the child’s and caregiver’s perspectives and inform best practices for use of PROMIS Pediatric and caregiver-proxy report in clinical research and care settings. We hypothesized that child and caregiver-proxy reports for the PROMIS Pediatric measures would have high congruence, that agreement would be better for observable domains such as physical functioning and pain and lower for less observable domains such as fatigue and depression, and that proxies’ own psychological and health symptoms would exacerbate differences.

**Methods**

**Participants**

Children with cancer and their primary caregivers from 9 pediatric oncology centers participated in this study. Sites included Children’s Healthcare of Atlanta, Georgia; Children’s Hospital Los Angeles, California; UPMC (University of Pittsburgh Medical Center) Children’s Hospital of Pittsburgh, Pennsylvania; Children’s National Hospital, Washington, DC; Dana-Farber Cancer Institute, Boston Children’s Hospital, Boston, Massachusetts; Duke University, Durham, North Carolina; The Hospital for Sick Children, Toronto, Ontario, Canada; St Jude Children’s Research Hospital, Memphis, Tennessee; and the University of North Carolina at Chapel Hill. All sites obtained institutional review board approval. Caregivers provided written informed consent, and children and adolescents provided assent. Participants received a $10 gift card at each of 2 assessment points. This study followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guideline.

Eligible children and adolescents were aged 7 to 18 years (hereinafter referred to as children), had a first diagnosis of cancer of any type, and were receiving cancer therapy consisting of surgery, chemotherapy, radiotherapy, and/or bone marrow transplant. Children must have completed 1 month of frontline treatment. If surgery was part of treatment, 3 to 6 weeks must have passed from the date of surgery for the child to be eligible. The child must have been able to read (or listen to) and understand English and have no clinically significant cognitive or memory impairment based on the judgement of the site
investigator. The eligible child’s caregiver must have agreed to participate.

Study Design
Children and caregivers completed questionnaires at 2 points. One point was when the child was expected to have relatively low symptom burden and higher functional performance (T1); the second was when the child was expected to have higher symptom burden and functional limitations (T2). The window between T1 and T2 was selected to allow for treatment schedules with different expected symptom patterns. For patients receiving chemotherapy, T1 was within 72 hours before the start of a chemotherapy cycle, and T2 was 7 to 17 days later. For patients receiving radiotherapy, T1 was within 72 hours before the start of radiotherapy and T2 was 4 to 6 weeks later.

Questionnaires were completed on tablet computers during health care visits, with paper backup if the network was down. A study team member (including M.M., J.S.W., N.R.L., and D.T.) was present to assist the child and caregiver and to ensure that the child and caregiver did not influence each other’s answers. If the child and caregiver were not in the clinic for T2, electronic survey links were sent to the caregiver requesting completion from home; for home surveys, participants were instructed to complete surveys separately, but no additional oversight was provided.

Measures
The PROMIS Pediatric and parent-proxy report measures include assessments of the child’s physical function (mobility) and symptoms, including pain interference, fatigue, depressive symptoms, anxiety, and psychological stress. Each question’s recall period is the past 7 days; questions have 5 response categories. PROMIS Pediatric and proxy report scores are linked together on a common T-score metric with a mean of 50 (SD, 10) in the original calibration sample. Higher scores for symptom domains represent worse symptom experiences; higher scores for mobility, better functioning. A 3-point difference in PROMIS pediatric scores is considered a minimally important difference.

Caregivers also completed the PROMIS-29 about their own HRQOL at T1 and T2. Domains match those in PROMIS Pediatric except for psychological stress, which is not included in PROMIS-29. Like PROMIS Pediatric measures, PROMIS-29 uses a recall period of the past 7 days with 5 response categories. PROMIS-29 scores are scaled on a T-score metric mean of 50 (SD, 10) relative to the US general population.

Statistical Analysis
We used a 1-way random-effects model to calculate intraclass correlations (ICCs) and corresponding 95% CIs for assessing absolute agreement between child self-report and caregiver-proxy report for each HRQOL domain. The ICCs were evaluated separately by age group (7-12, 13-15, and 16-18 years) and across all ages combined. The ICC magnitudes were classified as poor (<0.50), moderate (0.50 to <0.75), good (0.75-0.90), and excellent (>0.90). Because of nonnormality of score distributions, we performed a sensitivity analysis to examine the association between child self-report and caregiver-proxy report on PROMIS measures using nonparametric ICC methods described in Rothery and the R nprocago package for nonparametric concordance coefficient, version 1.0.6 (R Foundation for Statistical Computing). Throughout the article, we present results for T2, because this is when the child was expected to be more symptomatic. Results at T1 are provided in eTables 1 and 2 in the Supplement.

