Published in final edited form as:

Qual Life Res. 2011 November ; 20(9): 1497–1505. doi:10.1007/s11136-011-9874-y.

Using the PedsQL™ 3.0 asthma module to obtain scores comparable with those of the PROMIS pediatric asthma impact scale (PAIS)

David Thissen,

Department of Psychology, CB# 3270, Davie Hall, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

James W. Varni,

Department of Pediatrics, College of Medicine, Texas A&M University, College Station, TX, USA;
Department of Landscape Architecture and Urban Planning, College of Architecture, Texas A&M University, College Station, TX, USA

Brian D. Stucky,

Department of Psychology, CB# 3270, Davie Hall, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

Yang Liu,

Department of Psychology, CB# 3270, Davie Hall, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

Debra E. Irwin, and

Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

Darren A. DeWalt

Division of General Medicine and Clinical Epidemiology and Cecil G. Sheps Center for Health Services Research, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

David Thissen: dthissen@email.unc.edu

Abstract

Background—The National Institutes of Health's Patient-Reported Outcomes Measurement Information System (PROMIS) has developed several scales measuring symptoms and function for use by the clinical research community. One advantage of PROMIS is the ability to link other scales to the PROMIS metric.

Objectives—The objectives of this research are to provide evidence of validity for one of the PROMIS measures, the Pediatric Asthma Impact Scale (PAIS), and to link the PedsQL™ Asthma Symptoms Scale with the metric of the PAIS.

Method—Descriptive statistics were computed describing the relationships among scores on the PAIS, the PedsQL™ Asthma Symptoms, Treatment, Worry, and Communication Scales, and the DISABKIDS Asthma Impact and Worry Scales for approximately 300 children ages 8–17. A

novel linkage method based on item response theory (IRT), *calibrated projection*, was used to link scores on the PedsQL™ Asthma Symptoms Scale with the metric of the PAIS.

Results—The PAIS exhibited strong convergent validity with the PedsQL™ Asthma Symptoms Scale, and less strong relations with the other five scales. The linkage system uses scores on the PedsQL™ Asthma Symptoms Scale to produce relatively precise score estimates on the metric of the PAIS.

Conclusions—Results of this study provide evidence for the validity of the PAIS, and a method to use scores on the PedsQL™ Asthma Symptoms Scale to estimate scores on the metric of the PAIS, in partial fulfillment of the PROMIS goal to provide a *lingua franca* for health-related quality of life.

Keywords

PROMIS; HRQOL; PRO; Scale development; Pediatrics; Asthma

Introduction

The National Institutes of Health Roadmap for Medical Research initiative included a large-scale effort to develop the Patient-Reported Outcomes Measurement Information System (PROMIS). One goal of PROMIS was to provide a common set of health-related quality of life (HRQoL) scales and metrics for the clinical research community [1, 2]. To accomplish part of that goal, the PROMIS pediatric project developed patient-reported outcome (PRO) item banks across several health domains for youth aged 8–17. These included scales for measurement of generic health domains that are important across a variety of illnesses, including physical function, pain, fatigue, emotional distress, and social function [3–6]. In addition, one disease-specific item bank was developed for children with asthma [7], which was chosen because it is the most common chronic childhood disease, and PRO measurement is an essential component in evaluating outcomes for children with asthma [8–10].

The PROMIS scales all provide measurement with metrics defined using item response theory (IRT). A goal of PROMIS has been that those scales, and their metrics, will become a *lingua franca* (bridge or common language) for HRQoL measurement in clinical research. Although that goal may be accomplished by widespread adoption of the PROMIS scales themselves, even more use of the PROMIS metrics can be obtained if scores on suitable existing scales can be “linked” with the PROMIS scales, to provide score estimates on the new metrics.

This research describes the relations of scores on one of the PROMIS measures, the Pediatric Asthma Impact Scale (PAIS) [7], with scores from two previously existing measures of HRQoL for children with asthma (PedsQL™ 3.0 and DISABKIDS Asthma Modules), to provide evidence of convergent and divergent validity for the PAIS. Further, the PedsQL™ Asthma Symptoms Scale [11] is sufficiently strongly related with the PAIS that it can be linked with the metric of the PAIS.

Methods

Measures

PROMIS pediatric asthma impact scale (PAIS)—The field-test item set for the PROMIS PAIS item bank was created using a strategic item generation methodology developed by the PROMIS Network [12]. Some questionnaire items were developed as variants of items on existing scales measuring asthma-related HRQoL, while others were

generated de novo. Focus groups and cognitive interviews [13, 14] were used to refine item wording and content. Items successfully screened through the focus groups and cognitive interviews were retained for field testing. The final PROMIS pediatric asthma field test item set contained 34 items. The response scale for each item is “never” (0), “almost never” (1), “sometimes” (2), “often” (3), and “almost always” (4).

