Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Qual Life Res. Author manuscript; available in PMC 2022 April 01.

Published in final edited form as:

Qual Life Res. 2021 April ; 30(4): 997–1004. doi:10.1007/s11136-020-02700-5.

Using Nationally Representative Percentiles to Interpret PROMIS Pediatric Measures

Adam C. Carle,
Cincinnati Children's Hospital Medical Center, Department of Pediatrics; University of Cincinnati College of Medicine, Department of Pediatrics; University of Cincinnati College of Arts and Sciences, Department of Psychology; Cincinnati, OH, USA

Katherine B. Bevans,
Temple University College of Public Health

Carole A. Tucker,
Temple University College of Public Health, Department of Health and Rehabilitation Sciences

Christopher B. Forrest
Applied Clinical Research Center, Children's Hospital of Philadelphia, Philadelphia, PA, USA

Abstract

Purpose: This study's aim was to use a representative sample of the US pediatric population to estimate percentiles for several PROMIS pediatric measures: Anger, Anxiety, Depressive Symptoms, Family Relationships, Fatigue, Global Health, Life Satisfaction, Meaning and Purpose, Pain Behavior, Pain Interference, Physical Activity, Physical Function Mobility, Physical Function Upper Extremity, Physical Stress Experiences, Positive Affect, Psychological Stress Experiences, Sleep Disturbance, Sleep Impairment, and Peer Relationships.

Methods: We used two separate, nationally representative samples of parents and children aged 5–17 years drawn in different years from the GfK Knowledge Panel, a dual-frame online probability panel.

Results: All measures that were developed using a representative sample had a median at or near the expected value of 50. For the other measures, the 50th percentile was often 10 points or more from 50. Several domains had high floors or low ceilings. No domain's percentiles completely corresponded to the percentiles associated with a normal distribution with a mean of 50 and standard deviation of 10.

Corresponding Author: Adam Carle, adam.carle@cchmc.org.

Publisher's Disclaimer: This Author Accepted Manuscript is a PDF file of an unedited peer-reviewed manuscript that has been accepted for publication but has not been copyedited or corrected. The official version of record that is published in the journal is kept up to date and so may therefore differ from this version.

Conflicts of interest/Competing interests: None of the authors have any conflicts of interest.

Ethics approval:
This study was reviewed and approved by The Children's Hospital of Philadelphia IRB. This study was conducted in full accordance with the ethical standards in the 1964 Declaration of Helsinki, all applicable Children's Hospital of Philadelphia Research Policies and Procedures, and all applicable Federal and state laws and regulations including 45 CFR 46.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Conclusions: This work allows users to interpret a child's self-reported quality of life relative to children in the US general population. When attempting to evaluate whether a child falls above or below other children in the US, one should use the values presented in this study. In addition, we recommend that users should focus on whether a child's score falls into one of a few broad severity groups rather than on specific percentile scores.

Keywords

Pediatrics; Patient reported outcomes; PROMIS; Percentiles; Interpretation

Introduction

Recent years have seen an increase in the use of the Patient Reported Outcome Measurement Information System (PROMIS) pediatric measures.[1–19] However, some users grapple with interpreting scores. Because PROMIS measures were developed using item response theory (IRT),[2; 20] PROMIS T-scores can theoretically be interpreted relative to a normal distribution with a mean of 50 and a standard deviation of 10. However, in practice, a T-score of 50 will reflect the mean of the sample in which the IRT item parameters were estimated. If the parameters were estimated using a convenience sample that included a disproportionately large number of children with impaired health, a T-score of 50 will indicate higher impairment than the average child in the US experiences. Although some PROMIS measures' parameters were estimated using a nationally representative sample, many were not. Yet, some users and researchers refer to T=50 as "the average" without context,[14] even though some PROMIS investigators have clearly stated that a "score of 50 does not represent any one group."[8]

One potential solution to this problem would be to collect new, nationally representative data for each measure and re-estimate the item parameters. This would make scores estimated from the parameters interpretable relative to the general population. For example, a T score of 50 would be more meaningful for each measure given that it would correspond to the average score of US children. Alternatively, rather than re-estimating parameters, one could use nationally representative data to "center" the T-scores so that 50 corresponded to the general population's mean. However, both solutions are problematic because much work has already been done using PROMIS. Either solution would create substantial confusion when comparing scores estimated at different times across studies, or even across time for a given individual. As an additional complication, even though PROMIS measures have good reliability and validity, they are not perfect. Some levels are measured more precisely than others. Thus, even though the underlying latent variable may be standard normal, the distribution of estimated scores, even in a large sample of the general population, may not be normal.

