

Assessing responsiveness over time of the PROMIS[®] pediatric symptom and function measures in cancer, nephrotic syndrome, and sickle cell disease

Bryce B. Reeve^{1,2} · Lloyd J. Edwards³ · Byron C. Jaeger³ · Pamela S. Hinds^{4,5} · Carlton Dampier^{6,7} · Debbie S. Gipson⁸ · David T. Selewski⁸ · Jonathan P. Troost⁸ · David Thissen⁹ · Vaughn Barry^{6,7} · Heather E. Gross¹ · Darren A. DeWalt¹⁰

Accepted: 29 August 2017 / Published online: 7 September 2017
© Springer International Publishing AG 2017

Abstract

Purpose Previous studies provided evidence for the validity of the PROMIS Pediatric measures in cross-sectional studies. This study evaluated the ability of the PROMIS Pediatric measures to detect change over time in children and adolescents with cancer, nephrotic syndrome (NS), or sickle cell disease (SCD).

Methods Participants (8–17 years) completed measures of fatigue, pain interference, anger, anxiety, depressive

symptoms, mobility, upper extremity, and peer relationships at three or four time points (T1–T4). Between T1 and T2, children with cancer received chemotherapy and children with SCD experienced a pain exacerbation. Children with NS were first assessed during active disease (T2), with T3 and T4 conducted at disease remission. For the primary analysis of responsiveness, we expected better scores at T3 (recovery) compared to T2 (event) for all diseases. T1 and T4 are also expected to have better scores than T2. Linear mixed models were used and adjusted for time, gender, age, race/ethnicity, education, comorbid conditions, and disease.

Results Enrolled were 96 children with cancer, 121 children with SCD, and 127 children with NS. Fatigue, pain interference, mobility, and upper extremity scores worsened from T1 (baseline) to T2 (event) ($p < 0.01$), and significantly improved from T2 to T3 and T4 ($p < 0.01$). Similarly, anxiety and depressive symptoms significantly improved from T2 to T3 and T4 ($p < 0.01$).

Conclusions This study provides evidence for the responsiveness of seven PROMIS Pediatric measures to clinical disease state in three chronic illnesses. The findings support use of PROMIS Pediatric measures in clinical research.

Keywords Patient-reported outcomes · Responsiveness · Pediatrics · Cancer · Nephrotic syndrome · Sickle cell disease

Introduction

The National Institutes of Health (NIH) initiated the patient-reported outcomes measurement information system[®] (PROMIS[®]) project in 2004 with the goal to provide researchers and clinicians access to standardized,

✉ Bryce B. Reeve
bbreeve@email.unc.edu

¹ Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, NC, USA

² Department of Health Policy and Management, Gillings School of Global Public Health, University of North Carolina, 1101-D McGavran-Greenberg Hall, CB#7411, 135 Dauer Dr, Chapel Hill, NC 27599-7411, USA

³ Department of Biostatistics, University of North Carolina, Chapel Hill, NC, USA

⁴ Department of Nursing Research and Quality Outcomes, Children's National Health System, Washington, DC, USA

⁵ Department of Pediatrics, George Washington University, Washington, DC, USA

⁶ Department of Pediatrics, Emory University School of Medicine, Atlanta, GA, USA

⁷ AFLAC Cancer and Blood Disorders Center, Children's Healthcare of Atlanta, Atlanta, GA, USA

⁸ C.S. Mott Children's Hospital, University of Michigan, Ann Arbor, MI, USA

⁹ Department of Psychology and Neuroscience, University of North Carolina, Chapel Hill, NC, USA

¹⁰ Division of General Medicine and Clinical Epidemiology, UNC School of Medicine, Chapel Hill, NC, USA

psychometrically robust, patient-reported measures (pediatric and adult) of key symptom and function domains. A central principal is that these measures could be useful across a broad range of conditions and diseases to capture the impact of the disease and treatment on the lives of the patients. The non-disease specific measures of the PROMIS measurement system allow the comparison of scores from one group of individuals to another to evaluate relative disease burden. Given the broad scope of the PROMIS measures, it is critical to evaluate the psychometric properties in multiple disease populations to provide evidence that the measures are valid and reliable assessments of the symptom and function domains they measure.

The PROMIS Pediatric measures are designed for children and adolescents between 8 and 17 years of age. Extensive qualitative and quantitative methods have been used to design and evaluate, in multiple diverse patient groups, the PROMIS measures of: physical function—mobility [1], physical function—upper extremity [1], pain interference [2], fatigue [3], depressive symptoms [4], anxiety [4], anger [5], and peer relationships [6]. In addition, a multi-site study collected data from 1447 children and adolescents with chronic health conditions, including sickle cell disease, kidney disease, cancer, rehabilitative needs, obesity, and rheumatic disease, in order to validate the PROMIS Pediatric measures in cross-sectional studies [7–11].

It is also critical to evaluate the responsiveness of the PROMIS Pediatric measures over time and changing clinical situations. Responsiveness is an aspect of validity that indicates the ability of a PRO measure to detect change over time when it is expected. For PROMIS Pediatric measures to be adopted for use in clinical trials, there must be evidence of their responsiveness to inform the efficacy evaluation of the intervention or the disease treatment impact under study. The goal of this study was to evaluate the responsiveness of eight PROMIS Pediatric measures in diverse samples of children and adolescents with cancer, sickle cell disease (SCD), or nephrotic syndrome (NS).

Methods

Participants and study design

The University of North Carolina (UNC) served as the central coordinating center to support the following sites focused on specific diseases: Children’s National Health System—Cancer; University of Michigan—Nephrotic Syndrome, and Emory University—Sickle Cell Disease. The study reported here represents a secondary analysis of three separate studies of responsiveness of the PROMIS Pediatric measures at their respective sites. Each site had

the freedom to select the number of assessment points, timing of the assessment points, sample size, and PROMIS domains; thus, there is heterogeneity across the three diseases because of the differential experience of symptom burden and functional status by disease. However, each had a common goal to assess responsiveness over time of the PROMIS Pediatric measures for children experiencing changing health status.

