Agreement was not found in adolescents' quality of life rated by parents and adolescents

I-Chan Huang[a,*], **Elizabeth A. Shenkman**[a], **Walter Leite**[b], **Caprice A. Knapp**[a], **Lindsay A. Thompson**[c], and **Dennis A. Revicki**[d]

[a]Department of Epidemiology and Health Policy Research, College of Medicine, University of Florida, Gainesville, FL, USA

[b]Department of Educational Psychology, College of Education, University of Florida, Gainesville, FL, USA

[c]Department of Pediatrics, School College of Medicine, University of Florida, Gainesville, FL, USA

[d]The Center for Health Outcomes Research, the United BioSource, Bethesda, MD, USA

Abstract

Objective—Limited evidence examines agreement between the ratings of adolescents' health-related quality of life (HRQOL) by parents and adolescents especially accounting for measurement properties. This study aimed to test whether agreement exists between the dyads using a methodology of measurement invariance.

Study Design and Setting—We conducted a telephone survey to collect data from parents and their adolescents enrolled in Florida's Children's Medical Services (376 pairs) using the PedsQL 4.0. We used multi-group confirmatory factor analysis to test measurement invariance, including configural (equivalent HRQOL construct), metric (equivalent item–domain associations), scalar (equivalent starting value of ratings), and residual (equivalent nonsystematic measurement errors of ratings) invariance. We also investigated correlates of discrepancies in the dyadic ratings.

Results—There were equivalent HRQOL constructs and item–domain associations between the dyads. However, some items show different starting values and nonsystematic errors in the dyadic ratings. After adjusting for noninvariant items, adolescents reported significant higher HRQOL scores than parents in all domains ($P < 0.05$). Parents' rating of adolescents' health significantly contributed to discrepancies in the dyadic ratings ($P < 0.05$).

Conclusions—Adolescents rated HRQOL higher than their parents. This discrepancy was associated with severe health conditions. Without assuring measurement invariance, comparisons of the dyadic HRQOL ratings can be misleading.

Keywords

*Corresponding author. P.O. Box 100147, Epidemiology and Health Policy Research, University of Florida, College of Medicine, Gainesville FL 32610-0147, USA. Tel.: 352-265-7220 ext. 86287; fax: 352-265-7221. ichuang@ufl.edu I.-C. Huang.

1. Introduction

Pediatric health-related quality of life (HRQOL) is an important indicator, complementing objective clinical measures, to assess pediatric health outcomes [1,2]. Parents' ratings of pediatric HRQOL are often used as proxies, especially if children are too young or cognitively impaired to understand HRQOL concepts [3,4]. The use of a parent's rating, however, raises the question of what level of agreement and discrepancy exists between the dyadic HRQOL reports [4–6].

A recent review suggests that the agreement between parents' and children's HRQOL ratings is inconsistent [7]. Some studies show that parents report higher HRQOL scores than do their children across all domains [6,8]. However, others show that compared to children's self-ratings, parents tend to report better physical and social functioning, but more depression and bodily pain [7]. In contrast to children's own ratings, parents may underestimate HRQOL for sick children [9], but overestimate for healthy children [6,10].

In extant studies, parents' and children's pediatric HRQOL ratings are often compared using *t*-tests or regression methods [6,11–14]. Methodologically, these comparisons are limited because they cannot assure whether HRQOL measures between the dyads are comparable. It is possible that parents and children perceive the construct of HRQOL or meanings of specific items differently. If we are unable to assure the measurement constructs and metrics between the parents' and children's ratings are comparable (i.e., measurement invariance), our dyadic comparisons will be meaningless and interpretations may be misleading [15,16].

Meredith proposed an analytic framework explicitly testing whether measurements between groups are invariant, and included the components of configural, metric, scalar, and residual measurement invariance [15]. Configural invariance means the conceptual frameworks between parents' and children's assessments of pediatric HRQOL are equivalent. Metric invariance means, given configural invariance, the associations of items with the underlying HRQOL domain are equivalent in dyads. Scalar metric means, given metric invariance, parents' and children's ratings of the item are based on the same starting value. Residual invariance means, given scalar invariance, parents and children are subject to the same form of nonsystematic measurement errors in item response. In HRQOL studies, the application of measurement invariance methods is limited [16,17]. Most of the HRQOL studies merely emphasized the metric invariance [18–22] although residual invariance is a necessary condition for meaningful comparisons among groups [23–25].