Empirically-based percentiles offer a solution.[21; 22] Based on an observed vs. a theoretical distribution of scores, the pth percentile gives the score at which p percent of the observations in the representative sample fall at or below. As a result, they allow one to compare a child or a group of children with that of a reference population. Empirical percentiles for PROMIS measures would not depend on a theoretical distribution. Instead

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

they would be based an empirical distribution of estimated scores. In this paper, we describe the first study to use a representative sample of the US non-institutionalized pediatric population to estimate percentiles for several PROMIS pediatric child- and proxy-report measures: Anger, Anxiety, Depressive Symptoms, Family Relationships, Fatigue, Global Health, Life Satisfaction, Meaning and Purpose, Pain Behavior, Pain Interference, Physical Activity, Physical Function Mobility, Physical Function Upper Extremity, Physical Stress Experiences, Positive Affect, Psychological Stress Experiences, Sleep Disturbance, Sleep Impairment, and Peer Relationships.

Methods

Sample

We used two separate, nationally representative samples of parents and children aged 5–17 years. Both samples from were drawn in different years from the GfK Knowledge Panel, a dual-frame (random-digit dial and address-based) online probability panel.[23] Sample 1 measured all but Sleep Disturbance and Impairment, which were developed after Sample 1 was collected. Raked probability weights, which also account for oversampling of individuals living in minority communities and Spanish-language dominant areas and other sources of error, make Sample 1's weighted distributions[23] of age, gender, race, ethnicity, education, census region, metropolitan area, household internet access, and language (English/Spanish) match those in the 2013 Current Population Survey (CPS).[24] In Sample 2, which measured Sleep Disturbance and Impairment, similarly created weights make the weighted distributions match the 2015 CPS. We used the weights for all analyses in this paper.

Procedures

GfK emailed adult participants known to care for children aged 5–17 years to notify them of eligibility. Caregivers who provided informed consent were emailed a link to an online questionnaire. The majority of caregivers were mothers (71%). After completing proxy report, caregivers of children aged 8–17 years asked their child to complete the child questionnaires. These procedures resulted in a sample of 5,206 caregivers and 4,005 children aged 8–17 years for Sample 1 and 1,260 caregivers and 941 children aged 8–17 years for Sample 2. Rates of missing data were ≤ 1% for all items. To reduce respondent burden in Sample 1, short forms were grouped in blocks (e.g., Anxiety, Positive Affect, Pain Intensity, and Anger were blocked together) and children were randomly assigned to one of the blocks. Thus, the sample size for any given domain was approximately 1,000.

Score Estimation

All scores were estimated using their 8 item short forms, with the exception of Anger (6 items), Anxiety and Fatigue (10 items each), and Global Health, which is a 7 item measure and does not have a short form. For each measure, we used the published item parameters[3; 4; 7; 9; 11; 12; 14; 16; 17; 25–32] and estimated Bayesian Expected A Posteriori (EAP) IRT scores[33] using Mplus 8.[34] We transformed these into T-scores (T = IRT-score*10+50). While all of the pediatric measures use the graded response model,[35] the samples in which

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

the parameters were estimated differ.[3; 4; 7; 9; 11; 12; 14; 16; 17; 25–32] We provide a brief summary given a sample's impact on interpretation.