A series of statistical analyses were performed to confirm the unidimensionality of the measure, to identify and set aside items that exhibited local dependence (LD) or differential item functioning (DIF), and finally to calibrate the remaining items onto an IRT scale using Samejima's graded item response model [15, 16]. The final PAIS item pool comprises seventeen items, with item parameters for IRT scoring, and a scoring table for an eight-item short form [7]. The PAIS is also available from the PROMIS Assessment Center, which contains the item pool and allows the user to select and administer the 17-item pool, an 8-item short form, or a computerized adaptive test (CAT) [17].

PedsQL™ 3.0 asthma module—The PedsQL™ 3.0 Asthma Module [11] comprises 4 scales: Asthma Symptoms (11 items), Treatment Problems (11 items), Worry (3 items), and Communication (3 items). The response scale for each item is the same as that for the PAIS: “never” (0), “almost never” (1), “sometimes” (2), “often” (3), and “almost always” (4). For conventional scoring of the PedsQL™ 3.0 Asthma Module, the responses are transformed to 100, 75, 50, 25, and 0, respectively, resulting in a scale range of 0–100, with higher scores indicating better HRQoL. Scores are computed as the mean of non-missing items, unless more than 50% of the responses for a scale or subscale are missing data, in which case there is no score.

DISABKIDS asthma module—The DISABKIDS Asthma Module [18] includes 2 scales: Impact (6 items) and Worry (5 items). The response scale is “never” (0), “seldom” (1), “quite often” (2), “very often” (3), and “always” (4). Scores on the DISABKIDS Asthma Module are the average of each scale's reversed responses multiplied by 25, yielding a score range of 0–100, with higher scores indicating better HRQoL. There is no score if responses are missing for more than one of the items on a scale.

Sampling plan

The overall data collection design for calibration of the PROMIS pediatric measures has been described elsewhere [3]. Participants were recruited in hospital-based outpatient general pediatrics and subspecialty clinics and in public school settings between January 2007 and May 2008 in North Carolina and Texas. Parental informed consent and minor assent were obtained for all children taking the survey.

The 34 PROMIS field-test asthma items were administered along with the items measuring other domains (e.g., emotional distress, fatigue) from the PROMIS pediatric item tryout set, as well as from the PedsQL™ 3.0 Asthma Module [11] and the DISABKIDS asthma module [18]. Children ($N = 622$) completed one of the two testing forms (Form Asthma 1, Form Asthma 2). Form Asthma 1 included the 34 PROMIS field-test asthma items, and items measuring other PROMIS domains (the latter are not considered further here). Form Asthma 2 included the same 34 asthma items, and the PedsQL and DISABKIDS asthma modules. The seventeen asthma items that were ultimately retained for the PAIS item pool [7] had a 7 day recall period. The details of the sampling plan are described elsewhere [3].

Analysis

Scores for the PedsQL and DISABKIDS scales were computed according to published procedures for each scale. The items of the PAIS have been calibrated using Samejima's

graded IRT model [15, 16], which describes the probability of each item response as a function of a set of item parameters (a_s and c_s), and θ , the latent variable measured by the scale, as follows: The conditional probability of response $u = 0, 1, \dots, 4$ is

$$T_u(\theta) = T_u^*(\theta) - T_{u+1}^*(\theta)$$

in which $T_u^*(\theta)$ is a curve tracing the probability of a response in category u or higher:

$$T_0^*(\theta) = 1, \\ T_u^*(\theta) = \frac{1}{1 + \exp[-(a'\theta + c_u)]}$$

for $u = 1, 2, \dots, m - 1$, and

$$T_m^*(\theta) = 0$$

For calibration of the PAIS, θ is unidimensional. Values of the item parameters (a_s and c_s) are tabulated in the original description of the PAIS [7]. For use in this research, *expected a posteriori* (EAP) estimates for response patterns [19] were computed for each respondent. Response-pattern EAP estimates may be computed in the presence of any pattern of missing item responses, so they are available for all respondents.

Descriptive statistics were computed for the scores on all seven asthma scales using all available data.

Linkage

Holland provides a modern framework for test score linking; he writes that “*linking* refers to the general class of transformations between the scores from one test and those of another, ... linking methods can be divided into three basic categories called *predicting*, *scale aligning*, and *equating*” [20]. For linking scores from disparate tests, such as the PAIS and the PedsQL Asthma Symptoms Scale (hereafter the PedsQL Symptoms Scale), only *predicting* scores on the PAIS scale (from the PedsQL Symptoms Scale), and *aligning* the two scales, are viable candidates.