The overall aim of this study was to investigate measurement invariance in adolescents' HRQOL rated by parents and their adolescents. We focused on adolescents between 15 and 18 years who have special health care needs. Our selection of this age group is important because they are involving in "transition planning," where providers begin to assist adolescents in making more independent decision making, in conjunction with their parents, about their health care when adolescents are becoming young adults [26]. During the transition-planning phase, input of adolescents' HRQOL can be useful to promote shared decision making with the provider and to plan care. We used Meredith's analytic framework [15] to test measurement invariance. We also examined, after adjusting for noninvariant items, whether the discrepancy persists between the dyadic ratings. Finally, we investigated the correlates contributing to the discrepancy in the dyadic ratings.

2. Methods

2.1. Data collection, sample, and data sources

This is a cross-sectional study using data collected from the parents and their adolescents enrolled in 2005 Florida's Children's Medical Services Network program. Children's Medical Services Network is Florida's Title V Program for Children with Special Health Care Needs and is designed to provide specialty health care for children who are less than 19 years old. All children in this sample were also enrolled in Medicaid.

We identified 700 families and adolescents from enrollment files maintained by Children's Medical Services Network. We conducted a telephone survey using a statewide random sample of families that had an adolescent 15 through 18 years old and lived with them between 12/2005 and 03/2006. Multiple callbacks (at maximum of 10 times) were performed if phone numbers were busy or not answered. Twenty-five percent of parents reported that their children were physically or mentally unable to complete the survey and 6% refused to allow their adolescents to be interviewed. Additionally, 4% of the cases adolescents subsequently refused to participate or their parents refused after initially granting permission. The overall response rate was about 65%, which is similar to other telephone surveys with Medicaid populations [27,28]. The study sample consisted of 376 matched pairs of parents and their adolescents who completed the survey. However, 12 dyads that did not complete HRQOL survey or had over 50% of missing items were further excluded, leaving 364 dyads for final analyses.

Florida's Agency for Health Care Administration provided claims and encounter data for these children for the year preceding the survey. The claims and encounter data were used to classify the children's health status. Specifically, we used the clinical risk groups (CRGs) systems [29–31], an ICD-9-CM-based severity classification system that groups diagnoses into the following categories based on predicted resource consumption: (1) nonacute health care needs, (2) significant acute conditions, (3) minor chronic conditions (e.g., attention deficit/hyperactivity disorder), (4) moderate chronic conditions (e.g., asthma), or (5) major chronic conditions (e.g., cystic fibrosis) [29].

2.2. HRQOL measure

We used the PedsQL 4.0 to assess the child's HRQOL [32,33]. An adolescent self-report and a parent proxy-report module, designed for adolescents between 13 and 18 years old, was used. The PedsQL consists of 23 items measuring four domains—physical functioning (eight items), emotional functioning (five items), social functioning (five items), and school functioning (five items). Each item has five response categories. A specific domain score is calculated for corresponding items, ranging from 0 (worst HRQOL) to 100 (best HRQOL). The PedsQL has acceptable psychometric properties and is suitable for healthy children and children with acute and chronic conditions [33,34].

2.3. An analytic framework for testing measurement invariance

HRQOL is a latent variable that cannot be directly observed, but can be measured using observed variables (i.e., designed items). A latent variable is assumed to influence a subject's response to the designed items. We applied the methodology of multiple-group confirmatory factor analysis (MG-CFA) to test measurement invariance in HRQOL measure across groups. We conducted serial tests of nested models, which began with the less constrained model (i.e., configural invariance), then sequentially placed cross-group equality constraints on the parameters, and ended up with the most constrained model (i.e., residual invariance) (see Fig. 1) [15,35].

Table 1 shows the interpretations for invariance tests using MG-CFA. Specifically, configural measurement invariance means that the conceptual frameworks of HRQOL used by parents and adolescents to rate the adolescent's HRQOL are equivalent. MG-CFA investigates this invariance by testing equality in the item clusters of a specific HRQOL domain between the dyads.

Metric measurement invariance means, given the configural invariance, the magnitudes of the associations between items and underlying HRQOL domain perceived by parents and adolescents are comparable. If metric invariance is achieved, the measurement units for the dyadic HRQOL ratings are equivalent because the differences in item responses associated with the differences in domain scores are equal between the two groups. MG-CFA investigates this invariance by testing equality in factor loadings of items between the dyads.