The Anger, Anxiety, Depressive Symptoms, Fatigue, Mobility, Peer Relationships, and Upper Extremity measures were administered to a racially diverse convenience sample of 4,128 children aged 8–17.[13] Participants were recruited from hospital-based outpatient general pediatrics clinics, subspecialty clinics, and public schools. The Pain Interference measure's items were administered to a racially diverse convenience sample of 3,048 children aged 8–17 recruited from hospital-based outpatient general and subspecialty clinics and public schools.[14] The Pain Behavior measure was administered to 450 pediatric patients aged 8–17 years with chronic pain at 3 pediatric medical centers the Midwest, Northeast and Southeast United States. The Global Health measure's parameters were estimated using a quota sample of 3,635 children and youth aged 8–17 years old from a national internet panel maintained by Op4G.[4] None of these samples used methods designed to make estimates representative of the non-institutionalized US pediatric population.

The item parameters for the Family Relationships, Life Satisfaction, Meaning and Purpose, Physical Activity, Physical Stress Experiences, Psychological Stress Experiences, and Positive Affect measures were estimated using a multigroup IRT approach and data from Sample 1 above and a sample comprised of 1) individuals purposively sampled from an opt-in internet panel (~55%), 2) students (and their parents) from school districts in New Hampshire, Vermont, and Texas (~37%), and 3) a sample of children from primary and specialty clinics and the Emergency Department at Children's Hospital of Philadelphia clinics (~9%). The item parameters for the Sleep Disturbance and Sleep Impairment measures were estimated using Sample 2.[31] As a result of the samples and calibration methods, the IRT parameters and estimated scores for these measures can be interpreted relative to the non-institutionalized US pediatric population.

Percentile Calculation

Let $x_{(j)}$ index the x values of a given set of PROMIS scores in ascending order ($j = 1,2, \ldots, n$) and let $w_{(j)}$ refer to the weight associated with each $x_{(j)}$ observation. The weighted sample size (N) equals $\sum_{j=1}^{n} w_{(j)}$. To calculate the pth percentile ($x_{[p]}$), let P equal $\frac{N_p}{100}$ and $W_{(i)} = \sum_{j=1}^{i} w_{(j)}$ and then identify the first index, i, where $W_{(i)} > P$. In turn, the pth percentile is $x_{[p]} = \begin{cases} \frac{x_{(i+1)} + x_{(i)}}{2} \\ x_{(i)} \end{cases}$. Consequently, a given percentile corresponds to the percent of scores that fall at or below a given score.

Severity Cut-Points

Consistent with previous research that has noted the utility of categorizing children into severity groups using percentiles,[36; 37] we created severity groups based on percentile cut-points. We selected severity cut-points based on the authors' expert opinions, face validity, and consistency with other measures in medicine (e.g., both body mass index [BMI]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

and hypertension have a cut-point at the 95th percentile). For domains where higher scores reflect poorer functioning, we defined the severity groups as Minimal (≤ 74th percentile), Moderate (75–94th percentile), and Severe (≥ 95th percentile). For domains where higher scores reflect greater functioning, we defined the severity groups as Good (≥ 26th percentile), Fair (6th–25th percentiles), and Poor (≤ 5th percentile). When ties occurred, we used the midpoint. For example, a Global Health T-score of 41 covered the 24th, 25th, and 26th percentiles. We considered those with scores below 41.5 in the fair category and those greater than or equal to 41.5 in the good category.

Results

Table 1 presents the sample characteristics. Table 2 displays severity cut-points for each domain and for child and proxy report. The full percentile tables are available online as supplemental tables. As shown in Table 2, across domains no single common score defines the cut-point between the severity categories; Minimal-Moderate ranged 47.5 to 57.5, Moderate-Severe ranged 55 to 66, Fair-Good ranged 41.5 to 51.5, and Poor to Fair ranged 33 to 42. While the cut-points varied across domains, across reporters the cut-points were often very similar or identical, though this was not true for the Minimal-Moderate cut-points for Fatigue and Pain Behavior and the Fair-Good and Fair-Poor cut-points for Upper Extremity Function.