A commonly used method of scale aligning has been *calibration*, which uses IRT to place the items from each of two scales on the same metric. After that is done, standard computation of IRT scaled scores from any subset of the items (which includes all of the items on one scale, or the other) yields comparable scores. However, calibration has heretofore been limited to situations in which a unidimensional IRT model is suitable for the aggregate set of items from both scales—that is, both scales measure exactly the same construct.

For two scales that measure different (if highly related) constructs, *predicting* scores on one scale from those of the other has been the norm. Such predictions are based on regression models, but often the regression model is elaborated to attend to the fact that the prediction is not so much a point as it is a distribution. When the entire conditional (predicted) distribution is considered, predicting scores on one scale from the score on the other measure is called *projection*. Usually projection has been based on variations of standard regression models, which consider the values of the predictor variable(s) fixed.

Because the greatest threat to the validity of test linkage that is readily detectable empirically is a lack of invariance of the relation between the two scales across subgroups of the population, Dorans and Holland recommend the use of a *root mean squared deviation* (RMSD) statistic to check invariance of subgroup differences between scales to be linked [21]. Dorans and Holland found values of RMSD of 1–8% to be associated with linkages that have proved useful; those were associated with test scores correlated 0.83–0.92. In this research, the RMSD statistics were computed for each scale for the differences between boys and girls, and between younger and older children.

Calibrated projection—*Calibrated projection* is a new statistical procedure that uses IRT to link two measures, without considering the scores on the predictor scale to be fixed, and without the demand of conventional calibration that the two are measures of the same construct. In calibrated projection, the multidimensional version of Samejima's graded model is fitted to the item responses from the two measures: θ_1 represents the underlying construct measured by the PAIS, with estimated slopes a_1 for each of the PAIS items and fixed values of 0.0 for the items of the PedsQL Symptoms Scale. θ_2 represents the underlying construct measured by the PedsQL Symptoms Scale. The correlation between θ_1 and θ_2 is estimated simultaneously with the a s and c s for the items on the two scales together. The model may also include additional latent variables, if required to fit LD that may arise between pairs of items within or across scales.

Subsequently, the MIRT model may be used to provide IRT scaled score estimates on the scale of one measure, using only the item responses from the other measure. Figure 1 illustrates calibrated projection: The x -axis variable is θ_2 , the underlying construct measured by the PedsQL Symptoms Scale, and the y -axis variable is θ_1 , the underlying construct measured by the PAIS. The two latent variables are highly correlated, as indicated by the density ellipses around the regression line. Given the item responses on the PedsQL Symptoms Scale, IRT methods may be used to compute the implied distribution on θ_2 ; two of those are shown along the x -axis in Fig. 1, for summed scores of 13 and 44 on the PedsQL Symptoms Scale. The estimated relation between θ_1 and θ_2 may then be used to project those distributions onto the y -axis, to yield the implied distributions on θ_1 , the PAIS construct, given those scores on the PedsQL Symptoms Scale.

The means of the implied distributions on the θ dimensions are the IRT-based scaled scores, and the standard deviations of those distributions are reported as the standard errors of those scores. The projection links the scales in the sense that each score on the PedsQL Symptoms Scale yields a score on the PAIS metric.

Estimation of the MIRT parameters and computation of the scaled scores was done using the IRTPRO software [22], implementing two-tier methods [23].

Results

Sample description

The two test forms containing PROMIS pediatric asthma items were completed by a diverse sample of 622 respondents. Demographic statistics for the entire PROMIS pediatric calibration sample have been described elsewhere [3], as have those for the asthma form respondents [7]. Specifically for the asthma samples, the children participating in the survey had a range of asthma severity and asthma control, with approximately half reporting symptoms characteristic of children with mild persistent asthma or more severe asthma.

Comparisons among the asthma HRQoL measures

Descriptive statistics for scores on the PAIS, the PedsQL Asthma Symptoms, Treatment, Worry, and Communication Scales, and the DISABKIDS Asthma Impact and Worry Scales are provided in Table 1. For the PAIS the sample size is 622, because the IRT response pattern EAP estimates could be computed for all respondents. For the other scales, the sample size varies around 300: approximately half the sample responded to the form that included those legacy scales, and a few respondents had missing scores on the PedsQL or DISABKIDS scales.

