CrossMark

Determining a transitional scoring link between PROMIS® pediatric and adult physical health measures

David S. Tulsky[1,2] · Pamela A. Kisala[1] · Aaron J. Boulton[1] · Alan M. Jette[3] · David Thissen[4] · Pengsheng Ni[5] · Darren A. DeWalt[4] · I-Chan Huang[6,7] · Yang Liu[8] · M. J. Mulcahey[9] · Mary Slavin[5] · Brooke Magnus[10] · Holly Crump[4] · Robin Hanks[11] · Susan Charlifue[12] · Bryce B. Reeve[13]

Abstract

Purpose Having independent versions of the PROMIS® scales (for Pediatric and Adults) is problematic as scores cannot be evaluated longitudinally as individuals move from childhood into adulthood. The primary aim of this research project is to use item response theory (IRT) to develop a transitional scoring link (or "crosswalk") between the PROMIS adult and pediatric physical health measures.

Setting Sample 1 was collected at 6 rehabilitation hospitals in the U.S., and participants in Sample 2 were recruited from public health insurance programs and an online research panel.

Methods PROMIS pediatric and adult physical function, mobility, upper extremity, fatigue, and pain measures were administered to a sample of 874 individuals aged 14–20 years old with special health needs and a sample of 641 individuals aged 14–25 years with a disability. IRT-based scores were linked using a linear approximation to calibrated projection.

Results Estimated latent variable correlations ranged between 0.84 and 0.95 for the PROMIS pediatric and adult scores. Root Expected Mean Square Difference values were below the 0.08 threshold in all cases except when comparing genders on the Mobility (0.097) and Pain (0.10) scales in the special health care needs sample. Sum score conversion tables for the pediatric and adult PROMIS measures are presented.

Conclusions The linking coefficients can be used to calculate scale scores on PROMIS adult measures from pediatric measure scores and vice versa. This may lead to more accurate measurement in cross-sectional studies spanning multiple age groups or longitudinal studies that require comparable measurement across distinct developmental stages.

Keywords Patient-reported outcome measures · Psychometrics · Mobility limitation · Pain · Fatigue · Test equating · Test linking

✉ David S. Tulsky
dtulsky@udel.edu

1 Center for Health Assessment Research and Translation, University of Delaware, STAR Health Sciences Complex, 540 South College Ave, Newark, DE 10713, USA

2 Departments of Physical Therapy and Psychological and Brain Sciences, University of Delaware, Newark, DE, USA

3 MGH Institute of Health Professions, Boston, MA, USA

4 University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

5 Boston University School of Public Health, Boston, MA, USA

6 Department of Health Outcomes and Policy, College of Medicine, University of Florida, Gainesville, FL, USA

7 Department of Epidemiology and Cancer Control, St. Jude Children's Research Hospital, Memphis, TN, USA

8 University of Maryland, College Park, MD, USA

9 Department of Occupational Therapy, School of Health Professions, Thomas Jefferson University, Philadelphia, PA, USA

10 Marquette University, Milwaukee, WI, USA

11 Rehabilitation Institute of Michigan, Detroit, USA

12 Craig Hospital, Englewood, CO, USA

13 Duke University, Durham, NC, USA

Abbreviations

CHIP	Children's Health Insurance Program
CP	Cerebral palsy
EAP	Expected a Posteriori
HRQOL	Health-related quality of life
IRB	Institutional Review Board
IRT	Item response theory
LACP	Linear approximation to calibrated projection
OP4G	Opinions for good
PRO	Patient-Reported Outcomes
PROMIS	Patient-Reported Outcomes Measurement Information System®
REMSD	Root Expected Mean Square Difference
SCI	Spinal cord injury
SF	Short form
SHCN	Special health care needs
SMD	Standardized mean difference
TBI	Traumatic brain injury

Introduction

The transition to adulthood is a tumultuous period of biological, physical, social, and emotional changes [1]. In this period, adolescents consolidate their identity, achieve independence from parents, establish adult relationships outside the family, and find a vocation. For young people with disabilities, the transition period may be particularly difficult as they may be at a disadvantage due to their physical, cognitive, or psychosocial impairments, the extra health maintenance skills they need to acquire, lack of experience in activities and participation, social isolation, or by other environmental, family, and personal factors [2, 3]. Failure to make a successful transition to adulthood may result in unnecessary lifelong dependency, unemployment, lack of achievement, and poor quality of life [4].

Use of patient-reported outcomes (PRO) instruments in pediatric clinical medicine has increased in recent years but has yet to become standard practice as a way to assess physical, mental, and social outcomes [5]. Pediatric PRO instruments often have been developed in relative isolation from adult measures of similar constructs; the focus of pediatric PRO instrument development typically is on developing age-appropriate, contextually relevant items [5] rather than ensuring continuity of measurement throughout the lifespan. In cases where multiple versions of a measure are created for use with different age cohorts, it is common to have a parent proxy report, child, and adolescent versions of the same instrument (e.g., PedsQL) [6, 7] that contain different items and may measure slightly different aspects of the underlying construct. A major limitation to approaching pediatric PROs in this way is the inability to compare scores across the age groupings [5]. As the

child transitions into adult care and responds to adult PRO instruments, there may be no parallel "adult" measure and therefore no mechanism to compare scores from previous pediatric PRO instruments to those from adult PRO instruments. Lack of comparability renders it impossible to track changes in health outcomes across the lifespan for children aging with a disability. Currently, the comparison of outcomes between pediatric and adult PRO instruments is not possible nor is there a mechanism to monitor outcomes of children as they age through childhood and adolescence and into adulthood. These limitations are major barriers to evaluating and comparing treatment effectiveness and prognosticating long-term outcomes, especially for children with disabilities.

Starting in 2004, the NIH launched the Patient-Reported Outcomes Measurement Information System® (PROMIS®) initiative to design state-of-the-art PRO measurement instruments for a wide range of physical, mental, and social outcomes for children and adults [8, 9]. However, the overlap between pediatric and adult instruments was limited and the development and calibration of the pediatric and adult instruments were largely independent research activities. While the PROMIS pediatric investigators developed newly written items, they also reviewed the adult PROMIS item banks (and created items with content that was relevant to children) and other existing pediatric measures (e.g., PedsQL) to supplement their original item pools. The PROMIS items underwent extensive psychometric testing including a large quantitative study of 8000 children with a range of chronic conditions and children in the general population [10]. Meanwhile, the PROMIS Version 1.0 adult instruments were similarly developed and administered to a sample of over 20,000 adults and calibrated using graded response model IRT using a subsample of participants with demographic characteristics representative of the 2000 US Census [11]. In effect, the PROMIS pediatric and PROMIS adult versions have been evaluated separately and are discrete sets of instruments. Having two independent versions of the PROMIS instruments (for Pediatric and Adults) is problematic as child research participants cannot be followed longitudinally as they move from childhood into adulthood.