Common eligibility criteria across all diseases included children and adolescents between the ages of 8 and 17 years of age, ability to read and speak English (because at the time of the study, there were no translations of the PROMIS Pediatric measures), functional computer skills (defined as the ability to see and interact with a computer screen, keyboard, and mouse), and willingness to give written assent/permission for study participation. Excluded were children and adolescents who had any concurrent medical or psychiatric condition that precluded study participation, or cognitive or other (e.g., visual) impairments that interfered with completing a self-administered, computer-based questionnaire. Additional eligibility criteria specific to a disease are provided below. All sites received approval from their respective Institutional Review Boards.

Cancer

Eligible children and adolescents were diagnosed with a childhood cancer, scheduled to receive a course of anti-neoplastic (not biologic agent only) chemotherapy from course 2 forward, and not currently enrolled on a Phase 1 clinical trial. The study included three time points. Time 1 (T1) occurred 1–2 days before an early course of chemotherapy. Time 2 (T2) occurred 7–16 days following chemotherapy initiation at the time when the patient’s nadir was projected. T3 occurred within 1–2 days preceding the next course of scheduled chemotherapy or approximately 2 weeks following T2. More study details are provided elsewhere [12].

Based on clinical experience, we hypothesized that PROMIS physical symptoms and function scores at T1 would be within the normal range; however, children would have elevated emotional distress (depression and anxiety) given the proximity of the assessment relative to the cancer diagnosis and start-up of chemotherapy. Following chemotherapy and at the projected time of the nadir (T2), we hypothesized physical symptom and function scores would be worse compared to T1 (except depression and anxiety), and that the symptom and function scores would be improved at T3 compared to T2. The child’s relationship with peers was not expected to change due to the relatively short treatment cycles.

Table 1 Characteristics of children and adolescents with cancer, nephrotic syndrome, or sickle cell disease

Characteristics	Cancer (<i>N</i> = 96) <i>N</i> (%)	Nephrotic syndrome (<i>N</i> = 127) <i>N</i> (%)	Sickle cell disease (<i>N</i> = 121) <i>N</i> (%)
Gender			
Female	44 (46)	44 (35)	68 (56)
Male	52 (54)	83 (65)	53 (44)
Age (years)			
Mean (SD)	13.09 (2.96)	12.31 (2.77)	12.50 (3.05)
Race			
Non-white	48 (50)	61 (48)	121 (100)
White	48 (50)	66 (52)	0 (0)
Hispanic			
Yes	19 (20)	10 (8)	3 (3)
Comorbid conditions			
0	70 (73)	86 (68)	53 (44)
1+	22 (23)	41 (32)	36 (30)
Maternal education			
High school or less	33 (34)	73 (58)	53 (44)
Some college or greater	63 (66)	54 (43)	68 (56)
Completion rates ^a			
T1 (baseline)	96 (100)	–	121 (100)
T2 (event)	84 (88)	127 (100)	55 (45)
T3 (recovery)	86 (90)	112 (88)	38 (31)
T4 (follow-up)	–	90 (71)	80 (66)

^a For each disease, a separate study design was used to assess responsiveness of the PROMIS pediatric measures. Nephrotic syndrome did not have a T1 assessment when the child was in remission before active disease state (at study enrollment). Cancer did not have a final T4 assessment point after the recovery period. Sickle Cell disease enrolled children when their symptoms were well controlled and followed them for a possible pain exacerbation (requiring hospitalization) over the course of 18 months; thus, only a subset of children with sickle cell disease provided T2 and T3 data

Nephrotic syndrome

Eligible children and adolescents included those with active nephrotic syndrome defined as the presence of nephrotic range proteinuria (≥ 2 + urinalysis and edema or urine protein/creatinine ratio >2 g/g) at the time of the first PROMIS assessment. Participants were recruited from 14 academic medical centers across the US and Canada.

Children completed questionnaires at three time points. For purposes of this study's analyses, there was no T1 (baseline) comparable measure similar to the children with cancer or sickle cell disease. Children with NS were first assessed during active disease (T2), with T3 conducted when they reached complete remission or at 3 months after T2 if remission did not occur and a subsequent follow-up (T4) at 12 months after T2. More study details are provided elsewhere [13].

We hypothesized that T2 would be the worst symptom and poorest functional status for the children experiencing NS activity. We hypothesized that symptom levels and functioning would be better at T3 and T4 follow-up periods when the children were in remission. Peer relationships

scores were not expected to change because of change in health status.

Sickle cell disease

A convenience sample of SCD patients was recruited during routine clinic visits at three clinical sites that were part of the same large SCD program (Children's Healthcare of Atlanta). Eligibility criteria included one or more acute care visits for pain in the previous year. More details on the study are provided elsewhere [14, 15].