Table 2 also demonstrates the scores that correspond to the 5th, 50th, and 95th percentiles. For Fatigue, Mobility, Pain Behavior, Pain Interference, and Upper Extremity, the 50th percentile (median) was nearly 10 points and sometimes more from T=50. For Anger, Anxiety, and Depressive Symptoms, the median was nearly 5 points and sometimes more from T=50. All of the measures developed using a representative sample had a median at or close to the expected value of 50. Table 2 also shows that, while the 95th percentile theoretically corresponds to a T-score of 66, empirically this was true only for Physical Stress Experiences, Psychological Stress Experiences, and Positive Affect (though Anger, Sleep Disturbance, Sleep Impairment, Physical Activity, and Peer Relationships were within 1 point of 66). Analogous issues occurred for the 5th percentile, which theoretically corresponds to a T-score of 33. Table 2 also shows that for Pain Behavior, Mobility, and Upper Extremity Function, none of the T-scores associated with the observed percentiles corresponded to the theoretically expected T-scores. For example, Pain Behavior's median T-score was 32 for child report, while the T-scores for the 5th and 95th percentiles were 26 and 55 respectively. Again, except for Fatigue, Pain Behavior, and Upper Extremity Function, the percentiles were relatively similar or equal across reporters within a domain.

Finally, Table 3 shows the percent of children with scores at the floor for domains where higher scores reflect poorer functioning and the percent of children with scores at ceiling for domains where higher scores reflect greater functioning. As Table 3 reveals, several of the domains had high floors. For example, 44% of children had Pain Behavior scores at the floor. Several domains also had low ceilings. For example, 63% of children had Upper Extremity Function scores at the ceiling, while 66% of children's Mobility scores were at the ceiling. For all domains, the floors and ceilings respectively were across the "better" functioning range.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Discussion

In this study, we examined the distribution of T-scores for multiple PROMIS pediatric short form measures using data representative of the non-institutionalized US pediatric population. Our results highlight several things. Importantly, the empirical distribution of percentiles often did not correspond to the theoretical distribution. When attempting to evaluate whether a child falls above or below the other children and understand a score's severity relative to the general population, one should use the values presented in Table 2. Generally, the difference between the theoretical and observed percentiles was likely driven by whether a convenience sample was used for calibration or not. However, differences also likely occurred because of measurement imprecision and because we estimated T-scores using short-forms rather than a domain's full item set.

Floor and ceiling effects also impacted the empirical distribution of percentiles. All of the domains where higher scores reflect poorer functioning demonstrate floor effects and nearly all of the domains where higher scores reflect better functioning showed ceiling effects. The floor and ceiling effects occurred in the healthy range of the scores, which lessens the impact of this problem on clinical evaluation. Nevertheless, users should be aware that for some PROMIS pediatric short form measures, 45–65% or more of the scores will fall at the ceiling or floor.

We also found that generally, while the distribution of percentiles differed across domains, the distribution across reporters within a domain was relatively similar. Although all PROMIS domains were developed using IRT, the parameters were not estimated in a way that placed the child and proxy report measures on an equivalent metric. Despite this, it appears that one can interpret the severity of a T-score similarly whether or not it was estimated using child report or proxy report, though this does not mean that one can interchange scores across reporters within a statistical analysis.

In general, we recommend that users focus less on specific percentiles (65 vs. 45) and instead consider whether a child's scores falls into one of a few broad groups. This is partly because, as we note in more detail below, any single score for a given child lacks sufficient precision to make strong conclusions based on the specific score. For symptom-type measures (Anger, Anxiety, Depressive Symptoms, Family Relationships, Fatigue, Pain Behavior, Pain Interference, Physical Function Mobility, Physical Function Upper Extremity, and Sleep Impairment), we categorized scores as minimal, moderate, or severe. For the evaluative type measures (Global Health, Life Satisfaction, Meaning and Purpose, Physical Activity, Physical Stress Experiences, Positive Affect, Psychological Stress Experiences, Sleep Disturbance, and Peer Relationships), we categorized scores as good, fair, and poor.