The purpose of this study, therefore, is to use IRT to develop a transitional scoring link between the PROMIS adult and pediatric physical health item banks so that studies that follow individuals through the child–adult transition can compare scores on the pediatric instrument with scores from the adult instrument as the study population ages into adulthood. In addition, any study in which both pediatric and adult forms were used would be able to test hypotheses on a common metric. For this investigation, we included item banks assessing physical *functioning* (i.e.,

mobility and upper extremity)[1] as well as physical *symptoms* (i.e., pain and fatigue). This manuscript describes the administration of PROMIS pediatric and adult physical health instruments to two independent samples of children and young adults and the subsequent use of IRT to produce score linking coefficients to transform PROMIS pediatric scores to their adult equivalency scores (and vice versa) for physical health domains. The linking coefficients for the PROMIS emotional health domains (i.e., adult and pediatric versions of PROMIS Depression, Anxiety, and Anger) have been reported in Reeve et al. [12].

Method

As part of the research consortium developing and validating PROMIS projects, two research studies (one at the University of North Carolina research site and another at the University of Michigan/Boston University site) performed linking studies of pediatric PROMIS instruments with the corresponding adult instruments. The samples, research design, and analytic methods have been described in a previously published study by Reeve et al. [12]. We will briefly review the methodological information below.

Research participants and data collection

Sample 1: Individuals with physical or cognitive disabilities

In the first sample, 188 adolescents (14–17 years old) and 453 young adults (18–25 years old) living with physical and/or cognitive disabilities due to spinal cord injury (SCI), traumatic brain injury (TBI), or cerebral palsy (CP) were recruited into the study at six participating sites: the University of Michigan, Boston University, Craig Hospital (Colorado), Rehabilitation Hospital of Michigan, and the Shriners Hospitals for Children (Philadelphia, Chicago). The University of Michigan served as the primary coordinating institution, and local site personnel were responsible for recruitment. Institutional Review Board (IRB) approval was obtained from all participating sites. Individuals were eligible to participate if (a) a confirmed diagnosis of SCI, TBI, or CP, (b) ability to read and understand English, and (c) ability to respond to self-report scales (e.g., by speaking, using a communication board, or gesturing). Participants with non-traumatic SCI and uncomplicated mild TBI (i.e., a

---

[1] The mobility item bank measures the amount of difficulty a youth has performing tasks requiring primarily leg movement (e.g., walking up stairs, playing with other children). Conversely, the Upper Extremity item bank measures the amount of difficulty a youth has performing tasks requiring arm and hand movement (e.g., opening a jar, buttoning a shirt).

Glasgow Coma Scale score between 13 and 15 with no positive neuroimaging findings) were not eligible to participate. Prior to beginning the study, assent was obtained from adolescents and informed consent from their parents; informed consent was obtained from young adults.

Data were collected by trained interviewers between June 1, 2011 and April 10, 2012. The interview format allowed the research team to include individuals with higher levels of physical or neurocognitive impairment while maintaining a consistent assessment modality across all participants. Interviewers met with participants in person or via telephone and entered their responses to items directly into the Assessment Center[SM] data collection platform [13]. Response formats varied across items and measures; printed cards with item-specific response scales were provided by the interviewer to match each item when administered. Response format cards were sent to participants interviewed by phone, who were instructed by interviewers on which to use for each question.

Sample 2: Individuals with "special health care needs"

The second sample was comprised of adolescents ($n = 415$) and young adults ($n = 459$) living with health conditions that require specialized health services (e.g., hypertension, cancer, mental health conditions) [14] These individuals, similar to those recruited in the first sample, are prime candidates for HRQOL assessment given the potential of their condition(s) to influence quality of life domains. Adolescents and young adults were recruited from public health insurance programs (Medicaid and Children's Health Insurance Program [CHIP] in Florida) and the Opinions for Good (Op4G) panel, a research company that maintains an online participant pool and asks participants to donate a portion of their proceeds to charitable organizations. Participants were eligible for participation if they were identified with special health care needs (SHCN; defined by the Clinical Risk Groups [15] in the Medicaid/CHIP sample and by the Special Care Needs Screener [16] in the Op4G sample), were 14–20 years of age, able to read, write, and speak English, and able to access an internet-enabled computer. The University of North Carolina at Chapel Hill served as the coordinating site and the University of Florida was responsible for data collection. The study protocol was approved by the IRB at each institution. Data were collected between April 1, 2012 and September 30, 2013. Assent was obtained from adolescent participants and informed consent was obtained from young adult participants and parents of adolescent participants.

Measures

Demographic information (e.g., age, sex, race, ethnicity, education level) was obtained for all participants. Additionally,

Table 1 Pediatric and adult PROMIS® physical function measures and example items

Domain	PROMIS® Pediatric Short Form(s), v1.0	PROMIS® Adult Short Form(s), v1.0
Physical function	**Mobility 8a**[a]	**Physical function 10a**[b,c]
	"I could walk up stairs without holding on to anything"	"Does your health now limit you in climbing one flight of stairs?"
	"I could keep up when I played with other kids"	"Does your health now limit you in walking more than a mile?"
	Upper extremity 8a[a]	
	"I could open a jar by myself"	"Are you able to shampoo your hair?"
	"I could button my shirt or pants"	"Are you able to dress yourself, including tying shoelaces and buttoning your clothes?"
Fatigue	**Fatigue 10a**[d]	**Fatigue 8a**[e,f]
	"I got tired easily"	"I feel fatigued"
	"I had trouble finishing things because I was too tired"	"How often did you have trouble finishing things because of your fatigue?"
Pain	**Pain Interference 8a**[d]	**Pain Interference 8a**[e]
	"I had trouble sleeping when I had pain"	"How much did pain interfere with your day to day activities?"
	"It was hard to have fun when I had pain"	"How much did pain interfere with the things you usually do for fun?"

Adult Physical Function items have no context/timeframe. All items on all other forms listed here begin with the context/timeframe "In the past 7 days…" or "During the past 7 days…". Additional information on PROMIS measures may be found at http://www.healthmeasures.net

Response options: [a] 1 = not able to do, 2 = with a lot of trouble, 3 = with some trouble, 4 = with a little trouble, 5 = with no trouble

[b] 1 = not at all, 2 = very little, 3 = somewhat, 4 = quite a lot, 5 = cancot do

[c] 1 = without any difficulty, 2 = with a little difficulty, 3 = with some difficulty, 4 = with much difficulty, 5 = unaule to do

[d] 1 = never, 2 = almost never, 3 = sometimes, 4 = often, 5 = almost always

[e] 1 = not at all, 2 = a little bit, 3 = somewhat, 4 = quite a bit, 5 = very much

[f] 1 = never, 2 = rarely, 3 = sometimes, 4 = often, 5 = always

participants with disabilities were asked to provide information on their methods of mobility (e.g., wheelchair use), and secondary medical complications (e.g., neurogenic bowel/bladder). Participants in the SHCN sample provided additional information related to their health condition(s). All participants completed the PROMIS pediatric and adult short forms for the following domains: Pain Interference, Fatigue, Peer Relationships, Depression, Anxiety, and Anger. Participants also completed the adult Physical Functioning and pediatric mobility and pediatric upper extremity scales because there are two pediatric short forms based on two physical functioning sub-domains (Mobility and Upper Extremity), whereas there is only a single adult short form for Physical Function. All items utilize a 5-category Likert-type format; higher item responses reflect greater functional ability on the Physical Function short forms and more severe symptoms on the Pain Interference and Fatigue short forms. The measures examined in this study are listed in Table 1 along with example items and associated response options. Measures were administered in random order to minimize the likelihood of order effects. Expected a Posteriori (EAP)[2]

scores for the item response patterns were calculated for each measure and transformed to the standard PROMIS T-score metric ($M = 50$, SD = 10) [10].