Children completed questionnaires at up to four time points. T1 provided a baseline assessment of the child's health status. T2 assessment occurred at the end of a subsequent hospitalization for a pain exacerbation, 16.6 ± 19.1 months from baseline visit. T3 pain recovery assessment occurred at a median interval of 20 days (range 7–67 days) from their hospitalization assessment. T4 occurred at a subsequent routine clinic visit 1.5 ± 0.56 years following T1. Not all children enrolled in the study provided T2 and T3 data as only 45% of the children experienced a pain exacerbation that led to a

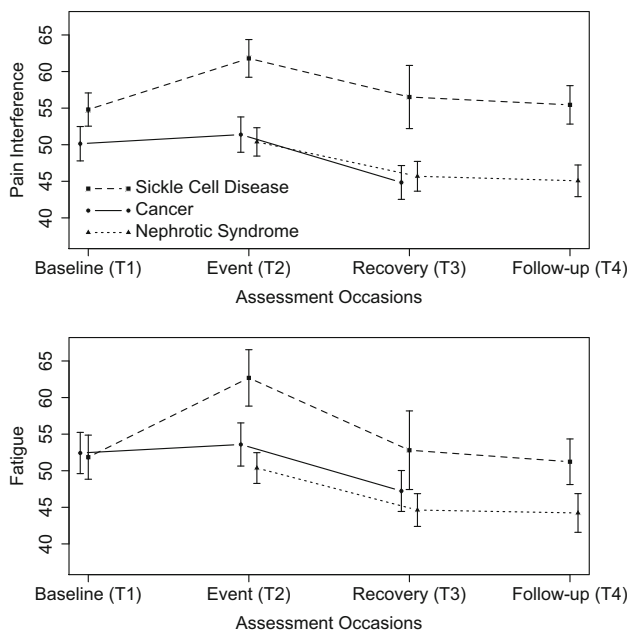


Fig. 1 Upper panel average PROMIS pediatric pain interference scores at T1–T4 for the sickle cell disease, cancer, and nephrotic syndrome studies. The vertical bars around each average are twice the standard error of the means, for approximately 95% confidence intervals. Lower panel as above, for PROMIS pediatric fatigue scores

hospitalization. Thus, sample sizes for T2 and T3 will be lower than T1 and T4 as not all children experienced a pain exacerbation.

We hypothesized that children with SCD at T2 to have the worst symptom levels and poorest functioning due to the pain exacerbation compared with T1, T3, and T4. The greatest change would occur between T2 and T3 (recovery phase) with pain interference showing the biggest change relative to other symptoms. Because of the short duration of the pain episode, peer relationships status were not expected to change.

Measures

The vast majority of participating children completed computerized-adaptive testing (CAT) versions of the PROMIS Pediatric measures of pain interference, mobility, upper extremity, fatigue, depressive symptoms, anxiety, anger, and peer relationships. The CAT and other data were collected using the Assessment Center platform (<https://www.assessmentcenter.net>). CAT tailors the questionnaire for the participant by selecting appropriately informative questions based on the individual's responses to previously completed questions [16]. The result is a reliable assessment with minimal response burden. If the participants did not have access to the web, then the PROMIS pediatric

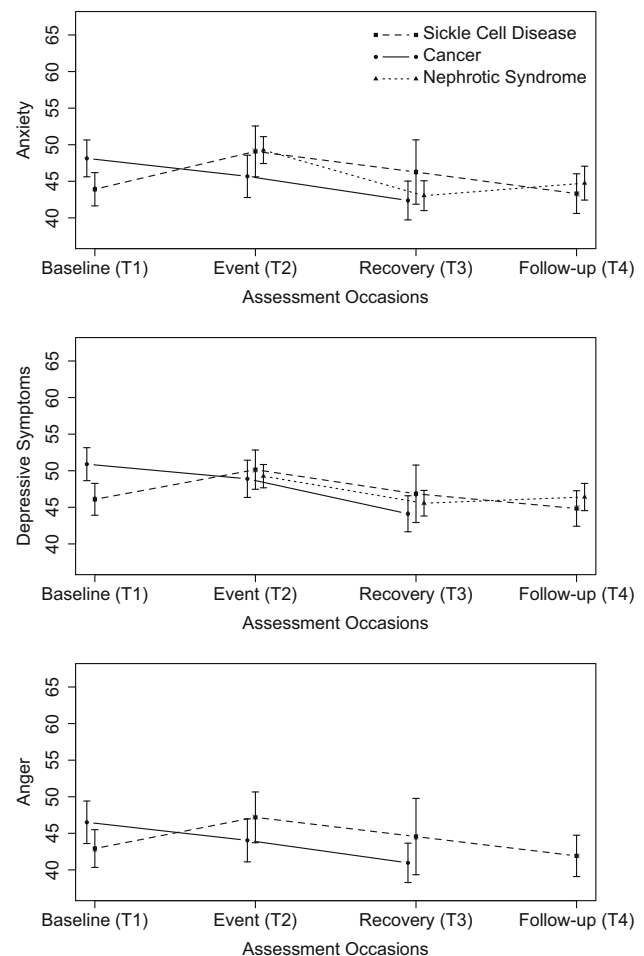


Fig. 2 Upper panel average PROMIS pediatric anxiety scores at T1–T4 for the sickle cell disease, cancer, and nephrotic syndrome studies. The vertical bars around each average are twice the standard error of the means, for approximately 95% confidence intervals. Center panel as above, for PROMIS pediatric depressive symptoms scores. Lower panel as above, for PROMIS pediatric anger scores

8-item short form for each health domain was used. PROMIS pediatric measures are scored on a T-score metric with mean of 50 and standard deviation of 10 in the original PROMIS pediatric item bank calibration [17]. Higher scores for symptom measures (i.e., fatigue, pain interference, depressive symptoms, anxiety, anger) indicate worse symptom experiences, whereas higher scores for function measures (i.e., mobility, upper extremity) and peer relationships indicate better functioning or relationships, respectively. In a previous study involving children diagnosed with one of five chronic diseases including cancer, SCD, and NS, a 3-point change was determined to be a minimally important difference (MID) [18]. The MID is defined as “the smallest difference in scores of a PRO measure that is perceived by patients as beneficial or harmful, and which would lead the clinician to consider a change in treatment” [19]. The MID was used to identify