The mean and standard deviation of the PAIS scores are close to the standardizing values of 50 and 10, because the original reference sample makes up the majority of the current data. The means and standard deviations for the PedsQL scales are close to those reported for the standardization sample [11], with the possible exception of the Worry scale (mean 70.7 in this sample, and 76.3 in the standardization sample). Coefficient α reliability values are higher for the PedsQL scales in this sample (Table 1) than they were in the original reference sample (symptoms = 0.85; treatment = 0.58; worry = 0.72; communication = 0.70). The values of coefficient alpha for the DISABKIDS scales are consistent with the range of alphas observed across several countries (impact range = 0.72–0.91; worry range = 0.61–0.88). [18].

All correlations of the PAIS scores with those from the other scales are negative, because the PAIS is scored with “asthma impact” (low HRQoL) high, while the other scales are scored so that high HRQoL corresponds with high scores. The PAIS scores are correlated -0.85 with PedsQL Symptoms Scale scores; the disattenuated value of the correlation is -0.94 , suggesting that the PAIS and the PedsQL Symptoms Scale measure nearly the same underlying construct, so linkage between the two scales may be useful. The correlations of the PAIS and PedsQL Symptoms Scale scores with the scores on the remaining five scales are very similar to each other, providing further evidence that those two scales measure approximately the same construct. Those correlations are generally lower, as are the correlations among the remaining five scales, suggesting that distinct aspects of HRQoL are measured by each of the remaining scales, and providing evidence of divergent validity for the PAIS and the PedsQL Symptoms Scale.

The rightmost columns of Table 1 show the standardized mean differences (effect sizes) for girls vs. boys, and younger (ages 8–11) vs. older (ages 12–17) respondents for each scale. Ignoring the (necessarily) opposite signs of the differences, those values are especially similar for the PAIS and PedsQL Symptoms Scale, again suggesting that linkage of the two metrics may be viable. RMSD values are 6% for the sex-related difference and 0 for the age difference, within the range of values for linkages between educational measures considered plausible [21].

Linking the PedsQL asthma symptoms scale with the PAIS metric

Calibrated projection was used to link scores on the PedsQL Symptoms Scale to the IRT metric of the PAIS. The projection is based on a MIRT model with item parameters tabulated in Table 2 (for the discrimination (a) parameters) and 3 (for the intercept (c) parameters). In this model, θ_1 represents the underlying construct measured by the PAIS, with estimated slopes a_1 for each of the PAIS items and corresponding fixed values of 0.0 for the items of the PedsQL Symptoms Scale. θ_2 represents the underlying construct measured by the PedsQL Symptoms Scale. The item parameters are tabulated on the conventional IRT scale for which the latent variables are standard (multivariate) normal; the estimated correlation between θ_1 and θ_2 is 0.96. The model also includes six additional uncorrelated latent variables, θ_3 to θ_8 ; each of those latent variables has non-zero (equal)

slopes (a) for a particular pair of items that are nearly identical on the PAIS and the PedsQL Symptoms Scale. Those pairs of items arose from the fact that six PAIS items were derived by rewording corresponding PedsQL Symptoms Scale items. Some degree of local dependence (LD) is induced by those nearly identical item pairs, and the additional factors fit that LD so that the model is not misspecified. All of the a parameter estimates in Table 2 exceed six times their standard errors, indicating that the corresponding relationships differ significantly from zero.

In Table 3 the items are sorted within each of the two scales by their degree of severity. For the graded item response model, an overall index of severity is the parameter b , which is the average of the item's threshold parameters $b_u = -c_u/a$ [25].

An alternative procedure using joint unidimensional calibration of the two scales was considered and rejected. The unidimensional model replaced the first two factors in the model described above with one-dimensional θ for both the PAIS and the PedsQL Symptoms Scale together. The likelihood ratio test for the difference in fit between the unidimensional model and the two-dimensional model was $\chi^2(1) = 50.9$, $P < 0.0001$, rejecting the unidimensional model (or joint calibration as a strategy).

Using the MIRT graded model parameter estimates in Tables 2 and 3 for the items of the PedsQL Symptoms Scale, *expected a posteriori* (EAP) values of θ_1 , the PAIS latent variable, were computed for each summed score on the PedsQL Symptoms Scale. These values are in Table 4, converted to the PAIS reporting scale (with a mean of 50 and standard deviation of 10 in the original PAIS calibration sample). Table 4 may be used as a scoring table to convert scores on the PedsQL Symptoms Scale into scores on the PAIS metric.

Discussion

These results represent evidence of convergent and divergent validity for the PAIS; its scores measure very nearly the same construct as that measured by the PedsQL Symptoms Scale; and that construct is relatively distinct from those measured by the PedsQL Asthma Treatment, Worry, and Communications scales, and the DISABKIDS Asthma Worry and Impact scales (although the PAIS construct is closer to the latter).