Analysis

Analyses were conducted on each sample separately and results were compared. Descriptive statistics and correlations between T-scores on the pediatric and adult measures were used initially to indicate which type of linking was most applicable. In educational testing, a correlation of 0.866 or higher between measures has been proposed as a prerequisite of unidimensional linking [17], which benchmarks a 50% or greater reduction in the uncertainty that results from predicting one measure from another (i.e., the scale score standard deviation). Although we employ this criterion in the current study so that estimations can be made on an individual level, it is important to state that a less stringent criterion (e.g., $r = 0.75$–0.80) is likely acceptable for lower stakes settings where decisions are not made on individual people (e.g., health outcomes research at the group level) [17, 18].

The Root Expected Mean Square Difference (REMSD) was calculated to further evaluate the tenability of linking. The REMSD statistic reflects the degree to which linking

---

[2] Expected a Posteriori (EAP) scores are IRT-based scale score estimates obtained by computing the mean of each individual's latent trait distribution, which is determined by each individual's response pattern and the IRT model employed.

is invariant for important subgroups within a population. Given two subgroups (e.g., males and females, adolescents and young adults), the REMSD can be calculated by computing the standardized mean difference (SMD) between the subgroups on one measure and subtracting it from the SMD computed between the subgroups on the second measure. Subgroup invariance is achieved when the subgroups differ by about the same amount—and in the same direction—on each measure; that is, the difference between SMDs for the two measures is close to zero. In the present study, subgroup invariance was evaluated between males and females as well as between adolescents and young adults. Using data from college admissions (i.e., high-stakes) testing, Dorans and Holland [19] concluded that REMSD values below 0.08 were generally supportive of subgroup invariance.

Disattenuated correlations were also calculated via the estimation of two-dimensional IRT models with pediatric items loading onto the first latent variable and adult items loading onto the second latent variable. Specifically, two-dimensional graded-response models were fit in the IRT-PRO [20] software using maximum likelihood estimation. As Reeve et al. [12] found with the PROMIS emotional distress measures, the pediatric and adult measures were not expected to be perfectly correlated, even after obtaining disattenuated estimates. If correlations were trivially different from 1, we planned to proceed with symmetrical (e.g., scale alignment) linking methods. However, we anticipated that the PROMIS pediatric and adult physical health measures would be highly correlated yet below the level at which symmetrical linking is justified. We therefore planned to proceed with an asymmetrical linking method called *calibrated projection* which relies on predicting one measure from another, and vice versa, in order to unify metrics. Calibrated projection [21] is a newly developed linking method that does not require the assumption that underlying constructs to be linked are the same, as is the case with symmetrical linking procedures. Calibrated projection does not require values on the predictor variable to be fixed; score distributions (as opposed to point estimates) are projected from one scale to another. Thissen et al. [22] simplified this technique by introducing a linear approximation that is computationally simpler and makes explicit the use of regression in the linking predictions. Therefore, in this article we used the *linear approximation to calibrated projection* (LACP) method described by Thissen et al. [22]. We briefly review this procedure here and direct interested readers to Thissen et al. [22] and Reeve et al. [12] for further information.

To proceed with the LACP given two scales, we first let θ_1 represent an arbitrary score on the latent construct underlying the first scale and let θ_2 represent an arbitrary score on the latent construct underlying the second scale. Given θ_1, the goal of calibrated projection is to predict an associated value θ_2 on the metric of the second scale, and vice versa when given θ_2 to predict θ_1. The prediction is accomplished via the linear regression:

$$\widehat{\mathrm{EAP}}\left[\theta_2\right] = \beta_0 + \beta_1 \mathrm{EAP}\left[\theta_1\right]. \qquad (1)$$

In Eq. 1, the regression coefficients β_0 and β_1 (intercept and slope) are hereafter referred to as 'linking' coefficients. Because the linkage is asymmetrical, the coefficients used to predict $\widehat{\mathrm{EAP}}\left[\theta_1\right]$ will differ from those used to predict $\widehat{\mathrm{EAP}}\left[\theta_2\right]$. The prediction in Eq. 1 is computed implicitly using IRT methods following the estimation of a multidimensional IRT (MIRT) model. In this example, a two-dimensional model is used such that items from the first scale load onto the first latent variable, and items from the second scale load onto the second latent variable. The prediction in Eq. 1 can be carried out by explicitly calculating the linking coefficients using estimates of the mean and covariance matrix derived from the MIRT model, which is the approach used in the LACP method. The standard deviations of the projected scores, which are taken as estimates of the projected score standard errors (and thus used to compute confidence intervals around individual scores), are automatically produced using IRT methods in the calibrated projection approach originally proposed [21]. Subsequently, Thissen et al. [22] suggested the SDs can be approximated using the formula:

$$\widehat{\mathrm{SD}}\left[\theta_2\right] = \sqrt{\beta_1^2 \mathrm{SD}^2\left[\theta_1\right] + \mathrm{MSE}}, \qquad (2)$$

where MSE is the mean squared error computed from Eq. 1 and $\mathrm{SD}^2[\theta_1]$ is the variance of the posterior distribution for observed θ_1. It can be seen in Eq. 2 that the approximation linearly combines two sources of error: the error variance associated with the observed θ_1 estimate, and the error variance associated with the projection procedure (Eq. 1).

The scales used in this study were based on published item banks calibrated in independent samples from those reported here. To preserve the metric of the original scales, parameters in the MIRT model were fixed to those estimated in the original calibrations. The resulting mean and covariance matrix estimates for θ_1 and θ_2 were then estimated with the current data and used to calculate the coefficients for each linkage [22]. A summary of the LACP procedure as used in this study is provided in Fig. 1.

Linkages were evaluated by comparing participants' projected scores on a given scale with their actual scores on the same scale (made possible by the fact that in both samples participants completed all PROMIS pediatric and adult measures). Comparisons were made in terms of confidence interval coverage; specifically, linkages were considered efficacious if approximately 68% of the observed values were within 1 SD of the associated projected values and approximately 95% of

Summary of LACP Procedure for Two Measures

1. Estimate two-dimensional MIRT model to all items from both measures. Parameters (discrimination, thresholds) are fixed to PROMIS calibration estimates to maintain metric.