We calculated mean child- and caregiver-reported scores and 95% CIs for each HRQOL domain. Scores were calculated by age group and across all ages at T1 and T2 and compared using independent sample 2-tailed t tests. Bland-Altman plots were used to display agreement between child and caregiver proxy scores as a function of the mean between child and caregiver scores.

We used multivariable mixed-effect models with both T1 and T2 data and random intercepts to identify child and caregiver characteristics associated with the magnitude of difference between child self-report and caregiver-proxy report for each HRQOL domain. Each domain was modeled independently. Models were adjusted for child age, sex, race/ethnicity, and treatment modality and caregiver educational level, sex, and marital status. In addition, each model was adjusted for caregiver self-reported PROMIS-29 score that was similar to the HRQOL domain being analyzed; for example, the model for magnitude of difference in fatigue scores between child and caregiver reports was adjusted for caregiver PROMIS-29 fatigue score. Caregiver self-reported scores for anxiety were used when psychological stress was the outcome, because there is no equivalent PROMIS-29 domain. In regression models, caregiver PROMIS-29 scores were centered on a mean score of 50, age was centered on the sample median of 13.5 years, and the most prevalent category was used as the reference group in every categorical child or caregiver characteristic variable to aid interpretation of the model intercept. We also explored an interaction term between child sex and caregiver PROMIS-29 score, but the term was not statistically significant and was not included in final models. All other characteristics were included in models regardless of statistical significance.

Analyses were conducted with SAS software, version 9.4, of the SAS System for UNIX (SAS Institute Inc). Two-tailed P < .05 indicated statistical significance.

Results

Patient and Caregiver-Proxy Characteristics
Of 580 child-caregiver dyads approached for participation, 88 (15.2%) declined, and 10 (1.7%) withdrew before completing the T1 survey. In total, 482 child-caregiver dyads completed surveys at T1, representing 83% of those approached, and 403 dyads (84%) completed surveys at T2. Two hundred twenty participating children were female (46%) and 262 were male (54%), 80 (17%) were Black, and 71 (15%) were Hispanic (Table 1). Patients represented diverse cancer types; 436 children (90%) received chemotherapy. Most caregivers were female (407 [84%]) and married (355 [74%]) but were well distributed across educational and income levels.
ICCs Between Child Self-report and Caregiver-Proxy Report
PROMIS Pediatric Scores

Table 2 shows ICCs between child self-report and caregiver-proxy report scores for the 6 PROMIS Pediatric domains at T2, stratified by age group. For all symptom domains, ICCs showed poor agreement (<0.50). The ICCs for mobility were in the moderate agreement range (0.48 [95% CI, 0.37-0.60] to 0.68 [95% CI, 0.58-0.77]). There were no statistical differences by age group, although the youngest age group had some of the lower ICCs (range, 0.25 [95% CI, 0.14-0.41] to 0.48 [95% CI, 0.37-0.60]). We saw similar patterns at T1 (eTable 1 in the Supplement).

In sensitivity analyses, nonparametric ICCs were slightly lower than parametric ICCs with associations from 0.43 (95% CI, 0.42-0.60) to 0.48 (95% CI, 0.47-0.60) in the total sample (eTables 2 and 3 in the Supplement).

Differences Between Child Self-report and Caregiver-Proxy Report
PROMIS Pediatric Scores

Figure 1 shows means and 95% CIs for PROMIS Pediatric scores at T2 for children and caregivers, stratified by child age group. Within stratified age groups, children consistently reported lower scores than caregivers for pain interference,
fatigue, depressive symptoms, anxiety, and psychological stress and higher scores for mobility (for mobility among those aged 16-18 years, \( t = 2.36211, P = .02 \); for depressive symptoms among those aged 16-18 years, \( t = -2.16211, P = .03 \)).

When all age groups were combined, independent-samples \( t \) tests demonstrated statistically significant differences (e.g., for mobility, \( t = 6.84811, P < .001 \)) between self-report and proxy report for all PROMIS measures. Findings were similar at T1 (eFigure 1 in the Supplement). Bland-Altman plots for T2 (Figure 2) and T1 (eFigure 2 in the Supplement) show differences between child and caregiver scores for each domain as a function of the mean between child and caregiver scores.