Calibrated projection has been used to provide linkage of scores on the PedsQL Symptoms Scale with the metric of the PAIS. This linkage takes into account the slight difference between the constructs measured by the two scales, as well as LD that arises from six pairs of items with very similar members on both scales.

Linkage of the PedsQL Symptoms Scale with the PAIS does not reduce the usefulness of the PAIS. For research in which the PAIS measures the intended construct, the PAIS short form provides more precise measurement with eight items than does the PedsQL eleven-item Symptoms Scale; in addition, the PAIS comprises a bank of IRT-calibrated items that may be used to construct other forms or in computerized adaptive testing. So the PAIS itself would be preferred for new research involving its construct.

However, there is a large body of research that has already used the PedsQL Symptoms Scale; that research can be retro-scored onto the PAIS metric for comparability. In addition, some research may have use for the entire four-scale PedsQL Asthma Module; in such contexts the PedsQL Asthma Module may be a more useful choice, with linkage used to provide scores comparable to those obtained with the PAIS in other research.

Test linking has not received extensive treatment in the health outcomes literature. Among health outcomes measures, researchers have linked alternate forms of similar measures of

functional health status [26] and self-regulation [27]. Others have attempted linkage of disparate health outcomes measures [28–30]. Those earlier attempts at linking demonstrate the power of IRT, but they also raise questions about when and what to link in measures of health outcomes (in particular, the feasibility and utility of linking apparently dissimilar instruments). Calibrated projection, as used in this study, makes fewer demands on the data (i.e., it does not require unidimensionality across scales). In addition, this research used RMSD statistics to evaluate the viability of linkage, something that has not previously been done with health outcomes scales.

Calibrated projection is not a panacea that permits any scale to be linked with any other. A limitation of linking is that it only makes sense between scales that measure very nearly the same construct. In this research, only one of the legacy scales was linked to the metric of the PAIS; the other five scales appear to measure constructs too different from that of the PAIS to link. An additional limitation of calibrated projection is that it requires item-level data on both scales from the same sample of respondents. Other linkage methods can use data collected with other designs, such as equivalent groups, but those procedures do not permit the validity of the linkage to be empirically checked. Further, projected scores from some other scale are less precise—that is, they have larger standard errors—than scores on the primary scale.

Overall this research provides results illustrating the usefulness of both the PAIS in particular and the PROMIS approach to integrated measurement of HRQoL in general. We anticipate that calibrated projection may be useful to link other legacy scales to the PROMIS metrics as well, permitting health outcomes research to integrate findings from earlier research that used previously existing scales with newer research that uses the more modern scales developed by the PROMIS network.

Acknowledgments

We would like to acknowledge the contribution of Harry A. Guess, MD, PhD to the conceptualization and operationalization of this research prior to his death. We are grateful to Li Cai for the theoretical development and implementation in software of the two-tier methods for item parameter estimation and the computation of scaled scores, and for his advice on their use in this project. This work was funded by the National Institutes of Health through the NIH Roadmap for Medical Research, Grant 1U01AR052181-01, and by SBIR contract HHSN-2612007-00013C with the National Cancer Institute of the National Institutes of Health. Information on the Patient-Reported Outcomes Measurement Information System (PROMIS) can be found at <http://nihroadmap.nih.gov/> and <http://www.nihpromis.org>.

References

1. Ader DN. Developing the patient-reported outcomes measurement information system (PROMIS). *Medical Care*. 2007; 45(1):S1–S2. [PubMed: 18027399]
2. Cella D, Yount S, Rothrock N, et al. The patient-reported outcomes measurement information system (PROMIS): Progress of an NIH roadmap cooperative group during its first two years. *Medical Care*. 2007; 45(Suppl 1):S3–S11. [PubMed: 17443116]
3. Irwin DE, Stucky BD, Thissen D, et al. Sampling plan and patient characteristics of the PROMIS pediatrics large-scale survey. *Quality of Life Research*. 2010; 19:585–594. [PubMed: 20204706]
4. Irwin D, Stucky BD, Thissen D, et al. An item response analysis of the pediatric PROMIS anxiety and depressive symptoms scales. *Quality of Life Research*. 2010; 19:595–607. [PubMed: 20213516]
5. Varni JW, Stucky BD, Thissen D, et al. PROMIS pediatric pain interference scale: An item response theory analysis of the pediatric pain item bank. *Journal of Pain*. 2010; 11:1109–1119. [PubMed: 20627819]
6. DeWitt EM, Stucky BD, Thissen D, et al. Construction of the PROMIS Pediatric Physical Function Scales: Built using item response theory. *Journal of Clinical Epidemiology*. in press.