2. Use estimated mean and covariance matrices from Step 1 to compute linking coefficients using linear approximation (see Thissen et al. [22] for computational details).

3. Compute SDs of projected values using linear approximation, which represent standard error estimates for individual projected scores.

Ex. Linking PROMIS Pediatric: Pain (P) to PROMIS Adult: Pain (A)

1.
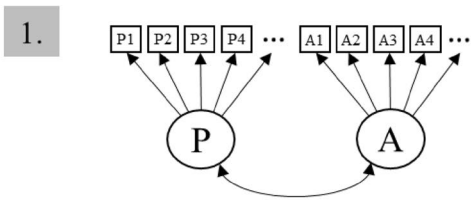
2.

$\widehat{\Sigma}$ = covariance matrix from step 1

$\widehat{\mu}$ = mean vector from step 1

$$\widehat{EAP}[\theta_P] = \boldsymbol{\beta_{0A>P}} + \boldsymbol{\beta_{1A>P}}EAP[\theta_A]$$
$$\widehat{EAP}[\theta_A] = \boldsymbol{\beta_{0P>A}} + \boldsymbol{\beta_{1P>A}}EAP[\theta_P]$$

3.

$$\widehat{SD}[\theta_P] = \sqrt{\boldsymbol{\beta_{1A>P}}SD^2[\theta_A] + MSE}$$
$$\widehat{SD}[\theta_A] = \sqrt{\boldsymbol{\beta_{1P>A}}SD^2[\theta_P] + MSE}$$

Fig. 1 Summary of Linking Procedure (Linear Approximation to Calibrated Projection). The diagram in step 1 of the example (Ex.) represents a path diagram—typically reserved for describing factor analytic or structural equation models—of the two-dimensional IRT model. In this diagram, circles represent the two latent variables underlying the set of combined items, which are represented by boxes. Terms in boldface in steps 2 and 3 of the example represent linking coefficients. Subscripts reflect whether the term is an intercept (0) or slope (1) as well as whether it is used to link the Pediatric form to the Adult form (P > A) or vice versa (A > P)

the observed values were within 2 SDs of the associated projected values. These criteria follow from the fact that the SDs were used as standard error estimates for the projected scores, and that scores are assumed to be normally distributed with repeated sampling. Thus, one would expect that $100(1 - \alpha)\%$ of the time the "true" scores (i.e., participants' actual scores on the measure) would reside within the projected scores' confidence intervals for a chosen width.

Results

Table 2 displays the demographic characteristics for the two samples with values displayed separately for the adolescent and young adult participants. Males made up a greater share of the sample of individuals with disability (62.9%), reflecting population estimates that show that SCI and TBI are more common in males than females [23,

24]. The sample of individuals with disability was largely white (80.5%) and non-Hispanic (84.2%), with 37.8% of participants diagnosed with SCI, 31.5% with TBI, and 30.7% with CP.

Descriptive statistics for the PROMIS pediatric and adult measures are shown in Table 3. Participants with a disability reported lower T-score averages on all measures. Correlations between the pediatric and adult versions in each sample overall ranged between 0.77 and 0.90 (Table 4; correlations are also provided for demographic subgroups); only one of the measures (Mobility, in the sample with a disability) exceeded Dorans' [17] criterion of 0.866 for unidimensional equating. Table 5 displays REMSD statistics, estimated latent variable correlations (ρ) and associated correlation SEs. REMSD values were below the 0.08 threshold in all cases except when comparing genders on the Mobility (0.097) and Pain (0.170) scales in the SHCN sample. Although exceeding the 0.08 upper limit commonly

Table 2 Demographic and clinical characteristics of the study samples

	Sample with a Disability		Sample with Special Health Care Needs	
	Adolescents (14–17 years) N = 188	Young adults (18–24 years) N = 453	Adolescents (14–17 years) N = 415	Young adults (18–20 years) N = 459
Age (mean, (SD))	15.6 (1.26)	21.43 (2.21)	15.63 (1.20)	18.93 (0.75)
Sex				
Male	112 (59.6%)	291 (64.2%)	214 (51.6%)	194 (42.3%)
Female	75 (39.9%)	161 (35.5%)	200 (48.2%)	265 (57.7%)
Missing	1 (0.5%)	1 (0.2%)	1 (0.2%)	0
Ethnicity				
Hispanic	13 (6.9%)	40 (8.8%)	138 (33.3%)	196 (42.7%)
Non-hispanic	153 (81.4%)	366 (80.8%)	277 (66.7%)	263 (57.3%)
Not provided	22 (11.7%)	47 (10.4%)	0	0
Race				
White	150 (79.8%)	366 (80.8%)	216 (52.0%)	228 (49.7%)
Black or African American	20 (10.6%)	44 (9.7%)	93 (22.4%)	89 (19.4%)
Asian	5 (2.7%)	9 (2.0%)	34 (8.2%)	40 (8.7%)
Other	11 (5.8%)	20 (4.4%)	51 (12.3%)	63 (13.7%)
Multiple races	2 (1.0%)	4 (0.8%)	14 (3.4%)	19 (4.1%)
Missing	0	8 (1.7%)	7 (1.7%)	20 (4.4%)
Health condition				
Cerebral palsy	99 (52.7%)	98 (21.6%)	6 (1.4%)	4 (0.9%)
Hemiplegia	36 (36.4%)	23 (23.5%)	nr	nr
Diplegia	52 (52.5%)	44 (44.9%)	nr	nr
Quadriparesis	10 (10.1%)	28 (28.6%)	nr	nr
Missing	1 (1.0%)	3 (3.1%)	nr	nr
Spinal cord injury	53 (28.2%)	189 (41.7%)	–	–
Paraplegia—incomplete	11 (20.8%)	33 (17.5%)	–	–
Paraplegia—complete	21 (39.6%)	63 (33.3%)	–	–
Tetraplegia—incomplete	7 (13.2%)	46 (24.3%)	–	–
Tetraplegia—complete	12 (22.6%)	45 (23.8%)	–	–
Missing	2 (3.8%)	2 (1.1%)	–	–
Traumatic brain injury	36 (19.1%)	166 (36.6%)	–	–
Complicated mild	14 (38.9%)	28 (16.9)	–	–
Moderate	4 (11.1%)	16 (9.6%)	–	–
Severe	14 (38.9%)	106 (63.9%)	–	–
Missing	4 (11.1%)	16 (9.6%)	–	–
ADHD	–	–	138 (33.3%)	108 (23.5%)
Allergies	–	–	98 (23.6%)	81 (17.6%)
Asthma	–	–	93 (22.4%)	104 (22.7%)
Blind	–	–	7 (1.7%)	6 (1.3%)
Born prematurely	–	–	18 (4.3%)	20 (4.4%)
Cancer	–	–	8 (1.9%)	18 (3.9%)
Chronic pain	–	–	88 (21.2%)	103 (22.4%)
Deaf	–	–	10 (2.4%)	8 (1.7%)
Diabetes	–	–	37 (8.9%)	52 (11.3%)
Epilepsy	–	–	15 (3.6%)	17 (3.7%)
Heart disease	–	–	11 (2.7%)	6 (1.3%)
Hypertension	–	–	81 (19.5%)	118 (25.7%)
Intestinal disease	–	–	19 (4.6%)	17 (3.7%)
Kidney disease	–	–	17 (4.1%)	15 (3.3%)