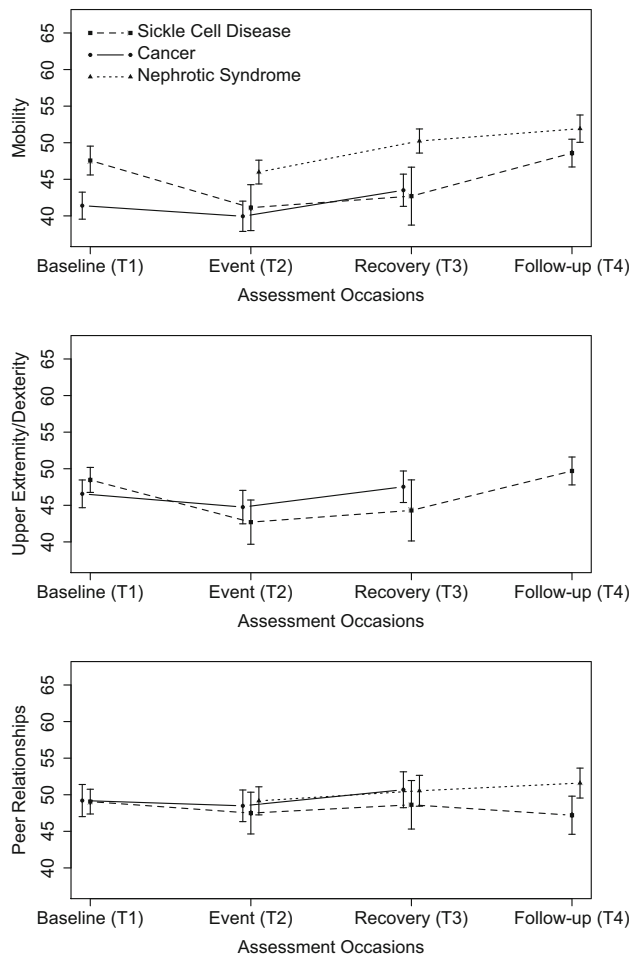


Fig. 3 Upper panel average PROMIS pediatric physical functioning: mobility scores at T1–T4 for the sickle cell disease, cancer, and nephrotic syndrome studies. The vertical bars around each average are twice the standard error of the means, for approximately 95% confidence intervals. Center panel as above, for PROMIS pediatric physical functioning: upper extremity/dexterity scores. Lower panel as above, for PROMIS pediatric peer relationships scores

meaningful change (responsiveness) scores, as determined by patients, parents, and clinicians, beyond findings from a statistical analysis.

Self-reported demographic data were also collected for each sample. For the analyses reported here, only the demographic data captured in common across the three sites were included. These variables include: gender (female or male), race (white vs. non-white), ethnicity (Hispanic or not), age (continuous), number of comorbid conditions (none vs. one or more conditions), and maternal education (high school or less vs. some college or more). The presence of other health conditions and highest level of maternal education were reported by the parent/guardian of the participating patient and captured on study case report forms.

Analyses

Descriptive statistics were computed for all study variables by disease and by time. The time duration for measurement occasions varied among the three disease types. However, the study design for each disease type consisted of assessment points based on disease events that aligned as indicators of study events, but the time intervals between these assessment points were different based on disease course. Linear mixed models were fit using the lme4 R package [20] for the analysis of longitudinal measures. Each PROMIS outcome (e.g., fatigue, mobility) was modeled independently. Models were adjusted for time, gender, age, race/ethnicity, maternal education, comorbid conditions, and disease type. Additionally, subject-specific random intercepts controlled for heterogeneity between participants. Goodness of fit criteria were used to facilitate model selection, and included AIC [21], BIC [22], and R^2 [23]. R^2 was calculated using the r2glmm R package [24]. Of three possible structures for the time variable (categorical, continuous, or both), semi-partial R^2 statistics [23] indicated that the categorical formulation explained more variability than each of the other approaches. Missing data were assumed missing completely at random (MCAR) based on sensitivity analyses and previous work [15].

With the alignment of events (time points) standardized across the three diseases, T2 is anticipated to be the point of worst symptom experience and functional impact. For primary analyses of responsiveness, lower symptom scores and higher functional scores are anticipated at T3 (recover phase) compared with T2 (event). For secondary analyses, we compare T1 (baseline) and T4 (follow-up) with T2 (event).

Results

The study included 96 children with cancer, 121 children with SCD, and 127 children with NS. Gender was approximately equally distributed in the cancer and SCD groups, but males comprised 65% of the NS group. The child's mean age was similar across all three disease groups but race differed across the groups. All SCD children were black or African American (except one was mixed race), children with nephrotic syndrome were 28% black, 13% Asian, 8% Hispanic, and 8% other race, and children with cancer were 24% black, 20% Hispanic, and 13% other race. Maternal education was lowest in the SCD group compared to the other two groups (Table 1).