### Table 3. ICC of Child Self-report and Caregiver-Proxy Report PROMIS Pediatric Scores at Time 2

<table>
<thead>
<tr>
<th>PROMIS Pediatric domain</th>
<th>Child age group, ICC (95% CI)</th>
<th>7-12 y (n = 176)</th>
<th>13-15 y (n = 123)</th>
<th>16-18 y (n = 106)</th>
<th>Total (N = 405)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mobility</td>
<td>0.48 (0.37-0.60)</td>
<td>0.68 (0.58-0.77)</td>
<td>0.58 (0.45-0.70)</td>
<td>0.57 (0.50-0.63)</td>
<td></td>
</tr>
<tr>
<td>Pain interference</td>
<td>0.37 (0.26-0.51)</td>
<td>0.39 (0.26-0.55)</td>
<td>0.41 (0.26-0.57)</td>
<td>0.39 (0.31-0.47)</td>
<td></td>
</tr>
<tr>
<td>Fatigue</td>
<td>0.27 (0.15-0.42)</td>
<td>0.40 (0.26-0.55)</td>
<td>0.33 (0.18-0.51)</td>
<td>0.32 (0.24-0.41)</td>
<td></td>
</tr>
<tr>
<td>Depressive symptoms</td>
<td>0.25 (0.14-0.41)</td>
<td>0.43 (0.30-0.58)</td>
<td>0.36 (0.19-0.52)</td>
<td>0.33 (0.25-0.42)</td>
<td></td>
</tr>
<tr>
<td>Anxiety</td>
<td>0.37 (0.25-0.50)</td>
<td>0.37 (0.23-0.53)</td>
<td>0.38 (0.24-0.55)</td>
<td>0.37 (0.29-0.46)</td>
<td></td>
</tr>
<tr>
<td>Psychological stress</td>
<td>0.44 (0.32-0.56)</td>
<td>0.40 (0.26-0.55)</td>
<td>0.41 (0.27-0.57)</td>
<td>0.42 (0.34-0.50)</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: ICC, intraclass correlation; PROMIS, Patient-Reported Outcomes Measurement Information System.

Characteristics Associated With the Magnitude of Difference Between Child Self-report and Caregiver-Proxy Report PROMIS Pediatric Scores

Table 3 shows multivariable models of factors associated with the magnitude of difference between child self-report and caregiver-proxy report PROMIS Pediatric scores for each HRQOL domain. The difference was calculated as proxy report minus child self-report PROMIS Pediatric score for each child. The reference group consisted of non-Hispanic White male children receiving chemotherapy with a college-educated, married, female caregiver. For mobility, caregivers reported the child’s mobility score as 6.00 points worse (95% CI, −7.45 to −4.51) than children’s self-report at T2, exceeding the minimally im-
important difference of 3 points. For mobility, positive estimates are associated with smaller differences between child and caregiver-proxy scores (e.g., increasing child’s age by 1 year results in a 0.47 decrease in child and parent discrepancy). For symptoms, caregivers overestimated the child’s symptom levels, ranging from 5.79 (95% CI, 3.99-7.60) points for psychological stress to 13.69 (95% CI, 11.60-15.78) points for fatigue. For symptoms, positive estimates are associated with larger differences between child and caregiver-proxy scores.

Consistently, the difference between child self-report and caregiver-proxy report was smaller at T1 than T2; for example, for mobility, the difference between child self-report and caregiver-proxy report diminished by 1.38 points (95% CI, 0.49-2.27) at T1 relative to T2. There was better agreement for mobility for older children (0.47-point decrement in difference per year of age [95% CI, 0.28-0.66]). There was also better agreement for fatigue and psychological stress for girls than boys (−2.73-point difference in fatigue score [95% CI, −4.58 to −0.87] and −1.89-point difference in psychological stress score [95% CI, −3.51 to −0.27]), whereas agreement got worse for pain interference and fatigue for Asian compared with White children (3.84-point difference in pain interference score [95% CI, 0.65-7.02] and 6.45-point difference in fatigue score [95% CI, 2.10-10.80]). Caregivers with an educational level of high school or less had better agreement with their child for mobility (1.81-point difference [95% CI, 0.02-3.59]), fatigue (0.21-point difference [95% CI, 0.13-0.30]), and anxiety (0.35-point difference [95% CI, 0.27-0.42]). Interestingly, the caregiver’s self-reported HRQOL had a significant influence on all domains except mobility, suggesting that when caregiver HRQOL is poorer, they have larger proxy-report differences with child self-report. For example, for fatigue, the magnitude of difference between child and caregiver-proxy scores increased by 0.21 points (95% CI, 0.13-0.30) for each 1-point increase in the caregiver’s own fatigue score.