Scalar measurement invariance means, given the metric invariance, the item scores rated by parents and adolescents are based on the same starting value. It is possible that parents and adolescents may report different scores for the same items, in part due to differential acquiescence response style (e.g., higher expectation by parents for their child's health than adolescents themselves) [36,37]. MG-CFA investigates this invariance by testing equality in the intercepts of item ratings between the dyads.

Residual measurement invariance means, given the scalar invariance, parents and adolescents have a similar degree of nonsystematic measurement errors in item responses. Very often, residuals of the item responses are correlated rather than conditionally independent. Therefore, the group differences cannot be truly attributed to differences in the common HRQOL factor [15,23]. MG-CFA investigates this invariance by testing equality in the residuals of item responses between the dyads.

Several indicators were applied to interpret the goodness of model fit for the MG-CFA, including change in Satorra–Bentler scaled χ^2 ($\Delta SB\chi^2$), root mean square error of approximation (RMSEA), comparative fit index (CFI), and change in CFI (ΔCFI) [38]. A model showing nonsignificant $\Delta SB\chi^2$, RMSEA ≤ 0.06, CFI ≥ 0.95, and $\Delta CFI \leq 0.01$ (i.e., the decrease of CFI for a more constrained vs. a less constrained model ≤ 0.01) will be considered as an acceptable model fit. It means that we should not reject the null hypothesis of measurement invariance between the dyads.

2.4. Partial measurement invariance

We specifically applied the concept of partial measurement invariance in the tests [24,39–41]. If all items in a HRQOL domain fail to demonstrate a specific level of measurement invariance (e.g., metric noninvariance), we sequentially removed the equality constraint on the parameter for each item between two groups in the MG-CFA to assess the degree of model improvement. If partial invariance was present, we calibrated items (i.e., allowing parameters of noninvariant items acting differently between groups, but treating parameters of invariant items equally between groups in MG-CFA) to calculate latent domain scores for parents' and adolescents' HRQOL.

2.5. Dyadic differences in HRQOL ratings

We compared the discrepancies in the dyadic ratings of HRQOL before and after adjusting for items with measurement noninvariance. The magnitude of the discrepancy was measured using effect size (defined as the difference in the dyadic HRQOL scores divided by a polled SD of both groups). We defined the magnitudes of < 0.2, $0.2–0.49$, $0.5–0.79$ and ≥ 0.8 as negligible, small, moderate, and large effect size, respectively [42,43].

We compared parents' and adolescents' domain scores through the Generalized Estimating Equations regression to explicitly addresses the influence of correlated dyadic HRQOL scores in the calculation of standard errors of regression coefficients [44]. We also investigated which factors were significantly associated with HRQOL ratings and the discrepancy in the dyadic ratings, including adolescents' age, gender, race (White, Black, Hispanic, and others), health status, and parents' level of education [4,45,46]. We measured adolescents' health status using CRGs and parents' self-reports (a single item with an option of excellent, very good, good, fair or poor). The use of a single item to measure health status has been shown to have good psychometric properties [47,48]. We conducted measurement invariance tests using Mplus [49] and other analyses using STATA 9.0 [50].

3. Results

3.1. Characteristics of subjects

Of the 364 matched pairs of parents and adolescents analyzed in this study, the mean age of parents and adolescents were 47 (SD = 9.9) and 17 (SD = 1.1) years, respectively. Fifty-one percent of adolescents were boys. For adolescents' race/ethnicity, 37% were white, 40% were black, 18% were Hispanics, and 6% were other. Thirty-two percent of parents had an education level below high school, 37% had a high school education, and 31% had some college, associate degree, or above. For parent-reported adolescent's health status, 13% were excellent, 18% were very good, 33% were good, 28% were fair, and 8% were poor. For CRGs, 23% of adolescents had nonacute health care needs, 5% had significant acute conditions, 10% had minor chronic conditions, 34% had moderate chronic conditions, and 28% had major chronic conditions.

3.2. Measurement invariance tests

Table 2 shows the findings of measurement invariance tests. The first model (M1), which places equality constraints of item clusters on HRQOL domains across two groups, reveals acceptable model fit in all domains. This suggests that the configural invariance of HRQOL measurements was held between the dyads.