Table 2 Mixed model results for PROMIS pediatric symptoms

	Fatigue	Pain interference	Anger	Anxiety	Depressive symptoms
Intercept (adjusted ¹)	51.9 (47.8, 55.9)	49.3 (45.9, 52.8)	46.4 (41.3, 51.5)	47.7 (44.0, 51.3)	50.3 (47.1, 53.6)
T1 (baseline)	−3.1** (−5.3, −0.9)	−2.6** (−4.4, −0.7)	1.0 (−1.1, 3.0)	0.2 (−1.5, 1.9)	0.9 (−0.7, 2.4)
T3 (recovery)	−6.4*** (−8.3, −4.5)	−5.5*** (−7.1, −3.9)	−3.3** (−5.5, −1.1)	−4.3*** (−5.7, −2.9)	−3.7*** (−5.05, −2.43)
T4 (follow-up)	−6.1*** (−8.3, −3.9)	−4.0*** (−5.8, −2.1)	−2.0 (−4.9, 1.0)	−3.3*** (−4.9, −1.6)	−2.9*** (−4.46, −1.37)
Female	3.4** (0.9, 6.0)	1.9 (−0.2, 4.0)	2.0 (−1.3, 5.4)	1.6 (−0.7, 3.9)	0.4 (−1.58, 2.47)
Age	0.3 (−0.2, 0.7)	0.1 (−0.2, 0.5)	−0.3 (−0.8, 0.3)	−0.3 (−0.7, 0.1)	−0.1 (−0.45, 0.24)
1+ conditions	−1.2 (−3.9, 1.4)	1.1 (−1.1, 3.4)	−1.7 (−5.3, 1.9)	−1.5 (−4.0, 0.9)	−0.4 (−2.58, 1.75)
Not white	−1.7 (−4.7, 1.2)	−0.2 (−2.7, 2.3)	−3.0 (−7.7, 1.8)	−2.5 (−5.2, 0.3)	−2.7* (−5.11, −0.35)
Hispanic	3.4 (−0.9, 7.7)	1.2 (−2.5, 4.8)	3.9 (−1.8, 9.6)	3.4 (−0.6, 7.4)	1.3 (−2.17, 4.83)
H.S. or less	1.9 (−0.8, 4.6)	2.5* (0.2, 4.8)	1.7 (−2.2, 5.5)	1.8 (−0.7, 4.2)	1.8 (−0.40, 3.95)
Incident NS	−3.3* (−6.5, −0.1)	−1.6 (−4.3, 1.1)		2.1 (−0.8, 5.0)	0.2 (−2.3, 2.8)
Sickle cell	2.2 (−1.5, 5.9)	5.7*** (2.6, 8.9)	−0.7 (−5.1, 3.7)	0.2 (−3.2, 3.6)	−1.5 (−4.5, 1.5)
Model R^2	0.16	0.17	0.10	0.14	0.14
Observations	744	743	442	740	744
Log likelihood	−2889.4	−2756.7	−1658.9	−2715.5	−2663.1
AIC	5806.8	5541.4	3343.9	5459.1	5354.1
BIC	5871.4	5605.9	3397.0	5523.5	5418.7

AIC Akaike information criteria, BIC Bayesian information criteria

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

^a The intercept can be interpreted as the mean symptom score at T2 (event) for the reference group who are non-Hispanic white, 8-years old boys with cancer, no other health condition and a mother with education more than high school

Evaluation of responsiveness of the PROMIS pediatric measures

Figures 1, 2 and 3 show the average PROMIS Pediatric scores at T1–T4 for the SCD, cancer, and NS studies. In Fig. 1, peak average symptom scores are at T2, the event, as expected for the Pain Interference and Fatigue measures. In Fig. 2, average emotional distress scores (Anxiety, Depressive Symptoms, and Anger) are higher at T2 than at T3 or T4, and higher at T2 than T1 for the SCD study. For the cancer study, emotional distress is at the highest level at T1. Figure 3 shows the pattern of the functioning scores across T1–T4 for the three studies; for the Physical Functioning and Peer Relationships scores, the trends are more subtle, but there is a tendency for the lowest average functioning scores and relationship score to be at T2, as expected for the function domains.

Table 2 provides the mixed model results for each of the PROMIS Pediatric symptom domains of Fatigue, Pain Interference, Anger, Anxiety, and Depressive Symptoms. The reference time point was set at T2 for all diseases as this was the point when the child or adolescent was expected to have the worst health status relative to other

time points due to chemotherapy (cancer), active disease (nephrotic syndrome), or pain exacerbation (sickle cell disease). Thus, negative regression weights (b) for T1, T3, and T4 in Table 2 indicate symptoms were less severe than T2. From T1 (baseline) to T2 (event), Fatigue ($b = -3.1$, $p < 0.01$) and Pain Interference ($b = -2.6$, $p < 0.01$) scores significantly worsened. From T2 (event) to T3 (recovery), all symptoms improved: Fatigue ($b = -6.4$, $p < 0.001$), Pain Interference ($b = -5.5$, $p < 0.001$), Anger ($b = -3.3$, $p < 0.001$); Anxiety ($b = -4.3$, $p < 0.001$); Depressive Symptoms ($b = -3.7$, $p < 0.001$). All symptom mean change scores from T2 to T3 exceeded the MID of 3 points. From T2 (event) to T4 (follow-up), all symptoms except Anger improved ($p < 0.01$).

Table 3 provides the mixed model results for both of the PROMIS Pediatric function domains of Physical Function—Mobility, Physical Function—Upper Extremity, and Peer Relationships. From T1 (baseline) to T2 (event), both function domains and relationships decreased on average (represented by positive regression weights in PROMIS T-score units) ($p < 0.05$). From T2 (event) to T3 (recovery), the function scores improved: Mobility ($b = 3.7$, $p < 0.001$), Upper Extremity ($b = 3.1$, $p < 0.01$), Peer