Importantly, users should take into account the precision of a child's score. Any given T-score will have an estimate of the error associated with it and one can create a confidence interval around the score using it. For example, in our data one child had a Fatigue T-score of 41 based on their item responses. While this point estimate suggests the child is in the Minimal category and ~45% of the general population has greater fatigue, the upper bound of a 95% confidence interval using the score's standard error of 4.5 indicates that the child

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

may be in the Moderate category with a T score closer to 48.5 (76th percentile). Because PROMIS measure were designed using IRT, the precision of scores varies across levels of a domain. The precision of a score also depends partly on the extent to which a child's item responses correspond to what one would expect given the model. Given their basis in IRT, PROMIS measures allow one to more realistically and individually understand a child's health related quality of life than measures not based in IRT.

Limitations

First, to our knowledge, no studies have examined the relationship between specific PROMIS scores and children's clinical, educational, and psychosocial outcomes. Future research should seek to better understand the correlates and long-term outcomes of different scores and percentiles. This would be particularly helpful in evaluating the appropriateness of the severity categories' cut-points. Research may show that the cut-points should be refined. Second, the sample was drawn from an internet panel. A different frame that did not rely on the internet might lead to somewhat different results. Third, our sample was insufficiently large to estimate percentiles across socio-demographic or clinical subgroups. Fourth, participants only received the short form item sets. This limited the range of estimated scores and means that users with scores based on a CAT or full item bank should use the percentile tables in this manuscript with some caution.

Conclusion

This study presents percentiles for 19 PROMIS pediatric self- and proxy-report short form measures based on data nationally representative of non-institutionalized US children. Our findings should help users understand where a given child's score falls relative to other children in the US, as well as understand some of the limits of PROMIS T-scores.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments:

Research reported in this publication was supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases of the National Institutes of Health under Award Numbers U19AR069525 and U01AR057956. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

1. Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, Amtmann D, Bode R, Buysse D, & Choi S (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. [PubMed: 20685078]
2. Reeve BB, Hays RD, Bjorner JB, Cook KF, Crane PK, Teresi JA, Thissen D, Revicki DA, Weiss DJ, Hambleton RK, Liu H, Gershon R, Reise SP, Lai J. s., Cella D, & Group, P. C. (2007). Psychometric evaluation and calibration of health-related quality of life item banks: plans for the Patient-Reported Outcomes Measurement Information System (PROMIS). Medical care, 45(5 Suppl 1), S22–31. [PubMed: 17443115]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