The second model (M2), which adds equality constraints of factor loadings of items on HRQOL domains to model 1, reveals acceptable model fit in all domains. This suggests that the metric invariance was held and the dyadic ratings of HRQOL were based on the same unit of measurement.

The third model (M3) shows, after additionally imposing equality constraints on the intercepts of items between two groups, the scalar invariance was rejected by all HRQOL domains. The P-values of SB χ^2 were < 0.05 for all domains; RMSEAs were >0.06 for physical and social functioning; CFI was < 0.95 for physical functioning; and CFIs were >0.01 for all domains. After freely estimating the intercepts of items with noninvariance (5, 2, 3, and 2 items in physical, emotional, social, and school functioning, respectively), the fourth model (M4) shows acceptable model fit. This suggests that there was a partial scalar invariance for HRQOL measurements between the dyads.

The fifth model (M5) shows, by additionally imposing equality constraints on the residuals of item ratings between the two groups, the residual invariance was rejected by three domains: physical, emotional, and social functioning. The P-values of SB χ^2 were < 0.05 for all three domains; RMSEAs were >0.06, and CFIs were < 0.95 for physical and social functioning; CFIs were >0.01 for all three domains. After freely estimating the residuals of noninvariant items (4, 1, and 2 items for physical, emotional, and social functioning, respectively), the sixth model (M6) shows acceptable model fit. This suggests that there was

partial residual invariance for HRQOL measurements between the dyads. Those items identified with measurement noninvariance are shown in the Appendix.

3.3. Discrepancies in the dyadic HRQOL ratings before correlate adjustment

Table 3 shows the discrepancies in parents' and adolescents' HRQOL ratings. Before assuring measurement invariance, adolescents rated HRQOL higher than their parents in all domains (all with $P < 0.001$). The effect sizes in discrepancies across all domains were larger than 0.3, a cutoff indicating a minimally important difference [51,52]. The magnitudes in the discrepancies were small for school and emotional functioning (0.31 and 0.33, respectively), but moderate for physical and school functioning (0.54 and 0.72, respectively).

After adjusting for items with measurement noninvariance (see Section 2.4), adolescents also demonstrated higher HRQOL ratings compared to parents in all domains (all with $P < 0.001$). The effect sizes in discrepancies across all domains were larger than 0.3, indicating a minimally important difference. However, the magnitudes of the discrepancies were expanded, especially for emotional and school functioning.

3.4. Discrepancies in the dyadic HRQOL ratings after correlate adjustment

Table 4 shows the discrepancy in HRQOL after adjusting for items with measurement noninvariance and controlling for covariates (adolescent's age, gender, race, health status, and parent's level of education). We found that adolescents still demonstrated higher HRQOL ratings compared to parents in all domains (all with $P < 0.001$). The magnitudes of effect size were small for physical functioning (0.38), but moderate for emotional, school, and social functioning (0.50, 0.66, and 0.73, respectively).

Adolescents' gender and parent-reported health status were significantly correlated with HRQOL ratings in all domains. Compared to girls, boys had significantly higher HRQOL scores in all domains ($P < 0.01$), except school functioning. Adolescents with more impaired parent-reported health status had significantly lower HRQOL scores in all domains than healthy adolescents. This association was obvious for physical and emotional functioning ($P < 0.001$). Health status as classified by the CRGs system was less significantly associated with HRQOL ratings than parent-reported health status.

3.5. Correlates of discrepancy in the dyadic HRQOL ratings

Parents' report of their child's health status was the only variable significantly associated with the discrepancies in the dyadic ratings of HRQOL. Figure 2 shows that when the levels of parent-reported adolescent health status were more impaired, the discrepancies in the dyadic HRQOL ratings increased. This pattern was obvious for physical, social, and emotional functioning. For example, the effect sizes in physical functioning were large (above 0.80) for poor health status, and moderate (0.5–0.8) for fair health status.

4. Discussion

In this study, we demonstrated pediatric HRQOL measured by the PedsQL was not invariant between parents and adolescents. Although all items across all domains show equivalent HRQOL structure (configural invariance) and equivalent magnitude in the association between the items and underlying HRQOL (metric invariance), only some items show equivalent starting value in item ratings and equivalent nonsystematic measurement errors in item responses. We also demonstrated that after adjusting for items with measurement noninvariance, HRQOL across all domains were rated more poorly by parents than their

adolescents. Parents' rating of adolescent health, especially those with more impaired health, was the most important predictor for the discrepancies in the dyadic ratings.