**Discussion**

In this large cohort study of children with cancer, agreement between reports by children themselves and their caregiver proxies for symptoms was poor and for mobility was moderate. In a previous study including children with a broad range of acute and chronic health conditions, investigators found...
similar patterns of association between self-report and caregiver report for PROMIS Pediatric domains: Pearson r = 0.63 for mobility, Pearson r = 0.49 for pain interference, Pearson r = 0.41 for fatigue, Pearson r = 0.37 for depressive symptoms, and Pearson r = 0.41 for anxiety. Other studies in children with cancer using different HRQOL measures reported a similar low association between child self-report and proxy report.16-18 Consistent with the literature, patient-proxy agreement is generally best for concrete or observable domains (eg, mobility) than less observable domains (eg, depression).19,40 Use of PROMIS Pediatric proxy measures could offer advantages compared with alternative measures of HRQOL when child self-report is not possible. However, for less observable symptoms, such as depression, child self-report is vital.

Caregivers consistently rated symptoms and function worse than children themselves. These differences, adjusting for child and caregiver characteristics, exceeded the minimally important difference of 3 points35 and had a difference of more than 0.5 SD on the PROMIS T-score metric (5 points).

This finding raises the question as to whose report is most reflective of the child’s experience. Difficulties with understanding and interpreting questions could lead to systematic underreporting of symptoms by children. However, the PROMIS Pediatric items have undergone extensive qualitative and quantitative evaluation, including in children with cancer.26,31,33,41 Cognitive interviews have shown that the caregiver's own HRQOL is an important factor associated with the magnitude of caregiver-child differences in response to similar care conditions.42 Thus, child comprehension of PROMIS pediatric items cannot account fully for child-proxy discrepancies.

An alternative explanation relates to the lens through which caregivers view their ill children. As in previous work, we identified the caregiver's own HRQOL as an important factor associated with the magnitude of caregiver-child differences in responding to similar care conditions.

Table 3. Differences Between the Child Self-report and Caregiver-Proxy Report of Child’s PROMIS Pediatric Scores Adjusting for Child and Caregiver Characteristics

<table>
<thead>
<tr>
<th>Characteristic vs reference group</th>
<th>Domain, estimated difference (95% CI)</th>
<th>Mobility</th>
<th>Pain interference</th>
<th>Fatigue</th>
<th>Depressive symptoms</th>
<th>Anxiety</th>
<th>Psychological stress</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adjusted mean scorea</td>
<td>-6.00 (-7.45 to -4.51)</td>
<td>7.47 (5.84 to 9.10)</td>
<td>13.69 (11.60 to 15.78)</td>
<td>7.30 (5.27 to 9.34)</td>
<td>7.30 (5.48 to 9.12)</td>
<td>5.79 (3.99 to 7.60)</td>
<td></td>
</tr>
<tr>
<td>Time 1 vs time 2</td>
<td>1.38 (0.49 to 2.27)</td>
<td>-1.46 (-2.52 to -0.40)</td>
<td>-2.54 (-3.81 to -1.27)</td>
<td>-2.50 (-3.66 to -1.34)</td>
<td>-0.96 (-2.02 to 0.10)</td>
<td>-2.66 (-3.64 to -1.67)</td>
<td></td>
</tr>
</tbody>
</table>