3. Bevans KB, Riley AW, Landgraf JM, Carle AC, Teneralli RE, Fiese BH, Meltzer LJ, Ettinger AK, Becker BD, & Forrest CB (2017). Children's family experiences: development of the PROMIS® pediatric family relationships measures. Quality of Life Research, 1–13.
4. Forrest C, Bevans K, Pratiwadi R, Moon J, Teneralli R, Minton J, & Tucker C (2014). Development of the PROMIS® pediatric global health (PGH-7) measure. Quality of Life Research, 23(4), 1221–1231. [PubMed: 24264804]
5. Broderick J, Morgan DeWitt E, Rothrock N, Crane P, & Forrest C (2013). Advances in Patient Reported Outcomes: The NIH PROMIS Measures. eGEMs (Generating Evidence & Methods to improve patient outcomes), 1(1).
6. Forrest CB, Bevans KB, Tucker C, Riley AW, Ravens-Sieberer U, Gardner W, & Pajer K (2012). Commentary: The Patient-Reported Outcome Measurement Information System (PROMIS®) for children and youth: Application to pediatric psychology. Journal of Pediatric Psychology, 37(6), 614–621. [PubMed: 22362923]
7. Jacobson CJ, Kashikar-Zuck S, Farrell JE, Barnett K, Goldschneider K, Dampier C, Cunningham N, Crosby L, & DeWitt EM (2015). Qualitative Evaluation of Pediatric Pain Behavior, Quality, and Intensity Item Candidates and the PROMIS Pain Domain Framework in Children With Chronic Pain. Journal of pain, 16(12).
8. DeWalt DA, Gross HE, Gipson DS, Selewski DT, DeWitt EM, Dampier CD, Hinds PS, Huang I-C, Thissen D, & Varni JW (2015). PROMIS® 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. [PubMed: 25715946]
9. DeWalt DA, Thissen D, Stucky BD, Langer MM, Morgan DeWitt E, Irwin DE, Lai J-S, Yeatts KB, Gross HE, & Taylor O (2013). PROMIS Pediatric Peer Relationships Scale: development of a peer relationships item bank as part of social health measurement. Health Psychology, 32(10), 1093. [PubMed: 23772887]
10. Varni JW, Thissen D, Stucky BD, Liu Y, Gorder H, Irwin DE, DeWitt EM, Lai J-S, Amtmann D, & DeWalt DA (2012). PROMIS® Parent Proxy Report Scales: an item response theory analysis of the parent proxy report item banks. Quality of Life Research, 21(7), 1223–1240. [PubMed: 21971875]
11. Irwin DE, Stucky BD, Langer MM, Thissen D, DeWitt EM, Lai J-S, Yeatts KB, Varni JW, & DeWalt DA (2012). PROMIS Pediatric Anger Scale: an item response theory analysis. Quality of Life Research, 21(4), 697–706. [PubMed: 21785833]
12. Irwin DE, Stucky B, Langer MM, Thissen D, DeWitt EM, Lai J-S, Varni JW, Yeatts K, & DeWalt DA (2010). An item response analysis of the pediatric PROMIS anxiety and depressive symptoms scales. Quality of Life Research, 19(4), 595–607. [PubMed: 20213516]
13. Irwin DE, Stucky BD, Thissen D, DeWitt EM, Lai JS, Yeatts K, Varni JW, & DeWalt DA (2010). Sampling plan and patient characteristics of the PROMIS pediatrics large-scale survey. Quality of Life Research, 19(4), 585–594. [PubMed: 20204706]
14. Varni JW, Stucky BD, Thissen D, Dewitt EM, Irwin DE, Lai JS, Yeatts K, & Dewalt DA (2010). PROMIS Pediatric Pain Interference Scale: an item response theory analysis of the pediatric pain item bank. Journal of Pain, 11(11), 1109–1119.
15. Morgan EM, Mara CA, Huang B, Barnett K, Carle AC, Farrell JE, & Cook KF (2017). Establishing clinical meaning and defining important differences for Patient-Reported Outcomes Measurement Information System (PROMIS®) measures in juvenile idiopathic arthritis using standard setting with patients, parents, and providers. Quality of Life Research, 26(3), 565–586. [PubMed: 27913986]
16. Forrest CB, Ravens-Sieberer U, Devine J, Becker BD, Teneralli RE, Moon J, Carle AC, Tucker CA, & Bevans KB (2018). Development and evaluation of the PROMIS® pediatric positive affect item bank, child-report and parent-proxy editions. Journal of happiness studies, 19(3), 699–718. [PubMed: 29760578]
17. Cunningham NR, Kashikar-Zuck S, Mara C, Goldschneider KR, Revicki DA, Dampier C, Sherry DD, Crosby L, Carle A, & Cook KF (2017). Development and validation of the self-reported PROMIS pediatric pain behavior item bank and short form scale. Pain, 158(7), 1323–1331. [PubMed: 28394851]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