Our findings are consistent with earlier studies showing a higher level of discrepancy in the dyadic ratings when items ask about abstract concepts (e.g., emotional well being) or when the dyads have access to different information (e.g., social and school activities) [4,5,7]. The possible interpretations for the discrepancy in the dyadic ratings can be due to the fact that parents and adolescents may interpret the same item in different ways [53]. One qualitative study reported that for items relevant to psychosocial functioning (e.g., asking "happy at school" or "friends help each other"), children related these items to one simple example/ setting in their daily life, whereas parents often referred to several examples/settings [37]. Second, the discrepancy may reflect different expectations by parents and children about the child's health status. It means, parents tend to incorporate concerns and worries about their child's health conditions in the HRQOL ratings, and perceive an illness to have more negative consequences than their child experiences [54,55]. This phenomenon is especially significant for children with life-limiting conditions; for example, parents show greater concerns for the future impact of pediatric cancer, whereas children show greater concerns about the immediate consequence, such as lose of friends or hair [54].

We suggest that parents' and adolescents' ratings of adolescents' HRQOL may provide different information, which should be considered as complementary rather than substitutive. The use of dyadic reports are important because we found the discrepancies in the dyadic ratings were statistically significant (in all domains), and the effect sizes in the discrepancies were meaningfully important difference with the range between small (physical functioning) and moderate (other domains). The combined use of the dyadic ratings is particularly important for children with severe health conditions because we found that the discrepancy in the dyadic ratings was largest for children with more severe health conditions. Parents' perceptions of children's HRQOL determine the utilization of health services, and both parents' and adolescents' ratings influence the decision-making process with providers for treatments [34,56,57]. Psychometric evidence also shows that parents' and children's reports have different merits. Parents' ratings are expected to have greater reliability, whereas children's ratings are likely to have greater validity [58,59].

Although we recommend the dual use of parent's and adolescent's HRQOL ratings, the time required for both parents and adolescents to complete the forms may limit its application in busy clinical settings. In addition, clinicians may be unclear about how to manage potentially disparate scores derived from the dyadic ratings. The use of item response theory to develop item banks containing measurement invariant items, further combined with computerized adaptive test technology, has the potential to provide a solution [60,61]. Instead of completing the entire set of items, this methodology guides parents and adolescents in selecting a subset of items from item banks to better represent the adolescent's HRQOL. As a result, the dyadic HRQOL ratings can be efficient and precise, both vitally important in busy clinical practice. Because the items used by parents and adolescents are selected from the same metric (a calibrated item bank), clinicians can interpret and manage the estimated scores easily (e.g., taking an average score) when planning care.

Our study has limitations that merit attention. First, this study is restricted to children who were between 15 through 18 years old with health conditions of sufficient severity to warrant enrollment in the State Title V Children with Special Health Care Needs Program. This limits the generalizability of our findings to other pediatric populations. Second, we only tested measurement invariance in HRQOL using the PedsQL. Its items and domains, however, may be different from other pediatric HRQOL instruments [62]. Third, we did not

explicitly control for parents' mental health status, which may confound the dyadic comparisons [63–65]. Finally, our methods for measurement invariance tests can be improved. In this study, we treated item response as a continuous variable. Further studies should specify item response as a categorical variable in the MG-CFA tests.

5. Conclusions

Measurement invariance is a necessary condition for meaningful comparisons of pediatric HRQOL rated by parents and children. Using the PedsQL, we demonstrated that some items show measurement noninvariance. After adjusting for items with noninvariant features, adolescents significantly reported higher HRQOL scores than their parents in all domains. Given the fact that there were significant discrepancies in the dyadic ratings, and both parents and children provide the most comprehensive picture regarding the impacts of disease on children's well being, we recommend the use of dyadic HRQOL ratings to better measure children's HRQOL.