Table 3 Mixed model results for PROMIS Pediatric function domains

	Mobility	Upper extremity	Peer relationships
Intercept (adjusted ^a)	40.4 (37.5, 43.3)	42.5 (39.0, 45.9)	47.7 (44.5, 51.0)
T1 (baseline)	2.8*** (1.3, 4.4)	3.0*** (1.3, 4.8)	1.8* (0.2, 3.5)
T3 (recovery)	3.7*** (2.4, 5.0)	3.1** (1.2, 5.0)	1.6* (0.2, 3.0)
T4 (follow-up)	5.3*** (3.8, 6.8)	5.6*** (3.0, 8.1)	2.0* (0.4, 3.7)
Female	−1.7 (−3.5, 0.1)	−4.0*** (−6.3, −1.8)	−0.1 (−2.1, 2.0)
Age	−0.1 (−0.4, 0.2)	0.6** (0.2, 1.0)	0.6** (0.2, 0.9)
1+ Conditions	−0.8 (−2.7, 1.1)	0.6 (−1.8, 3.0)	−1.5 (−3.7, 0.7)
Not White	1.5 (−0.7, 3.6)	1.5 (−1.6, 4.6)	−2.2 (−4.6, 0.2)
Hispanic	−1.0 (−4.1, 2.1)	−3.1 (−6.9, 0.6)	−1.4 (−4.9, 2.1)
High school or less	−1.3 (−3.2, 0.7)	−0.8 (−3.4, 1.8)	−2.5* (−4.7, −0.3)
Incident NS	7.4*** (5.2, 9.7)	–	1.6 (−0.9, 4.2)
Sickle cell	4.3** (1.7, 7.0)	0.4 (−2.6, 3.3)	−0.3 (−3.3, 2.7)
Model R^2	0.22	0.18	0.07
Observations	744	446	737
Log likelihood	−2632.5	−1568.5	−2671.4
AIC	5292.9	3162.9	5370.8
BIC	5357.5	3216.2	5435.2

AIC Akaike information criteria, BIC Bayesian information criteria

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

^a The intercept can be interpreted as the mean function score at T2 (event) for the reference group who are non-Hispanic white, 8-years old boys with cancer, no other health condition and a mother with education more than high school. The domain of upper extremity was not assessed in children with nephrotic syndrome

Relationships ($b = 1.6$, $p < 0.05$). From T2 to T4 (follow-up), the function scores improved ($p < 0.05$). The physical function domains exceeded the MID of 3 points going from T2 to T3 and from T2 to T4.

Discussion

This multi-site, multi-disease longitudinal study provided evidence for the responsiveness of the PROMIS pediatric measures of fatigue, pain interference, anxiety, anger, depressive symptoms, mobility, and upper extremity. Using mixed modeling methods, we were able to combine data from children and adolescents with cancer, nephrotic syndrome, and sickle cell disease and examine how PROMIS scores in a more stable health state compared with scores when the child was experiencing a deteriorating health event, including chemotherapy (cancer), disease activity (nephrotic syndrome), and pain exacerbation (sickle cell disease). The magnitude of change from the event to the recovering phase exceeded the minimally important difference of 3 points [18] for all domains expected to change. These findings are consistent with our hypotheses about expected changes in symptoms and function in the 3 groups. Contrary to hypotheses, peer relationships did significantly get worse from T1 to T2 and

got better from T2 to T1, but the magnitude of change was below the MID and was the least affected compared to the other domains. It could be the collective decline of physical functioning and increased symptom burden experienced by children at T2 slightly decreased children's reported peer relationships.

The findings from this study add to the validity evidence from other studies that have examined the responsiveness of the PROMIS Pediatric measures. In an online cohort of 276 children with Crohn's disease (ages 9–17 years), children completed self-report measures at baseline and 6 months later including a measure of Crohn's disease activity and PROMIS measures of pain interference, fatigue, anxiety, depressive symptoms, and peer relationships [25]. Children with improved Crohn's disease activity from baseline to follow-up reported improved scores (larger than the established MID) on all PROMIS pediatric measures, and children with worse Crohn's disease activity from baseline to follow-up reported worse scores (larger than the MID) for all domains except Anxiety. In another study, 229 children (ages 8–17 years) from public insurance programs with asthma completed PROMIS pediatric measures of pain interference, fatigue, depressive symptoms, mobility, and peer relationships and measures of asthma control across four time points over 2 years [26]. The study found that children with worsened asthma control and poorer

overall health tended to report deteriorated function and more symptom burden on the PROMIS pediatric measures, with fatigue showing the greatest change.

The ability of the PROMIS pediatric fatigue measure to capture the largest changes compared to the other symptom measures is especially meaningful in children. Fatigue has been reported as the most troubling symptom during and following the recovery period by children experiencing a number chronic conditions including cancer, anemia, and surviving organ transplants. This means that the ability of the PROMIS pediatric fatigue measure to capture change is highly relevant to a number of illness groups.

This study had limitations, including limited locality as the cancer data were collected only from a single site and the SCD data were collected from three sites in the Atlanta area. Limited locality may raise concerns about the generalizability of findings. However, we do not have reason to believe that children in these localities with these conditions would vary from others in their response to these questionnaire items. The study was conducted in English language only. Lastly, the assessments did not include the collection of specific occurrence of other life events or stressors which could have had an impact on the PROMIS results beyond the influence of the disease under study.

Conclusions

The PROMIS Pediatric measures, as completed by 8–17 year olds experiencing one or more chronic conditions, are able to measure symptom and functional impact for the affected children and adolescents at specified time points and capture clinically meaningful change in health conditions as hypothesized. This means that these measures are able to quantify the impact of disease and treatment on a child or adolescent and further that these measures are now ready to be embedded into clinical trials for treatment of these diverse chronic illnesses.

The responsiveness of the PROMIS Pediatric measures has been documented here for three different pediatric chronic conditions and in the literature for two additional pediatric chronic conditions that vary in their clinical presentation and in their likely causative factors [25, 26]. In addition, there are ongoing efforts to further evaluate the responsiveness of the PROMIS pediatric measures in additional disease populations. The NIH-funded initiative, validation of pediatric patient-reported outcomes in chronic diseases (PEPR) Consortium [<http://grants.nih.gov/grants/guide/rfa-files/RFA-AR-15-014.html>] will examine how changes in PROMIS scores are associated with changes in disease status in populations of children with inflammatory bowel disease, cancer, juvenile idiopathic disease, systemic lupus erythematosus, and asthma. These studies will

include the PROMIS measures used in this study as well as newer PROMIS pediatric measures.