Table 2 (continued)

	Sample with a Disability		Sample with Special Health Care Needs	
	Adolescents (14–17 years) $N=188$	Young adults (18–24 years) $N=453$	Adolescents (14–17 years) $N=415$	Young adults (18–20 years) $N=459$
Mental health	–	–	104 (25.1%)	95 (20.7%)
Needs walking assist	–	–	7 (1.7%)	8 (1.7%)
Overweight	–	–	67 (16.1%)	85 (18.5%)
Rheumatic disease	–	–	12 (2.9%)	13 (2.8%)
Sickle cell disease	–	–	7 (1.7%)	8 (1.7%)
Thyroid disease	–	–	14 (3.4%)	18 (3.9%)

nr = not reported

Table 3 Descriptive statistics and correlations between PROMIS® Pediatric and adult measures

Domain	Sample	PROMIS® Pediatric				PROMIS® Adult			
		Mean	SD	Min	Max	Mean	SD	Min	Max
Mobility[a]	Disabilities[b]	38.9	14.2	14.2	58.5	43.5	13.3	13.9	61.7
	Adolescents	42.0	13.3	14.2	58.5	46.1	12.3	13.9	61.7
	Young adults	37.6	14.3	14.2	58.5	42.4	13.5	13.9	61.7
	SHCN[c]	44.1	11.3	14.2	58.5	46.8	10.4	13.9	61.7
	Adolescents	44.8	11.3	14.2	58.5	47.5	10.6	13.9	61.7
	Young adults	43.5	11.2	14.2	58.5	46.3	10.2	13.9	61.7
Upper extremity[a]	Disabilities	44.5	14.2	9.65	56.7	43.5	13.3	13.9	61.7
	Adolescents	45.9	12.7	9.65	56.7	46.1	12.3	13.9	61.7
	Young adults	44.0	14.7	9.65	56.7	42.4	13.5	13.9	61.7
	SHCN	44.9	12.8	9.6	56.7	46.8	10.4	13.9	61.7
	Adolescents	45.2	12.9	9.6	56.7	47.4	10.6	13.9	61.7
	Young adults	44.6	12.7	9.6	56.7	46.3	10.2	13.9	61.7
Fatigue	Disabilities	46.4	11.5	30.2	85.1	48.2	8.9	33.1	77.8
	Adolescents	44.5	10.1	30.2	69.4	46.5	8.9	33.1	77.8
	Young adults	47.2	12.0	30.2	85.1	49.0	8.8	33.1	77.8
	SHCN	54.7	13.5	30.2	85.1	52.5	10.6	33.1	77.8
	Adolescents	53.8	13.6	30.2	85.1	51.7	10.9	33.1	77.8
	Young adults	55.6	13.3	30.2	85.1	53.2	10.3	33.1	77.8
Pain	Disabilities	45.5	10.8	34.0	77.2	48.8	8.7	40.7	77.0
	Adolescents	43.3	9.9	34.0	72.5	46.6	7.5	40.7	77.0
	Young adults	46.3	11.0	34.0	77.2	49.7	9.0	40.7	77.0
	SHCN	52.0	11.5	34.0	78.2	54.4	9.9	40.7	77.0
	Adolescents	51.7	11.3	34.0	78.2	53.8	9.8	40.7	77.0
	Young adults	52.4	11.7	34.0	78.2	54.9	9.9	40.7	77.0

Age categorized by adolescent (14–17 years) and young adult (18–24 years)

[a] Adult version is PROMIS® Physical Function short form, same scores tabulated for Mobility and Upper Extremity

[b] Disabilities = adolescents and young adults with physical or cognitive disability

[c] SHCN = adolescents and young adults with special health care needs

used to establish subgroup invariance, these REMSD values were not replicated in the sample with a disability. Finally, estimated latent variable correlations (i.e., disattenuated correlations estimated from the 2-D IRT models) ranged between 0.84 and 0.95 for the PROMIS pediatric and adult scales. These values are large, suitable for health outcomes research, and most exceed the very conservative criterion of 0.866 that was recommended for high-stakes educational

Table 4 Correlations between pediatric and adult PROMIS forms by age, sex, disability type, special health care needs type, education

Factor	Level	Mobility[a]	Upper extremity[a]	Fatigue	Pain
Disabilities Sample[b]					
Overall		0.90	0.81	0.77	0.79
Age	Adolescent	0.86	0.78	0.72	0.69
	Young adult	0.91	0.83	0.79	0.82
Gender	Female	0.90	0.83	0.75	0.78
	Male	0.90	0.80	0.79	0.81
Disability type	CP	0.88	0.84	0.71	0.71
	SCI	0.78	0.87	0.78	0.82
	TBI	0.80	0.64	0.83	0.82
SHCN Sample[c]					
Overall		0.85	0.78	0.86	0.79
Age	Adolescent	0.86	0.82	0.85	0.76
	Young adult	0.84	0.74	0.86	0.81
Gender	Female	0.85	0.77	0.84	0.76
	Male	0.85	0.78	0.88	0.81
SHCN type[d]	ADHD	0.84	0.80	0.86	0.80
	Allergies	0.82	0.75	0.86	0.78
	Asthma	0.79	0.72	0.84	0.73
	Chronic pain	0.73	0.66	0.84	0.76
	Hypertension	0.82	0.75	0.85	0.72
	Mental health	0.81	0.74	0.84	0.82

Age categorized by adolescent (14–17 years) and young adult (18–24 years)

CP = cerebral palsy; SCI = spinal cord injury; TBI = traumatic brain injury

[a]Adult version is PROMIS® Physical Function short form

[b]Disabilities = adolescents and young adults with physical or cognitive disability

[c]SHCN = adolescents and young adults with special health care needs

[d]SHCN types are not mutually exclusive

Table 5 Root expected mean square difference (REMSD) by age and sex and estimated latent variable correlation ρ between the pediatric and adult constructs

Domain	Sample	REMSD Sex	REMSD Age	ρ	SE (ρ)
Mobility[a]	Disabilities[b]	0.015	0.015	0.95	0.01
	SHCN[c]	0.097	0.060	0.95	0.01
Upper extremity[a]	Disabilities	0.032	0.062	0.92	0.01
	SHCN	0.066	0.019	0.92	0.01
Fatigue	Disabilities	0.001	0.020	0.84	0.02
	SHCN	0.010	0.007	0.91	0.01
Pain	Disabilities	0.053	0.035	0.88	0.03
	SHCN	0.170	0.031	0.86	0.01

Age categorized by adolescent (14–17 years) and young adult (18–24 years)

[a]Adult version is PROMIS® Physical Function short form

[b]Disabilities = adolescents and young adults with physical or cognitive disability

[c]SHCN = adolescents and young adults with special health care needs

method were between 0.9 and 1.8 times those derived from calibrated projection. Therefore, the LACP worked well and the intercept and slope coefficients are presented in Table 6. Associated 95% confidence intervals as well as the estimated MSE from the projections are provided. Comparing the disabilities and SHCN samples, the regression slopes are quite similar. Across the entire range of observed scale scores, the regression lines never differed by more than one-third of a standard deviation (i.e., 3 T-score points). Given the apparent comparability, the linking coefficients were averaged across the two samples.