18. Carle AC, Cella D, Cai L, Choi SW, Crane PK, Curtis SMK, Gruhl J, Lai JS, Mukherjee S, & Reise SP (2011). Advancing PROMIS's methodology: results of the Third Patient-Reported Outcomes Measurement Information System (PROMIS®) Psychometric Summit. Expert Review of Pharmacoeconomics & Outcomes Research, 11(6), 677–684. [PubMed: 22098283]
19. Gerhardt WE, Mara CA, Kudel I, Morgan EM, Schoettker PJ, Napora J, Britto MT, & Alessandrini EA (2018). Systemwide Implementation of Patient-Reported Outcomes in Routine Clinical Care at a Children's Hospital. The Joint Commission Journal on Quality and Patient Safety, 44(8), 441–453. [PubMed: 30071964]
20. Hambleton RK (1985). Item Response Theory. Boston, MA: Kluwer.
21. Crawford JR, & Garthwaite PH (2009). Percentiles please: The case for expressing neuropsychological test scores and accompanying confidence limits as percentile ranks. The Clinical Neuropsychologist, 23(2), 193–204. [PubMed: 18609335]
22. Crawford JR, Garthwaite PH, & Slick DJ (2009). On percentile norms in neuropsychology: Proposed reporting standards and methods for quantifying the uncertainty over the percentile ranks of test scores. The Clinical Neuropsychologist, 23(7), 1173–1195. [PubMed: 19322734]
23. IPSOS. KnowledgePanel. Retrieved 2020-11-02, from https://www.ipsos.com/en-us/solutions/public-affairs/knowledgepanel
24. United States Census Bureau. Current Population Survey. Retrieved 2020-11-02, from https://www.census.gov/programs-surveys/cps.html
25. Bevans KB, Gardner W, Pajer KA, Becker B, Carle A, Tucker CA, & Forrest CB (2018). Psychometric Evaluation of the PROMIS® Pediatric Psychological and Physical Stress Experiences Measures. Journal of pediatric psychology, 43(6), 678–692. [PubMed: 29490050]
26. Forrest CB, Devine J, Bevans KB, Becker BD, Carle AC, Teneralli RE, Moon J, Tucker CA, & Ravens-Sieberer U (2018). Development and psychometric evaluation of the PROMIS Pediatric Life Satisfaction item banks, child-report, and parent-proxy editions. Quality of Life Research, 27(1), 217–234. [PubMed: 28828568]
27. Lai J-S, Stucky BD, Thissen D, Varni JW, DeWitt EM, Irwin DE, Yeatts KB, & DeWalt DA (2013). Development and psychometric properties of the PROMIS® pediatric fatigue item banks. Quality of Life Research, 22(9), 2417–2427. [PubMed: 23378106]
28. Irwin DE, Stucky BD, Thissen D, DeWitt EM, Lai JS, Yeatts K, Varni JW, & DeWalt DA (2010). Sampling plan and patient characteristics of the PROMIS pediatrics large-scale survey. Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care & Rehabilitation, 19(4), 585–594.
29. DeWitt EM, Stucky BD, Thissen D, Irwin DE, Langer M, Varni JW, Lai J-S, Yeatts KB, & DeWalt DA (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. [PubMed: 21292444]
30. Ravens-Sieberer U, Devine J, Bevans K, Riley AW, Moon J, Salsman JM, & Forrest CB (2014). Subjective well-being measures for children were developed within the PROMIS project: presentation of first results. Journal of clinical epidemiology, 67(2), 207–218. [PubMed: 24295987]
31. Forrest CB, Meltzer LJ, Marcus CL, de la Motte A, Kratchman A, Buysse DJ, Pilkonis PA, Becker BD, & Bevans KB (2018). Development and validation of the PROMIS Pediatric Sleep Disturbance and Sleep-Related Impairment item banks. Sleep, 41(6), zsy054.
32. Bevans KB, Meltzer LJ, De La Motte A, Kratchman A, Viél D, & Forrest CB (2018). Qualitative Development and Content Validation of the PROMIS Pediatric Sleep Health Items. Behavioral sleep medicine, 1–15.
33. Bock RD, & Mislevy RJ (1982). Adaptive EAP estimation of ability in a microcomputer environment. Applied psychological measurement, 6(4), 431–444.
34. Muthén LK, & Muthén BO (1998–2012). Mplus User's Guide. Seventh Edition. Los Angeles, CA: Muthén & Muthén.
35. Samejima F (1969). Estimation of latent ability using a response pattern of graded scores. Psychometrika Monograph Supplement, 34(4, Pt. 2), 100.
36. Daniels SRJP (2009). The use of BMI in the clinical setting. 124(Supplement 1), S35–S41.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

37. Carle AC, Blumberg SJ, Moore KA, & Mbwana K (2011). Advanced Psychometric Methods for Developing and Evaluating Cut-Point-Based Indicators. Child Indicators Research, 1–26.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 1:

Sample characteristics.