**Child characteristics**

- **Age per year**: 0.47 (0.28 to 0.66) for mobility, -0.03 (-0.24 to 0.17) for pain interference, -0.26 (-0.54 to 0.02) for fatigue, -0.16 (-0.44 to 0.11) for depressive symptoms, 0.13 (-0.11 to 0.38) for anxiety, 0.11 (-0.14 to 0.35) for psychological stress.
- **Female vs male**: 0.76 (-0.53 to 2.05) for mobility, -1.09 (-2.46 to -0.28) for pain interference, -2.73 (-4.58 to -0.87) for fatigue, -1.58 (-3.39 to -0.23) for depressive symptoms, -0.94 (-2.56 to 0.68) for anxiety, -1.89 (-3.51 to -0.27) for psychological stress.
- **Hispanic vs non-Hispanic**: -0.13 (-2.01 to 1.75) for mobility, -0.52 (-2.51 to 1.48) for pain interference, -1.36 (-4.06 to 1.35) for fatigue, -1.82 (-4.46 to 0.82) for depressive symptoms, -0.26 (-2.62 to 2.09) for anxiety, -0.32 (-2.67 to 2.04) for psychological stress.
- **Asian vs White**: -0.97 (-3.98 to 2.05) for mobility, 3.84 (0.65 to 7.02) for pain interference, 6.45 (2.10 to 10.80) for fatigue, 2.34 (-1.90 to 6.59) for depressive symptoms, 3.13 (-0.65 to 6.91) for anxiety, 3.37 (-0.43 to 7.17) for psychological stress.
- **Black vs White**: 2.13 (0.23 to 4.03) for mobility, -0.77 (-2.78 to 1.24) for pain interference, -2.02 (-4.77 to 0.74) for fatigue, -0.01 (-2.68 to 2.66) for depressive symptoms, -2.16 (-4.55 to 0.23) for anxiety, -0.72 (-3.11 to 1.66) for psychological stress.
- **Other vs White**: -0.73 (-3.02 to 1.55) for mobility, 0.40 (-2.03 to 2.84) for pain interference, -3.84 (-7.13 to -0.54) for fatigue, 0.62 (-2.61 to 3.85) for depressive symptoms, -1.77 (-4.64 to 1.10) for anxiety, -0.67 (-3.54 to 2.19) for psychological stress.
- **BMT vs chemotherapy**: -2.88 (-7.47 to 1.70) for mobility, 3.12 (-1.73 to 7.97) for pain interference, 3.34 (-3.26 to 9.94) for fatigue, 1.95 (-4.50 to 8.40) for depressive symptoms, -0.47 (-6.22 to 5.29) for anxiety, 3.44 (-2.39 to 9.28) for psychological stress.
- **Radiotherapy vs chemotherapy**: 0.55 (-1.99 to 3.09) for mobility, 0.37 (-2.32 to 3.05) for pain interference, -1.66 (-5.31 to 1.99) for fatigue, 1.33 (-2.24 to 4.90) for depressive symptoms, 0.60 (-2.59 to 3.79) for anxiety, 0.78 (-2.42 to 3.98) for psychological stress.

**Caregiver characteristics**

- **PRO per child unitb**: -0.02 (-0.09 to 0.06) for mobility, 0.16 (0.08 to 0.24) for pain interference, 0.21 (0.13 to 0.30) for fatigue, 0.27 (0.18 to 0.37) for depressive symptoms, 0.35 (0.27 to 0.42) for anxiety, 0.26 (0.19 to 0.33) for psychological stress.
- **Male vs female**: 2.13 (0.25 to 4.00) for mobility, -0.47 (-2.48 to 1.54) for pain interference, -1.99 (-4.70 to 0.73) for fatigue, -2.21 (-4.84 to 0.43) for depressive symptoms, -2.47 (-4.83 to -0.12) for anxiety, -2.64 (-4.99 to -0.28) for psychological stress.
- **High school educational level vs less college**: 1.81 (0.02 to 3.59) for mobility, -1.16 (-3.06 to 0.74) for pain interference, -3.23 (-5.80 to -0.67) for fatigue, -1.18 (-3.68 to 1.32) for depressive symptoms, -2.25 (-4.49 to -0.02) for anxiety, -1.82 (-4.06 to 0.42) for psychological stress.
- **Some college vs college**: -0.21 (-1.89 to 1.48) for mobility, -0.52 (-2.31 to 1.27) for pain interference, -1.20 (-3.63 to 1.23) for fatigue, -0.89 (-3.26 to 1.47) for depressive symptoms, -3.06 (-5.18 to -0.94) for anxiety, -2.25 (-4.36 to -0.13) for psychological stress.
- **Postgraduate educational level vs college**: 1.52 (-0.46 to 3.51) for mobility, 0.94 (-1.17 to 3.05) for pain interference, 1.61 (-1.25 to 4.47) for fatigue, 1.49 (-1.31 to 4.29) for depressive symptoms, 2.11 (-0.39 to 4.60) for anxiety, 1.82 (-0.69 to 4.32) for psychological stress.
- **Single vs married**: -1.07 (-2.65 to 0.51) for mobility, -0.10 (-1.78 to 1.58) for pain interference, -0.94 (-3.22 to 1.33) for fatigue, -1.23 (-3.45 to 0.99) for depressive symptoms, 0.47 (-1.51 to 2.45) for anxiety, 0.89 (-1.09 to 2.87) for psychological stress.