Appendix

Items of the PedsQL show measurement noninvariance

	Scalar noninvariance	Residual noninvariance
Physical functioning		
1. Walking more than one block		
2. Running	X	X
3. Participating in sports activity or exercise		
4. Lifting something heavy		X
5. Taking a bath or shower by him or herself	X	X
6. Doing chores around the house	X	X
7. Having hurts or aches	X	
8. Low energy level	X	
Emotional functioning		
9. Feeling afraid or scared		X
10. Feeling sad or blue		
11. Feeling angry		
12. Trouble sleeping	X	
13. Worrying about what will happen to him or her	X	
Social functioning		
14. Getting along with other teens		X
15. Other teens not wanting to be his or her friend	X	
16. Getting teased by other teens		
17. Not able to do things that other teens his or her age can do	X	
18. Keeping up with other teens	X	X
School functioning		
19. Paying attention in class		
20. Forgetting things	X	
21. Keeping up with schoolwork		
22. Missing school because of not feeling well		

	Scalar noninvariance	Residual noninvariance
23. Missing school to go to the doctor or hospital	X	

All items show configural and metric invariance.

References

1. Spieth LE, Harris CV. Assessment of health-related quality of life in children and adolescents: an integrative review. J Pediatr Psychol. 1996; 21:175–193. [PubMed: 8920152]

2. Eiser C, Morse R. A review of measures of quality of life for children with chronic illness. Arch Dis Child. 2001; 84:205–211. [PubMed: 11207164]

3. Addington-Hall J, Kalra L. Who should measure quality of life? BMJ. 2001; 322:1417–1420. [PubMed: 11397754]

4. Eiser C, Morse R. Can parents rate their child's health-related quality of life? Results of a systematic review. Qual Life Res. 2001; 10:347–357. [PubMed: 11763247]

5. Achenbach TM, McConaughy SH, Howell CT. Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. Psychol Bull. 1987; 101:213–232. [PubMed: 3562706]

6. Theunissen NC, Vogels TG, Koopman HM, Verrips GH, Zwinderman KA, et al. The proxy problem: child report versus parent report in health-related quality of life research. Qual Life Res. 1998; 7:387–397. [PubMed: 9691719]

7. De Civita M, Regier D, Alamgir AH, Anis AH, Fitzgerald MJ, Marra CA. Evaluating health-related quality-of-life studies in paediatric populations: some conceptual, methodological and developmental considerations and recent applications. Pharmacoeconomics. 2005; 23:659–685. [PubMed: 15987225]

8. Sturms LM, van der Sluis CK, Groothoff JW, ten Duis HJ, Eisma WH. Young traffic victims' long-term health-related quality of life: child self-reports and parental reports. Arch Phys Med Rehabil. 2003; 84:431–436. [PubMed: 12638113]

9. Ennett ST, DeVellis BM, Earp JA, Kredich D, Warren RW, Wilhelm CL. Disease experience and psychosocial adjustment in children with juvenile rheumatoid arthritis: children's versus mothers' reports. J Pediatr Psychol. 1991; 16:557–568. [PubMed: 1744805]

10. Russell KM, Hudson M, Long A, Phipps S. Assessment of health-related quality of life in children with cancer: consistency and agreement between parent and child reports. Cancer. 2006; 106:2267–2274. [PubMed: 16604563]

11. Britto MT, Kotagal UR, Chenier T, Tsevat J, Atherton HD, Wilmott RW. Differences between adolescents' and parents' reports of health-related quality of life in cystic fibrosis. Pediatr Pulmonol. 2004; 37:165–171. [PubMed: 14730662]

12. Chang PC, Yeh CH. Agreement between child self-report and parent proxy-report to evaluate quality of life in children with cancer. Psychooncology. 2005; 14:125–134. [PubMed: 15386781]

13. Klassen AF, Miller A, Fine S. Agreement between parent and child report of quality of life in children with attention-deficit/hyperactivity disorder. Child Care Health Dev. 2006; 32:397–406. [PubMed: 16784495]

14. Robitail S, Simeoni MC, Ravens-Sieberer U, Bruil J. Auquier Pfor the KIDSCREEN Group. Children proxies' quality-of-life agreement depended on the country using the European KIDSCREEN-52 questionnaire. J Clin Epidemiol. 2007; 60:469–478. [PubMed: 17419958]

15. Meredith W. Measurement invariance, factor analysis and factorial invariance. Psychometrika. 1993; 58:525–543.

16. Meredith W, Teresi JA. An essay on measurement and factorial invariance. Med Care. 2006; 44:S69–S77. [PubMed: 17060838]