In Table 7, the proportion of observed scores falling within ± 1 SD and ± 2 SDs of the projected scores are reported separately for the two samples. Estimates are provided both for the pediatric-to-adult and adult-to-pediatric linkages. Proportions based on same-sample coefficients (see Table 7 footnote) are provided in the first row block and reflect model fit. Proportions based on cross-sample coefficients are provided in the second row block and reflect cross-sample validation. Finally, proportions based on the averages of the linking coefficients across samples are presented in the third row block. Proportions were close to the confidence interval widths chosen (68% and 95%) with the exception of the Pediatric Upper Extremity—Adult Physical Functioning linkage. For this linkage and in both directions, proportions of observed values within 1 SD of projected values were well below 68% in most cases, with values ranging between 0.37 and 0.59. This is due to the distribution of scores for those scales: Over 20% of the respondents have perfect (maximum) scores on both scales, which produces a single point mass in the distributions with over 20% of the data. The fact that these large blocks, and adjacent nearly

testing. However, statistical comparisons between unidimensional and multi-dimensional models that contained pediatric and adult items in the same model (e.g., Pediatric PROMIS Fatigue and Adult PROMIS Fatigue items) revealed that multi-dimensional models provided a better fit to the data. These results suggest that symmetrical equating methods, such as those based on unidimensional IRT models, are inappropriate. However, asymmetrical linking methods, such as the LACP method used in this study, were considered applicable.

LACP results are presented in Table 6. To evaluate the precision of the approximation, sum score conversion tables of the PROMIS pediatric and adult measures were created using calibrated projection and approximate calibrated projection [12, 22]. Across the entire range of sum scores, the two methods produced scores that are virtually identical. Standard deviations created under the approximation

Table 6 Regression coefficients β_0 and β_1, and MSE for the calibrated projection from θ_1 to θ_2

Domain	Sample	Pediatric to adult			Adult to pediatric		
		β_0	β_1	MSE	β_0	β_1	MSE
Mobility	Disabilities	11.99 (9.46–14.51)	0.84 (0.76–0.91)	23.3	−9.00 (−12.03 to −5.96)	1.07 (0.99–1.15)	29.7
	SHCN	10.60 (9.03–12.17)	0.83 (0.79–0.87)	13.1	−6.91 (−8.95 to −4.87)	1.08 (1.03–1.13)	17.1
	Average	**11.29**	**0.83**	**18.2**	**−7.95**	**1.08**	**23.4**
Upper extremity	Disabilities	12.37 (10.53–14.22)	0.67 (0.62–0.71)	37.6	−7.98 (−11.16 to −4.81)	1.26 (1.18–1.34)	71.0
	SHCN	21.05 (19.76–22.34)	0.54 (0.51–0.57)	19.3	−25.04 (−29.28 to −20.80)	1.55 (1.45–1.65)	55.3
	Average	**16.71**	**0.60**	**28.4**	**−16.72**	**1.40**	**63.2**
Fatigue	Disabilities	24.29 (22.41–26.17)	0.53 (0.49–0.57)	23.7	−19.51 (−24.19 to 14.84)	1.34 (1.24–1.43)	59.6
	SHCN	17.72 (15.51–19.92)	0.64 (0.60–0.68)	21.0	−13.93 (−17.27 to −10.59)	1.31 (1.25–1.37)	43.4
	Average	**21.00**	**0.58**	**22.4**	**−16.72**	**1.32**	**51.5**
Pain	Disabilities	23.40 (18.63–28.18)	0.59 (0.49–0.68)	24.0	−21.44 (−27.87 to −15.01)	1.32 (1.21–1.44)	54.1
	SHCN	21.28 (17.35–25.21)	0.65 (0.60–0.70)	26.3	−11.62 (−25.56 to −17.32)	1.16 (1.08–1.24)	47.2
	Average	**22.34**	**0.62**	**25.1**	**−16.53**	**1.24**	**50.7**

All scores are on the standard T-score scale used for PROMIS® instruments. Linear regression can be used to convert scores across measures. For instance, a score of 55 on the pediatric Mobility measure would convert to a score of 56.94 ($11.29 + 0.83 \times 55$) on the adult Physical Function scale

perfect-score blocks, have residuals between 1 and 2 SDs from the mean reduces the observed proportion within ±1 SD from the nominal 0.68 to 0.37–59. Excluding this linkage from consideration, proportions ranged between 0.61 and 0.75 for 1 SD and between 0.92 and 0.96 for 2 SD when the linking coefficients were averaged, which is considered acceptable.

Discussion

The PROMIS measurement system was designed to address several needs in PRO assessment [8]. Primary among these was the lack of a common set of standardized and validated instruments that could be used to measure important PROs in the general population and across a wide variety of populations with health conditions. While the initial and subsequent releases of PROMIS have made just such a set of measures available for both research and clinical use, at a more fundamental level, a common *longitudinal* metric of overall health and functioning for children and young adults was lacking. In the current study, we set out to create a bridge to connect individuals' scores across PROMIS pediatric and adult physical health measures. This work allows investigators who enroll children in longitudinal studies to follow them over years, switching from the pediatric version of the scale to the adult version. Clinicians will also be able to administer either the pediatric or adult PROMIS measures as appropriate and convert scores from the pediatric version to equivalent adult version scores (and vice versa) using the newly created linking coefficients. A computer application has

been developed to assist researchers and clinicians compute these linkages and is available at https://sites.udel.edu/chs-chart/.

We first assessed the viability of linking PROMIS pediatric measures of Fatigue, Pain, Upper Extremity Functioning, and Mobility to commensurate adult measures in two healthcare-specific samples. Although pediatric and adult measures within each domain were highly correlated, construct equivalence could not be established. However, the REMSD index of subgroup invariance suggested that asymmetrical linking was possible. Therefore, we used the recently developed LACP procedure [22] to connect the pediatric and adult measures by establishing unidirectional prediction functions that can be used to convert scale scores from one instrument to another. The linear approximation was close, as observed scores and LACP-predicted scores on the same measure were nearly identical. Additionally, except for the PROMIS Pediatric Upper Extremity–PROMIS Adult Physical Functioning linkage, the approximation did not substantially inflate score error estimates.