Abbreviations: BMT, bone marrow transplant; PRO, patient-reported outcome; PROMIS, Patient-Reported Outcomes Measurement Information System.

* The adjusted mean scores (intercept) are the differences in scores between child self-report and caregiver proxy report with reference group of time 2, a male non-Hispanic child aged 12.9 years and receiving chemotherapy, and a married female caregiver with a college educational level. Income was not included in the model owing to collinearity issues with parent educational level. The parents' PROMIS-29 scores were centered on a mean of 50 (SD, 10).

b The caregiver’s PRO was the same outcome being evaluated in the model. For example, the model looking at the difference between the child's self-reported fatigue and the caregiver's proxy-reported fatigue included the PRO of the caregiver's own self-reported fatigue level. There is no adult version of the psychological stress domain, so we used the parent's anxiety score in the model.
porting. When caregivers felt more anxious themselves, for example, they considered their children to have symptoms of psychological stress in excess of the child's report. Caregivers who are especially sad or worried may ascribe more sadness or worry to their children owing to heightened concern about their well-being. However, this phenomenon also limits the extent to which caregiver reports offer the most accurate window into the child’s experience.

In addition, even after controlling for the caregiver’s own HRQOL, we still found discrepancies ranging from 5.79 points of difference (psychological stress) to 13.69 points of difference (fatigue), exceeding the 3-point minimally important difference. Calman43 has theorized that patient-proxy discrepancies in HRQOL reports may relate to different expectations. As patients adapt to illness, their expectations conform more closely to actual experiences, and reported HRQOL improves, a phenomenon sometimes referred to as recalibration response shift.44 Proxies, on the other hand, may hold higher expectations for HRQOL because their expectations fail to adapt to the same degree. This could be especially true for caregivers, whose expectations for child HRQOL may be determined in comparison with other healthy children or the child’s life before cancer. Understandably, the caregiver views the child through a caregiver’s eyes, and their reports should be understood as such.

Additional factors associated with the magnitude of difference between child and caregiver report are worthy of mention. We identified better child-proxy agreement at T1, when children were expected to have lower symptom burden, than at T2, suggesting that caregiver assessments may be more accurate when children are feeling relatively well. Caregivers of girls had closer alignment with child report in select domains, perhaps reflecting sex differences in caregiver-child communication about symptoms. Asian caregivers, in contrast, overestimated some symptoms relative to child report, as did caregivers with higher educational attainment. Both findings could be rooted in caregiver expectations for their child’s experience, with higher expectations for HRQOL or failure to adapt to experience among Asian and highly educated caregivers.

Limitations and Strengths
Limitations include the fact that this study enrolled only English-speaking children and caregivers in the United States and Canada; generalizability to other languages and settings requires further study. Children were recruited during initial cancer treatment; children with relapse were not eligible, although these children are often highly symptomatic. Most participating children received chemotherapy, with fewer participating after radiotherapy or surgery, limiting our ability to detect small differences in reporting in these settings. We lacked complete data on whether surveys were administered in the clinic or at home and were unable to adjust for administration mode in analyses. The study also has some notable strengths, including participation of a large and diverse cohort of children and caregivers from 9 cancer centers.

Conclusions
The PROMIS Pediatric measures allow children to report on their own experiences of symptoms and function during cancer therapy. This cohort study found that agreement between child and caregiver proxy was poor for symptoms and moderate for mobility. Caregivers tended to overestimate symptoms and underestimate function relative to children themselves. Our findings suggest that proxy reporting is influenced by the proxy’s personal experience of symptoms and function as well as the child’s experience. Although these findings may limit use of proxy reports when child report is possible, understanding the ways in which proxy reports differ from those of children can facilitate thoughtful incorporation of proxy report when child report is not possible. This is particularly important given that overestimation of symptoms by caregivers may result in overtreatment of symptoms.43 Our findings underscore the importance of allowing children to speak to their own experiences, about which they are the experts.
Effect of Patient-Reported Outcomes Monitoring on Health Outcomes in Pediatric Oncology:
A Systematic Review and Meta-Analysis

Funding/Support: This study was supported by award U19AR065952 from the National Institutes of Health, National Center on Sleep Disorders Research, and award R01AT007575 from the National Institute on Aging.

Role of the Funders/Sponsors: The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Additional Contributions: We thank the children, adolescents, and their families for participating in this study and giving us their time to help us create better measures to capture their experiences.

REFERENCES

READER RESPONSE: How do you think patient-reported outcomes monitoring can improve cancer care for children and adolescents? Please share your thoughts in the comments below.