		Sleep Measures %			All Other Measures %		
		f	Unweighted	% Weighted	f	Unweighted	% Weighted
Child's Gender	Male	480	51.0	50.9	2,021	50.8	50.8
	Female	461	49.0	49.1	1,955	49.2	49.2
Child's Race and Ethnicity	Non-Hispanic White	585	62.2	54.5	2,638	65.9	54.6
	Non-Hispanic Black or African American	78	8.3	13.9	310	7.7	13.6
	Non-Hispanic Asian	34	3.6	3.4	160	4.0	3.2
	Non-Hispanic American Indian or Alaskan Native	6	0.6	0.6	10	0.2	0.2
	Non-Hispanic Native Hawaiian or Pacific Islander	9	1.0	0.9	24	0.6	0.4
	Non-Hispanic Other or More Than One	49	5.2	4.3	256	6.4	5.5
	Hispanic	180	19.1	22.4	604	15.1	22.6
Child's Age in Years	8–12	515	54.7	48.2	1,936	48.7	48.4
	13–17	426	45.3	51.8	2,037	51.3	51.6
Child's Proxy's Education	High school or less	329	35.0	36.1	631	15.8	30.3
	Some college	295	31.3	29.4	1,415	35.3	34.4
	Bachelor's degree or higher	317	33.7	34.5	1,959	48.9	35.4

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 2:

Cut-points and Select Percentiles for Pediatric PROMIS Measures.

High Scores = Poorer Functioning	Minimal-Moderate Cut-Point		Moderate-Severe Cut-Point		5th Percentile		Median		95th Percentile	
	Child Report	Proxy Report	Child Report	Proxy Report	Child Report	Proxy Report	Child Report	Proxy Report	Child Report	Proxy Report
Anger	54.5	53	65	64	24	23	44	43	65	64
Anxiety	52.0	52.0	63.0	62	25	28	42	43	63	62
Depressive Symptoms	53.5	52	62.5	60.5	35	32	45	44	62	60
Fatigue	47.5	53.5	60	62	20	37	37	47	60	61
Pain Behavior	48.5	41	55	54	26	3	32	13	55	53
Pain Interference	49.0	52	58.0	59.5	25	34	36	42	58	59
Physical Stress Experiences	56.5	55.5	66.0	65	39	41	50	49	66	65
Psychological Stress Experiences	56.5	56.5	66	66	37	38	50	50	66	65
Sleep Disturbance	57.5	56.5	65.5	66	37	39	50	49	65	66
Sleep Impairment	57	57.5	65	66	38	38	50	50	65	66
High Scores = Greater Functioning	**Fair-Good Cut-Point**		**Poor-Fair Cut-Point**		**5th Percentile**		**Median**		**95th Percentile**	
Family Relationships	43	46.5	34	39	34	39	50	50	63	58
Global Health	41.5	44.5	36.5	35.5	36	35	48	49	61	63
Life Satisfaction	44.5	44.5	35.5	35.5	35	35	49	50	63	62
Meaning and Purpose	43	45.5	34	36	34	36	50	51	61	60
Mobility	51.5	52	42	40	42	40	63	61	63	61
Physical Activity	44.5	43	33	33	33	33	49	48	67	67
Positive Affect	44	45.5	36	36.5	36	36	49	49	66	67
Peer Relationships	43	43.0	33	33.0	33	33	48	49	67	63
Upper Extremity	49.5	39.5	36	31	36	31	63	54	63	54

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 3:

Percent of Scores at the Floor and Ceiling across Domains.

	Percent at floor	
High Scores = Poorer Functioning	**Child Report**	**Proxy Report**
Anger	16	13
Anxiety	26	23
Depressive Symptoms	32	32
Fatigue	29	39
Pain Behavior	44	45
Pain Interference	44	49
Physical Stress Experiences	24	35
Psychological Stress Experiences	19	21
Sleep Disturbance	18	23
Sleep Impairment	21	21
High Scores = Greater Functioning	**Percent at ceiling**	
Family Relationships	14	15
Global Health	3	2
Life Satisfaction	25	28
Meaning and Purpose	30	33
Mobility	66	74
Physical Activity	3	2
Positive Affect	14	10
Peer Relationships	22	29
Upper Extremity	63	56