The regression coefficients produced in Table 6, which can be used to convert PROMIS pediatric scale scores to PROMIS adult scale scores or vice versa, were evaluated by calculating the proportion of individuals' observed scale scores falling within ±1 or 2 SDs of associated projected scores (Table 7). The results provided general support for the pediatric-to-adult and adult-to-pediatric linkages: the proportion of observed scores within 1 SD (2 SDs) of the predicted scores were close to 68% (95%) as expected. One exception to this finding, again, was the linkage between the PROMIS Pediatric Upper Extremity scale and the adult PROMIS Physical Functioning scale: in both directions, the

Table 7 Proportions of EAP[θ_2] values within ± 1 and ± 2 SD of the values obtained using linear approximation to calibrated projection

Domain	Disabilities sample				Special health care needs (SHCN) sample			
	Pediatric to adult		Adult to pediatric		Pediatric to adult		Adult to pediatric	
	± 1 SD	± 2 SD	± 1 SD	± 2 SD	± 1 SD	± 2 SD	± 1 SD	± 2 SD
	Disabilities parameters				SHCN parameters			
Mobility	0.70	0.93	0.67	0.93	0.71	0.93	0.69	0.91
Upper extremity	0.59	0.96	0.58	0.96	0.54	0.95	0.53	0.95
Fatigue	0.68	0.94	0.74	0.97	0.71	0.94	0.75	0.94
Pain	0.73	0.93	0.80	0.97	0.75	0.92	0.77	0.95
Domain	Disabilities sample				Special health care needs (SHCN) sample			
	Pediatric to adult		Adult to pediatric		Pediatric to adult		Adult to pediatric	
	± 1 SD	± 2 SD	± 1 SD	± 2 SD	± 1 SD	± 2 SD	± 1 SD	± 2 SD
	SHCN parameters				Disabilities parameters			
Mobility	0.62	0.89	0.57	0.88	0.72	0.95	0.77	0.95
Upper extremity	0.37	0.81	0.40	0.86	0.56	0.97	0.47	0.97
Fatigue	0.68	0.94	0.68	0.92	0.67	0.93	0.75	0.95
Pain	0.74	0.93	0.79	0.97	0.65	0.91	0.78	0.95
Domain	Disabilities sample				Special health care needs (SHCN) sample			
	Pediatric to adult		Adult to pediatric		Pediatric to adult		Adult to pediatric	
	± 1 SD	± 2 SD	± 1 SD	± 2 SD	± 1 SD	± 2 SD	± 1 SD	± 2 SD
	Average parameters				Average parameters			
Mobility	0.66	0.92	0.61	0.92	0.73	0.94	0.70	0.92
Upper extremity	0.49	0.93	0.53	0.92	0.59	0.96	0.52	0.97
Fatigue	0.72	0.95	0.73	0.96	0.72	0.94	0.75	0.94
Pain	0.73	0.93	0.79	0.97	0.70	0.92	0.77	0.95

Proportions provided in the first row block are based on same-sample comparisons (e.g., EAP estimates for disability sample vs. projected EAP estimates for disability sample *based on the disability sample linking coefficients*). Proportions in the second row block are based on cross-sample comparisons (e.g., EAP estimates for disability sample vs. projected EAP estimates for disability sample *based on SHCN linking coefficients*). Finally, proportions in the bottom row are based on the averaged linking parameters (e.g., EAP estimates for disability sample vs. projected EAP estimates for disability sample *based on linking coefficients averaged across the disability and SHCN samples*)

calculated proportions were below those expected based on the normal distribution. Notably, this linkage also results in the highest inflation of estimation error resulting from the linear approximation as described earlier. Both of these phenomena likely are due to the point mass in the distributions for Upper Extremity and Physical Function. While the linking coefficients in Table 6 to link the Pediatric Upper Extremity and Adult Physical Function measures may be useful in research settings when large data sets are compared and group-level inferences are of interest, it is not recommended to use these linking coefficients in any clinical or other high-stakes setting where decisions about an individual person are made. The linking coefficients do not provide the level of accuracy and the large prediction errors could result in inaccurate decisions for an individual.

As shown in Table 6, the MSE is smaller from pediatric to adult than adult to pediatric. The implication is that the

adult-to-pediatric predictions will have more linking error than the converse. Although the Upper Extremity-Physical Function linkage results in the highest MSE discrepancy and lowest distributional overlap, this is likely due to a distributional artifact as explained above.

Given that Upper Extremity and Mobility-specific PROMIS adult items have been identified subsequent to the conduct of this study, and that Hays et al. [25] have identified a subset of these adult items that have corresponding pediatric items, future work could use the LACP projection procedure to produce coefficients to directly link the PROMIS Pediatric and Adult Mobility measures and, separately, link the PROMIS Pediatric and Adult Upper Extremity measures. It would be ideal to include a substantial subsample of individuals with upper extremity limitations in this work.

A clear strength of this study was the use of two samples which (a) allowed for descriptive comparisons and explicit

cross-validation, and (b) represent distinct and important healthcare populations whose constituents are the primary beneficiaries of PRO research. Furthermore, LACP was able to produce successful linkages while being flexible enough to accommodate the special case of linking two pediatric physical functioning measures (Upper Extremity and Mobility) to the single adult physical functioning measure. However, future work including participants from a wider age range (< 14 and > 20) would add insight to the generalizability of the findings reported here.

Ideally, the same PRO measure could be used consistently over time to compare scores between individuals and, especially, within individuals over time. However, given the rapid pace of advances in PRO measurement and the diversity of developmentally appropriate constructs and items over the lifespan, it is necessary to employ PRO linking procedures to approximate the results that would be obtained through the perpetual administration of a static measure. The results of this study and those reported in Reeve et al. [12] represent a significant contribution to longitudinal research that assesses children across their lifespan and as such offers a significant contribution to PROMIS and the PROMIS linkages published in recent years [18]. However, estimating scores through linking procedures never will be as accurate as directly comparing scores on the same measure. The projection procedures applied here will always involve prediction error, but is nevertheless one possible solution when the same measurement tool cannot be utilized across the lifespan. Linking efforts appear to be increasing in popularity among PRO researchers in a variety of health-related fields and, as a consequence, the field would benefit from more research on linking methods in the context of health outcomes assessment. Methodologies currently used (including those reported in this article) originally were designed for measure linking in educational and high-stakes testing settings and may need further development in the healthcare context [18]. For instance, one limitation of the current study is the use of the normal distribution to compare actual vs. predicted scores, which may be less informative when substantial floor/ceiling effects are present (as is common in health outcomes measures).

Conclusion

The results presented herein are part of a larger collective effort aiming to establish a common metric for PROs regardless of the measurement instrument administered or the characteristics of the person assessed. In particular, the linking coefficients provided in this study can be used by researchers to calculate scale scores on PROMIS pediatric or adult physical health scales given data on only one age-specific instrument. As a result, researchers may potentially achieve more accurate measurement in cross-sectional studies spanning multiple age groups or longitudinal studies that require comparable measurement across distinct developmental stages.