Acknowledgements PROMIS[®] was funded with cooperative agreements from the National Institutes of Health (NIH) Common Fund Initiative (Northwestern University, PI: David Cella, PhD, U54AR057951, U01AR052177; Northwestern University, PI: Richard C. Gershon, PhD, U54AR057943; American Institutes for Research, PI: Susan (San) D. Keller, PhD, U54AR057926; State University of New York, Stony Brook, PIs: Joan E. Broderick, PhD and Arthur A. Stone, PhD, U01AR057948, U01AR052170; University of Washington, Seattle, PIs: Heidi M. Crane, MD, MPH, Paul K. Crane, MD, MPH, and Donald L. Patrick, PhD, U01AR057954; University of Washington, Seattle, PI: Dagmar Amtmann, PhD, U01AR052171; University of North Carolina, Chapel Hill, PI: Harry A. Guess, MD, PhD (deceased), Darren A. DeWalt, MD, MPH, Bryce B. Reeve, PhD U01AR052181; Children's Hospital of Philadelphia, PI: Christopher B. Forrest, MD, PhD, U01AR057956; Stanford University, PI: James F. Fries, MD, U01AR052158; Boston University, PIs: Alan Jette, PT, PhD, Stephen M. Haley, PhD (deceased), and David Scott Tulsky, PhD (University of Michigan, Ann Arbor), U01AR057929; University of California, Los Angeles, PIs: Dinesh Khanna, MD (University of Michigan, Ann Arbor) and Brennan Spiegel, MD, MSHS, U01AR057936; University of Pittsburgh, PI: Paul A. Pilkonis, PhD, U01AR052155; Georgetown University, PIs: Carol M. Moinpour, PhD (Fred Hutchinson Cancer Research Center, Seattle) and Arnold L. Potosky, PhD, U01AR057971; Children's Hospital Medical Center, Cincinnati, PI: Esi M. Morgan DeWitt, MD, MSCE, U01AR057940; University of Maryland, Baltimore, PI: Lisa M. Shulman, MD, U01AR057967; and Duke University, PI: Kevin P. Weinfurt, PhD, U01AR052186). NIH Science Officers on this project have included Deborah Ader, PhD, Vanessa Ameen, MD (deceased), Susan Czajkowski, PhD, Basil Eldadah, MD, PhD, Lawrence Fine, MD, DrPH, Lawrence Fox, MD, PhD, Lynne Haverkos, MD, MPH, Thomas Hilton, PhD, Laura Lee Johnson, PhD, Michael Kozak, PhD, Peter Lyster, PhD, Donald Mattison, MD, Claudia Moy, PhD, Louis Quatrano, PhD, Bryce Reeve, PhD, William Riley, PhD, Peter Scheidt, MD, Ashley Wilder Smith, PhD, MPH, Susana Serrate-Sztejn, MD, William Phillip Tonkins, DrPH, Ellen Werner, PhD, Tisha Wiley, PhD, and James Witter, MD, PhD. The contents of this article uses data developed under PROMIS. These contents do not necessarily represent an endorsement by the US Federal Government or PROMIS. See www.HealthMeasures.net for additional information on the PROMIS[®] initiative.

Funding Funding for this study was provided by the National Institutes of Health (NIH) Common Fund Initiative: Grant #: U01AR052181. Grant #: U2CCA186878.

Compliance with ethical standards

Conflict of interest Dr. DeWalt was an author of some of the items in the PROMIS measures and owns the copyright for these items. Dr. DeWalt has given an unlimited free license for the use of the materials to the PROMIS Health Organization. Both Dr. DeWalt and Dr. Reeve are former members of the PROMIS Health Organization Board of Directors. There are no other conflicts of interest to report for other authors.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