7. Yeatts K, Stucky BD, Thissen D, et al. Construction of the pediatric asthma impact scale (PAIS) for the patient reported outcomes measurement information system (PROMIS). *Journal of Asthma*. 2010; 47:295–302. [PubMed: 20394514]
8. Chan KS, Mangione-Smith R, Burwinkle TM, et al. The PedsQL™: Reliability and validity of the short-form generic core scales and asthma module. *Medical Care*. 2005; 43:256–265. [PubMed: 15725982]
9. Guyatt GH, Juniper EF, Griffith LE, et al. Children and adult perceptions of childhood asthma. *Pediatrics*. 1997; 99:165–168. [PubMed: 9024440]
10. Juniper EF, Guyatt GH, Feeny DH, et al. Measuring quality of life in children with asthma. *Quality of Life Research*. 1997; 5:35–46. [PubMed: 8901365]
11. Varni JW, Burwinkle TM, Rapoff MA, et al. The PedsQL™ in pediatric asthma: Reliability and validity of the pediatric quality of life inventory™ generic core scales and asthma module. *Journal of Behavioral Medicine*. 2004; 27:297–318. [PubMed: 15259457]
12. Irwin DE, Varni JW, Yeatts K, et al. Cognitive interviewing methodology in the development of a pediatric item bank: A patient reported outcomes measurement information system (PROMIS) study. *Health Qual Life Outcomes*. 2009; 7:1–10. [PubMed: 19134191]
13. DeWalt DA, Rothrock N, Yount S, et al. Evaluation of item candidates: The PROMIS qualitative item review. *Medical Care*. 2007; 45(Suppl 1):S12–S21. [PubMed: 17443114]
14. Walsh TR, Irwin DE, Meier A, et al. The use of focus groups in the development of the PROMIS pediatrics item bank. *Quality of Life Research*. 2008; 17:725–735. [PubMed: 18427951]
15. Samejima F. Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monograph*. 1969; 34(Monograph Suppl):1–100.
16. Samejima, F. Graded response model. In: van der Linden, WJ.; Hambleton, RK., editors. *Handbook of modern item response theory*. New York: Springer; 1997. p. 85-100.
17. PROMIS Assessment Center web site. [Accessed 3 Sept 2010] 2010. Available at <http://www.nihpromis.org/Web%20Pages/Assessment%20Center.aspx>
18. Baars RM, Atherton CI, Koopman HM, et al. The European DISABKIDS project: Development of seven condition-specific modules to measure health related quality of life in children and adolescents. *Health Qual Life Outcomes*. 2005; 3:1–9. [PubMed: 15634354]
19. Thissen, D.; Nelson, L.; Rosa, K., et al. Item response theory for items scored in more than two categories. In: Thissen, D.; Wainer, H., editors. *Test scoring*. Mahwah, NJ: Lawrence Erlbaum Associates; 2001. p. 141-186.
20. Holland, PW. Framework and history for score linking. In: Dorans, NJ.; Pommerich, M.; Holland, PW., editors. *Linking and aligning scores and scales*. New York: Springer; 2007. p. 5-30.
21. Dorans NJ, Holland PW. Population invariance and the equitability of tests: Basic theory and the linear case. *Journal of Educational Measurement*. 2000; 37:281–306.
22. IRTPRO. Flexible professional item response theory modeling for patient-reported outcomes [Computer program]. Chicago, IL: SSI International; 2011.
23. Cai L. A two-tier full-information item factor analysis model with applications. *Psychometrika*. 2010; 75:581–612.
24. Green BF, Bock RD, Humphreys LG, et al. Technical guidelines for assessing computerized adaptive tests. *Journal of Educational Measurement*. 1984; 21:347–360.
25. Muraki E. Fitting a polytomous item response model to likert-type data. *Applied Psychological Measurement*. 1990; 14:59–71.
26. McHorney CA, Cohen AS. Equating health status measures with item response theory: Illustrations with functional status items. *Medical Care*. 2000; 38(Suppl):II43–II59. [PubMed: 10982089]
27. Masse LC, Allen D, Wilson M, et al. Introducing equating methodologies to compare test scores from two different self-regulation scales. *Health Education Research*. 2006; 21(Suppl 1):i110–i120. [PubMed: 16926231]
28. Fisher WP Jr, Eubanks RL, Marier RL. Equating the MOS SF36 and the LSU HSI physical functioning scales. *Journal of Outcome Measurement*. 1997; 1:329–362. [PubMed: 9661727]
29. Bond, TG.; Fox, CM. *Applying the Rasch model: Fundamental measurement in the human sciences*. Mahwah, NJ: Lawrence Erlbaum Associates; 2007.