17. Gregorich SE. Do self-report instruments allow meaningful comparisons across diverse population groups? Testing measurement invariance using the confirmatory factor analysis framework. Med Care. 2006; 44:S78–S94. [PubMed: 17060839]

18. Devins GM, Beiser M, Dion R, Pelletier LG, Edwards RG. Cross-cultural measurements of psychological well-being: the psychometric equivalence of Cantonese, Vietnamese, and Laotian translations of the affect balance scale. Am J Public Health. 1997; 87:794–799. [PubMed: 9184508]

19. Limbers CA, Newman DA, Varni JW. Factorial invariance of child self-report across age subgroups: a confirmatory factor analysis of ages 5 to 16 years utilizing the PedsQL 4.0 generic core scales. Value Health. in press.

20. Limbers CA, Newman DA, Varni JW. Factorial invariance of child self-report across healthy and chronic health condition groups: a confirmatory factor analysis utilizing the PedsQLTM 4.0 generic core scales. J Pediatr Psychol. 2008; 33:630–639. [PubMed: 18165221]

21. Nguyen HT, Kitner-Triolo M, Evans MK, Zonderman AB. Factorial invariance of the CES-D in low socioeconomic status African Americans compared with a nationally representative sample. Psychiatry Res. 2004; 126:177–187. [PubMed: 15123397]

22. Yao G, Wu CH. Factorial invariance of the WHOQOL-BREF among disease groups. Qual Life Res. 2005; 14:1881–1888. [PubMed: 16155775]

23. Lubke GH, Dolan CV. Can unequal residual variances across groups mask differences in residual means in the common factor model? Struct Equation Model. 2003; 10:175–192.

24. Millsap RE, Kwok OM. Evaluating the impact of partial factorial invariance on selection in two populations. Psychol Methods. 2004; 9:93–115. [PubMed: 15053721]

25. Deshon RP. Measures are not invariant across groups without error variance homogeneity. Psychol Sci. 2004; 46:137–149.

26. Lotstein DS, McPherson M, Strickland B, Newacheck PW. Transition planning for youth with special health care needs: results from the national survey of children with special health care needs. Pediatrics. 2005; 115:1562–1568. [PubMed: 15930217]

27. Anarella J, Roohan P, Balistreri E, Gesten F. A survey of medicaid recipients with asthma: perceptions of self-management, access, and care. Chest. 2004; 125:1359–1367. [PubMed: 15078746]

28. Dick AW, Brach C, Allison RA, Shenkman E, Shone LP, Szilagyi PG, et al. SCHIP's impact in three states: how do the most vulnerable children fare? Health Aff. 2004; 23:63–75.

29. Neff JM, Sharp VL, Muldoon J, Graham J, Popalisky J, Gay JC. Identifying and classifying children with chronic conditions using administrative data with the clinical risk group classification system. Ambul Pediatr. 2002; 2:71–79. [PubMed: 11888441]

30. Hughes JS, Averill RF, Eisenhandler J, Goldfield NI, Muldoon J, Neff JM, et al. Clinical risk groups (CRGs): a classification system for risk-adjusted capitation-based payment and health care management. Med Care. 2004; 42:81–90. [PubMed: 14713742]

31. Shenkman E, Knapp C, Sappington D, Vogel B, Schatz D. Persistence of high health care expenditures among children in medicaid. Med Care Res Rev. 2007; 64:304–330. [PubMed: 17507460]

32. Varni JW, Seid M, Rode CA. The PedsQL: measurement model for the pediatric quality of life inventory. Med Care. 1999; 37:126–139. [PubMed: 10024117]

33. Varni JW, Seid M, Kurtin PS. PedsQL 4.0: reliability and validity of the pediatric quality of life inventory version 4.0 generic core scales in healthy and patient populations. Med Care. 2001; 39:800–812. [PubMed: 11468499]

34. Varni JW, Seid M, Knight TS, Uzark K, Szer IS. The PedsQL 4.0 generic core scales: sensitivity, responsiveness, and impact on clinical decision-making. J Behav Med. 2002; 25:175–193. [PubMed: 11977437]

35. Cheung GW, Rensvold RB. Testing factorial invariance across groups: a reconceptualization and proposed new method. J Manage. 1999; 25:1–27.