Compliance with ethical standards

References

1. Donkervoort, M., Wiegerink, D. J., van Meeteren, J., Stam, H. J., Roebroeck, M. E., & Transition Research Group South West, N (2009). Transition to adulthood: validation of the Rotterdam Transition Profile for young adults with cerebral palsy and normal intelligence. *Developmental Medicine & Child Neurology, 51*(1), 53–62.

2. Stevens, S. E., Steele, C. A., Jutai, J. W., Kalnins, I. V., Bortolussi, J. A., & Biggar, W. D. (1996). Adolescents with physical disabilities: some psychosocial aspects of health. Journal of Adolescent Health, *19*(2), 157–164.

3. Ansell, B. M., & Chamberlain, M. A. (1998). Children with chronic arthritis: the management of transition to adulthood. *Baillieres Clinical Rheumatology, 12*(2), 363–373.

4. Magill-Evans, J., Wiart, L., Darrah, J., & Kratochvil, M. (2005). Beginning the transition to adulthood: the experiences of six families with youths with cerebral palsy. *Physical & Occupational Therapy, 25*(3), 19–36.

5. Matza, L. S., Swensen, A. R., Flood, E. M., Secnik, K., & Leidy, N. K. (2004). Assessment of health-related quality of life in children: a review of conceptual, methodological, and regulatory issues. *Value Health, 7*(1), 79–92.

6. Varni, J. W., Seid, M., & Kurtin, P. S. (2001). PedsQL 4.0: reliability and validity of the Pediatric Quality of Life Inventory version 4.0 generic core scales in healthy and patient populations. *Medical Care, 39*(8), 800–812.

7. Varni, J. W., Burwinkle, T. M., Seid, M., & Skarr, D. (2003). The PedsQL 4.0 as a pediatric population health measure: feasibility, reliability, and validity. Ambulatory *Pediatrics, 3*(6), 329–341.

8. Cella, D., Yount, S., Rothrock, N., Gershon, R., Cook, K., Reeve, B., Ader, D., Fries, J. F., Bruce, B., Rose, M., & Group, P. C. (2007). The Patient-Reported Outcomes Measurement Information System (PROMIS): progress of an NIH Roadmap cooperative group during its first two years. *Medical Care, 45*(5 Suppl 1), S3–S11.

9. Ader, D. N. (2007). Developing the Patient-Reported Outcomes Measurement Information System (PROMIS). *Medical Care, 45*(5 Suppl 1), S1–S2.

10. Irwin, D. E., Stucky, B. D., Thissen, D., Dewitt, E. M., Lai, J. S., Yeatts, K., Varni, J. W., & DeWalt, D. A. (2010). Sampling plan and patient characteristics of the PROMIS pediatrics large-scale survey. Quality of Life Research, *19*(4), 585–594.

11. Cella, D., Riley, W., Stone, A., Rothrock, N., Reeve, B., Yount, S., Amtmann, D., Bode, R., Buysse, D., Choi, S., Cook, K., Devellis, R., DeWalt, D., Fries, J. F., Gershon, R., Hahn, E. A., Lai, J. S., Pilkonis, P., Revicki, D., Rose, M., Weinfurt, K., & Hays, R. (2010). The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. Journal of Clinical Epidemiology, *63*(11), 1179–1194.

12. Reeve, B. B., Thissen, D., DeWalt, D. A., Huang, I. C., Liu, Y., Magnus, B., Quinn, H., Gross, H. E., Kisala, P. A., Ni, P., Haley, S., Mulcahey, M. J., Charlifue, S., Hanks, R. A., Slavin, M., Jette, A., & Tulsky, D. S. (2016). Linkage between the PROMIS(R) pediatric and adult emotional distress measures. Quality of Life Research, *25*(4), 823–833.

13. Gershon, R., Rothrock, N. E., Hanrahan, R. T., Jansky, L. J., Harniss, M., & Riley, W. (2010). The development of a clinical outcomes survey research application: Assessment Center. Quality of Life Research, *19*(5), 677–685.

14. McPherson, M., Arango, P., Fox, H., Lauver, C., McManus, M., Newacheck, P. W., Perrin, J. M., Shonkoff, J. P., & Strickland, B. (1998). A new definition of children with special health care needs. *Pediatrics, 102*(1 Pt 1), 137–140.

15. Neff, J. M., Sharp, V. L., Muldoon, J., Graham, J., Popalisky, J., & Gay, J. C. (2002). Identifying and classifying children with chronic conditions using administrative data with the clinical risk group classification system. *Ambulatory Pediatrics, 2*(1), 71–79.

16. Bethell, C. D., Read, D., Stein, R. E., Blumberg, S. J., Wells, N., & Newacheck, P. W. (2002). Identifying children with special health care needs: development and evaluation of a short screening instrument. *Ambulatory Pediatrics, 2*(1), 38–48.

17. Dorans, N. J. (2004). Equating, concordance, and expectation. *Applied Psychological Measurement, 28*, 227–246.

18. Choi, S. W., Schalet, B., Cook, K. F., & Cella, D. (2014). Establishing a common metric for depressive symptoms: Linking the BDI-II, CES-D, and PHQ-9 to PROMIS Depression. *Psychological Assessment, 26*(2), 513.

19. Dorans, N. J., & Holland, P. W. (2000). Population invariance and the equatability of tests: Basic theory and the linear case. *Journal of Educational Measurement, 37*(4), 281–306.

20. Cai, L., Thissen, D., & duToit, S. H. C. (2011). *IRTPRO for windows*. Lincolnwood: Scientific Software International.

21. Thissen, D., Varni, J. W., Stucky, B. D., Liu, Y., Irwin, D. E., & Dewalt, D. A. (2011). Using the PedsQL 3.0 asthma module to obtain scores comparable with those of the PROMIS pediatric asthma impact scale (PAIS). *Quality Life Research, 20*(9), 1497–1505.

22. Thissen, D., Liu, Y., Magnus, B., & Quinn, H. (2015). Extending the use of multidimensional IRT calibration as projection: many-to-one linking and linear computation of projected scores. In L. A. van der Ark, D. M. Bolt, W.-C. Wang, J. A. Doublas & S.-M. Chow (Eds.), Quantitative Psychology Research: The 79th annual meeting of the psychometric society, Madison, Wisconsin, 2014 (pp. 1–16). New York: Springer.

23. Center, N. S., & C. I., S. (2015). *Spinal Cord Injury (SCI) Facts and Figures at a Glance*. Birmingham: Spain Rehabilitation, University of Alabama.

24. Centers for Disease Control and Prevention. (1999). *TBI in the United States: A Report to Congress*. Atlanta: CDC.

25. Hays, R. D., Spritzer, K. L., Amtmann, D., Lai, J. S., Dewitt, E. M., Rothrock, N., Dewalt, D. A., Riley, W. T., Fries, J. F., & Krishnan, E. (2013). Upper-extremity and mobility subdomains from the Patient-reported outcomes measurement information system (PROMIS) adult physical functioning item bank. *Archives of Physical Medicine and Rehabilitation, 94*(11), 2291–2296.