30. Holzner B, Bode RK, Hahn EA, et al. Equating EORTC QLQ-C30 and FACT-G scores and its use in oncological research. *European Journal of Cancer*. 2006; 42:3169–3177. [PubMed: 17045472]

Fig. 1.

The x -axis variable is θ_2 , the underlying construct measured by the PedsQL Asthma Symptoms Scale, and the y -axis variable is θ_1 , the underlying construct measured by the PAIS (both in standard units). The correlation between the two latent variables is indicated by the density ellipses around the correlation line. The implied distribution on θ_2 for summed scores of 13 and 44 on the PedsQL Asthma Symptoms Scale are shown along the x -axis, along with the corresponding implied distributions on θ_1 , the PAIS construct along the y -axis

Table 1
Descriptive statistics for the pediatric asthma impact scale (PAIS), the PedsQL™ asthma symptoms scale (PedsQL S), treatment (PedsQL T), worry (PedsQL W), and communication (PedsQL C) scales, and the DISABKIDS asthma impact (DISAB I) and worry (DISAB W) scales

	<i>N</i>	Mean	SD	Correlations (reliability on the diagonal)							Standardized difference	
				1	2	3	4	5	6	7	Girls-Boys	Young-Old
1. PAIS	622	50.9	9.4	0.93							0.16	-0.13
2. PedsQL S	303	68.2	19.7	-0.85	0.88						-0.28	0.13
3. PedsQL T	302	79.5	15.0	-0.49	0.56	0.77					-0.18	0.05
4. PedsQL W	302	70.7	24.3	-0.59	0.61	0.59	0.73				-0.30	0.18
5. PedsQL C	304	72.9	25.5	-0.44	0.45	0.52	0.58	0.81			-0.25	0.16
6. DISAB I	295	66.0	23.9	-0.71	0.72	0.49	0.63	0.51	0.85		-0.25	0.21
7. DISAB W	300	76.2	23.5	-0.62	0.63	0.46	0.64	0.50	0.68	0.86	-0.36	0.03

Reliability is marginal reliability Green et al. [24] for the IRT-scored PAIS, and coefficient a for the traditionally scored PedsQL and DISABKIDS scales

Table 2
Discrimination (α) parameters for the graded MIRT model used in calibrated projection of PAIS scores from responses to the PedsQL™ asthma symptoms scale

Item	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8
PAIS—17 items								
8 I had trouble breathing because of my asthma	2.92	0	0	0	0	0	0	0
9 My asthma bothered me	2.65	0	0	0	0	0	0	0
10 It was hard for me to play sports or exercise because of my asthma	2.03	0	0	0	0	0	0	0
6 My chest felt tight because of my asthma	2.00	0	0	0	0	0	0	0
16 I had trouble walking because of my asthma	1.95	0	0	0	0	0	0	0
5 I felt scared that I might have trouble breathing because of my asthma	1.93	0	0	0	0	0	0	0
4 My body felt bad when I was out of breath	1.84	0	0	0	0	0	0	0
3 I was bothered by the amount of time I spent wheezing	1.78	0	0	0	0	0	0	0
2 My asthma bothered me when I was with my friends	1.53	0	0	0	0	0	0	0
17 I missed school because of asthma	1.47	0	0	0	0	0	0	0
1 I got tired easily because of my asthma	1.37	0	0	0	0	0	0	0
13 I had asthma attacks	3.89	0	3.85	0	0	0	0	0
12 I felt wheezy because of my asthma	3.86	0	0	2.37	0	0	0	0
15 It was hard to take a deep breath because of my asthma	3.49	0	0	0	2.04	0	0	0
7 It was hard for me to play with pets because of my asthma	2.99	0	0	0	0	4.06	0	0
11 I had trouble sleeping at night because of my asthma	2.79	0	0	0	0	0	1.66	0
14 I coughed because of my asthma	2.76	0	0	0	0	0	0	2.52
PedsQL asthma symptoms scale—11 items								
20 I have asthma attacks	0	4.09	3.85	0	0	0	0	0
19 I feel wheezy	0	3.90	0	2.37	0	0	0	0
24 It is hard to take a deep breath	0	3.42	0	0	2.04	0	0	0
27 It is hard for me to play with pets	0	3.36	0	0	0	4.06	0	0
26 I wake up at night with trouble breathing	0	3.11	0	0	0	0	1.66	0
23 I cough	0	2.63	0	0	0	0	0	2.52
18 My chest hurts or feels tight	0	2.31	0	0	0	0	0	0
22 I get out of breath	0	2.25	0	0	0	0	0	0
28 It is hard for me to play outside	0	2.19	0	0	0	0	0	0

Item	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8
21 I get scared when I have asthma attacks	0	1.70	0	0	0	0	0	0
25 I have a stuffy or runny nose	0	1.07	0	0	0	0	0	0