- DeWitt, E. M., Stucky, B. D., Thissen, D., Irwin, D. E., Langer, M., Varni, J. W., et al. (2011). Construction of the eight-item patient-reported outcomes measurement information system pediatric physical function scales: Built using item response theory. *Journal of Clinical Epidemiology*, *64*(7), 794–804. doi:10.1016/j.jclinepi.2010.10.012.
- Varni, J. W., Stucky, B. D., Thissen, D., Dewitt, E. M., Irwin, D. E., Lai, J. S., et al. (2010). PROMIS Pediatric Pain Interference Scale: An item response theory analysis of the pediatric pain item bank. *The Journal of Pain*, *11*(11), 1109–1119. doi:10.1016/j.jpain.2010.02.005.
- Lai, J. S., Stucky, B. D., Thissen, D., Varni, J. W., DeWitt, E. M., Irwin, D. E., et al. (2013). Development and psychometric properties of the PROMIS((R)) pediatric fatigue item banks. *Quality of Life Research*, *22*(9), 2417–2427. doi:10.1007/s11136-013-0357-1.
- Irwin, D. E., Stucky, B., Langer, M. M., Thissen, D., Dewitt, E. M., Lai, J. S., et al. (2010). An item response analysis of the pediatric PROMIS anxiety and depressive symptoms scales. *Quality of Life Research*, *19*(4), 595–607. doi:10.1007/s11136-010-9619-3.
- Irwin, D. E., Stucky, B. D., Langer, M. M., Thissen, D., DeWitt, E. M., Lai, J. S., et al. (2012). PROMIS Pediatric Anger Scale: An item response theory analysis. *Quality of Life Research*, *21*(4), 697–706. doi:10.1007/s11136-011-9969-5.
- Dewalt, D. A., Thissen, D., Stucky, B. D., Langer, M. M., Morgan Dewitt, E., Irwin, D. E., et al. (2013). PROMIS Pediatric Peer Relationships Scale: Development of a peer relationships item bank as part of social health measurement. *Health Psychology*, *32*(10), 1093–1103. doi:10.1037/a0032670.
- DeWalt, D. A., Gross, H. E., Gipson, D. S., Selewski, D. T., DeWitt, E. M., Dampier, C. D., et al. (2015). PROMIS((R)) pediatric self-report scales distinguish subgroups of children within and across six common pediatric chronic health conditions. *Quality of Life Research*, *24*(9), 2195–2208. doi:10.1007/s11136-015-0953-3.
- Gipson, D. S., Selewski, D. T., Massengill, S. F., Wickman, L., Messer, K. L., Herreshoff, E., et al. (2013). Gaining the PROMIS perspective from children with nephrotic syndrome: A midwest pediatric nephrology consortium study. *Health and Quality of Life Outcomes*, *11*, 30. doi:10.1186/1477-7525-11-30.
- Hinds, P. S., Nuss, S. L., Ruccione, K. S., Withycombe, J. S., Jacobs, S., DeLuca, H., et al. (2013). PROMIS pediatric measures in pediatric oncology: Valid and clinically feasible indicators of patient-reported outcomes. *Pediatric Blood & Cancer*, *60*(3), 402–408. doi:10.1002/pbc.24233.
- Selewski, D. T., Collier, D. N., MacHardy, J., Gross, H. E., Pickens, E. M., Cooper, A. W., et al. (2013). Promising insights into the health related quality of life for children with severe obesity. *Health and Quality of Life Outcomes*, *11*, 29. doi:10.1186/1477-7525-11-29.
- Selewski, D. T., Massengill, S. F., Troost, J. P., Wickman, L., Messer, K. L., Herreshoff, E., et al. (2014). Gaining the patient reported outcomes measurement information system (PROMIS) perspective in chronic kidney disease: A midwest pediatric nephrology consortium study. *Pediatric Nephrology(Berlin, Germany)*, *29*(12), 2347–2356. doi:10.1007/s00467-014-2858-8.
- Menard, J. C., Hinds, P. S., Jacobs, S. S., Cranston, K., Wang, J., DeWalt, D. A., et al. (2014). Feasibility and acceptability of the PROMIS measures in children and adolescents in active cancer treatment and survivorship. *Cancer Nursing*, *37*(1), 66–74. doi:10.1097/NCC.0b013e3182a0e23d.
- Selewski, D. T., Troost, J. P., Massengill, S. F., Gbadegesin, R. A., Greenbaum, L. A., Shatat, I. F., et al. (2015). The impact of disease duration on quality of life in children with nephrotic syndrome: A midwest pediatric nephrology consortium study. *Pediatric Nephrology(Berlin, Germany)*, *30*(9), 1467–1476. doi:10.1007/s00467-015-3074-x.
- Dampier, C., Barry, V., Gross, H. E., Lui, Y., Thornburg, C. D., DeWalt, D. A., et al. (2016). Initial evaluation of the pediatric PROMIS(R) health domains in children and adolescents with sickle cell disease. *Pediatric Blood & Cancer*, *63*(6), 1031–1037. doi:10.1002/pbc.25944.
- Dampier, C., Jaeger, B., Gross, H. E., Barry, V., Edwards, L., Lui, Y., et al. (2016). Responsiveness of PROMIS(R) pediatric measures to hospitalizations for sickle pain and subsequent recovery. *Pediatric Blood & Cancer*, *63*(6), 1038–1045. doi:10.1002/pbc.25931.
- Bjorner, J. B., Chang, C. H., Thissen, D., & Reeve, B. B. (2007). Developing tailored instruments: Item banking and computerized adaptive assessment. *Quality of Life Research*, *16*(Suppl 1), 95–108. doi:10.1007/s11136-007-9168-6.
- Irwin, D. E., Stucky, B. D., Thissen, D., Dewitt, E. M., Lai, J. S., Yeatts, K., et al. (2010). Sampling plan and patient characteristics of the PROMIS pediatrics large-scale survey. *Quality of Life Research*, *19*(4), 585–594. doi:10.1007/s11136-010-9618-4.
- Thissen, D., Liu, Y., Magnus, B., Quinn, H., Gipson, D. S., Dampier, C., et al. (2016). Estimating minimally important difference (MID) in PROMIS pediatric measures using the scale-judgment method. *Quality of Life Research*, *25*(1), 13–23. doi:10.1007/s11136-015-1058-8.
- Revicki, D., Hays, R. D., Cella, D., & Sloan, J. (2008). Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *Journal of Clinical Epidemiology*, *61*(2), 102–109. doi:10.1016/j.jclinepi.2007.03.012.
- Bates, D., Machler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.
- Akaike, H. (1974). New look at statistical-model identification. *IEEE Transactions on Automatic Control*, *Ac19*(6), 716–723. doi:10.1109/Tac.1974.1100705.
- Schwarz, G. (1978). Estimating dimension of a model. *Annals of Statistics*, *6*(2), 461–464. doi:10.1214/aos/1176344136.
- Edwards, L. J., Muller, K. E., Wolfinger, R. D., Qaqish, B. F., & Schabenberger, O. (2008). An R2 statistic for fixed effects in the linear mixed model. *Statistics in Medicine*, *27*(29), 6137–6157. doi:10.1002/sim.3429.
- Jaeger, B. (2016). r2glmm: Computes R squared for mixed (multilevel) models (LMMs and GLMMs). (R package version 0.1.0. ed.).
- Arvanitis, M., DeWalt, D. A., Martin, C. F., Long, M. D., Chen, W., Jaeger, B., et al. (2016). Patient-reported outcomes measurement information system in children with Crohn's disease. *Journal of Pediatrics*, *174*(153–159), e152. doi:10.1016/j.jpeds.2016.03.069.
- Howell, C. R., Thompson, L. A., Gross, H. E., Reeve, B. B., DeWalt, D. A., & Huang, I. C. (2016). Responsiveness to change in PROMIS((R)) measures among children with asthma: A report from the PROMIS((R)) pediatric asthma study. *Value Health*, *19*(2), 192–201. doi:10.1016/j.jval.2015.12.004.