Table 3
Intercept (c) parameters for the graded MIRT model used in calibrated projection of PAIS scores from responses to the PedsQL™ asthma symptoms scale; the items are sorted within scale by the parameter b , an index of severity

Item	c_1	c_2	c_3	c_4	b
PAIS—17 items					
14 I coughed because of my asthma	3.48	1.33	-1.76	-3.99	0.09
6 My chest felt tight because of my asthma	1.95	0.78	-1.37	-2.76	0.18
8 I had trouble breathing because of my asthma	2.44	0.87	-2.42	-4.37	0.30
4 My body felt bad when I was out of breath	1.66	0.51	-1.58	-2.86	0.31
9 My asthma bothered me	2.30	0.71	-2.25	-4.00	0.31
5 I felt scared that I might have trouble breathing because of my asthma	1.45	0.18	-2.03	-3.59	0.52
15 It was hard to take a deep breath because of my asthma	1.96	-0.19	-3.46	-5.65	0.53
1 I got tired easily because of my asthma	1.34	0.25	-1.57	-2.98	0.54
12 I felt wheezy because of my asthma	2.17	-0.21	-4.30	-7.23	0.62
10 It was hard for me to play sports or exercise because of my asthma	0.92	-0.16	-2.31	-3.58	0.63
3 I was bothered by the amount of time I spent wheezing	0.87	-0.28	-2.22	-3.54	0.73
11 I had trouble sleeping at night because of my asthma	0.13	-1.43	-3.74	-5.36	0.93
2 My asthma bothered me when I was with my friends	0.20	-0.81	-2.87	-4.44	1.29
13 I had asthma attacks	-1.51	-3.88	-7.57	-10.26	1.49
17 I missed school because of asthma	-0.56	-1.31	-2.97	-3.94	1.49
16 I had trouble walking because of my asthma	-1.19	-2.27	-3.71	-5.32	1.60
7 It was hard for me to play with pets because of my asthma	-1.47	-3.73	-6.60	-8.75	1.72
PedsQL asthma symptoms scale—11 items					
23 I cough	4.44	2.17	-1.70	-4.08	-0.08
25 I have a stuffy or runny nose	1.64	0.55	-1.29	-2.29	0.32
22 I get out of breath	2.05	0.69	-2.14	-3.82	0.36
24 It is hard to take a deep breath	1.79	-0.65	-4.59	-7.02	0.77
19 I feel wheezy	1.50	-1.05	-5.83	-8.24	0.87
18 My chest hurts or feels tight	0.77	-0.56	-3.19	-5.50	0.92
26 I wake up at night with trouble breathing	-0.17	-1.88	-4.11	-5.82	0.96
28 It is hard for me to play outside	0.14	-1.18	-3.44	-5.02	1.08
21 I get scared when I have asthma attacks	-0.48	-1.20	-2.84	-3.68	1.21

Item	c_1	c_2	c_3	c_4	b
27 It is hard for me to play with pets	-1.91	-4.02	-7.34	-9.21	1.67
20 I have asthma attacks	-2.04	-4.89	-9.10	-12.15	1.72

Table 4
Expected a posteriori (EAP) values of the PAIS latent variable for each summed score on the PedsQL™ asthma symptoms scale, on the PAIS reporting scale

Summed score	Original PedsQL score	PAIS scale	
		<i>T</i> -score	<i>T</i> -score SE
0	100.0	31	6
1	97.7	35	6
2	95.5	37	5
3	93.2	39	5
4	90.9	40	5
5	88.6	42	5
6	86.4	43	5
7	84.1	45	4
8	81.8	46	4
9	79.5	47	4
10	77.3	48	4
11	75.0	49	4
12	72.7	50	4
13	70.5	51	4
14	68.2	52	4
15	65.9	52	4
16	63.6	53	4
17	61.4	54	4
18	59.1	55	4
19	56.8	56	4
20	54.5	57	4
21	52.3	57	4
22	50.0	58	4
23	47.7	59	4
24	45.5	60	4
25	43.2	60	4
26	40.9	61	4
27	38.6	62	4
28	36.4	63	4
29	34.1	64	4
30	31.8	64	4
31	29.5	65	4
32	27.3	66	4
33	25.0	67	4
34	22.7	68	4
35	20.5	69	4
36	18.2	70	4

Summed score	Original PedsQL score	PAIS scale	
		<i>T</i> -score	<i>T</i> -score SE
37	15.9	71	4
38	13.6	72	4
39	11.4	73	4
40	9.1	74	4
41	6.8	75	5
42	4.5	77	5
43	2.3	78	5
44	0.0	81	5