36. Ungar WJ, Mirabelli C, Cousins M, Boydell KM. A qualitative analysis of a dyad approach to health-related quality of life measurement in children with asthma. Soc Sci Med. 2006; 63:2354–2366. [PubMed: 16887248]

37. Davis E, Nicolas C, Waters E, Cook K, Gibbs L, Gosch A, et al. Parent-proxy and child self-reported health-related quality of life: using qualitative methods to explain the discordance. Qual Life Res. 2007; 16:863–871. [PubMed: 17351822]

38. Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. Struct Equation Model. 2002; 9:233–255.

39. Byrne BM, Shavelson RJ, Muthen B. Testing for the equivalence of factor covariance and mean structures—the issue of partial measurement invariance. Psychol Bull. 1989; 105:456–466.

40. Reise SP, Widaman KF, Pugh RH. Confirmatory factor-analysis and item response theory—2 approaches for exploring measurement invariance. Psychol Bull. 1993; 114:552–566. [PubMed: 8272470]

41. Steenkamp JBEM, Baumgartner H. Assessing measurement invariance in cross-national consumer research. J Consum Res. 1998; 25:78–90.

42. Cohen, J. Statistical power analysis for the behavioral sciences. 2nd edition. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.

43. Fayers, PM.; Machin, D. Quality of life: the assessment, analysis, and interpretation of patient-reported outcomes. 2nd edition. Chichester/-Hoboken, NJ: John Wiley; 2007.

44. Zeger SL, Liang KY, Albert PS. Models for longitudinal data—a generalized estimating equation approach. Biometrics. 1988; 44:1049–1060. [PubMed: 3233245]

45. Varni JW, Burwinkle TM, Seid M, Skarr D. The PedsQL 4.0 as a pediatric population health measure: feasibility, reliability, and validity. Ambul Pediatr. 2003; 3:329–341. [PubMed: 14616041]

46. White-Koning M, Arnaud C, Dickinson HO, Thyen U, Beckung E, Fauconnier J, et al. Determinants of child-parent agreement in quality-of-life reports: a European study of children with cerebral palsy. Pediatrics. 2007; 120:e804–e814. [PubMed: 17908738]

47. de Boer AG, van Lanschot JJ, Stalmeier PF, van Sandick JW, Hulscher JB, de Haes JC, et al. Is a single-item visual analogue scale as valid, reliable and responsive as multi-item scales in measuring quality of life? Qual Life Res. 2004; 13:311–320. [PubMed: 15085903]

48. DeSalvo KB, Fisher WP, Tran K, Bloser N, Merrill W, Peabody J. Assessing measurement properties of two single-item general health measures. Qual Life Res. 2006; 15:191–201. [PubMed: 16468076]

49. Muthen, LK.; Muthen, BO. Mplus: User' guide. 4th edition. Los Angeles, CA: Muthen & Muthen; 2004.

50. STATCorp. Stata statistical software: Release 9.0. College Station, TX: Stata Corporation; 2005.

51. Guyatt GH, Osoba D, Wu AW, Wyrwich KW. Norman GRClinical Significance Consensus Meeting Group. Methods to explain the clinical significance of health status measures. Mayo Clin Proc. 2002; 77:371–383. [PubMed: 11936935]

52. Revicki DA, Cella D, Hays RD, Sloan JA, Lenderking WR, Aaronson NK. Responsiveness and minimal important differences for patient reported outcomes. Health Qual Life Outcomes. 2006; 4:70. [PubMed: 17005038]

53. Jokovic A, Locker D, Guyatt G. How well do parents know their children? Implications for proxy reporting of child health-related quality of life. Qual Life Res. 2004; 13:1297–1307. [PubMed: 15473508]

54. Vance YH, Morse RC, Jenney ME, Eiser C. Issues in measuring quality of life in childhood cancer: measures, proxies, and parental mental health. J Child Psychol Psychiatry. 2001; 42:661–667. [PubMed: 11464970]

55. Sprangers MA. Quality-of-life assessment in oncology. Achievements and challenges. Acta Oncol. 2002; 41:229–237. [PubMed: 12195741]

56. Janicke DM, Finney JW, Riley AW. Children's health care use: a prospective investigation of factors related to care-seeking. Med Care. 2001; 39:990–1001. [PubMed: 11502956]

57. Campo JV, Comer DM, Jansen-Mcwilliams L, Gardner W, Kelleher KJ. Recurrent pain, emotional distress, and health service use in childhood. J Pediatr. 2002; 141:76–83. [PubMed: 12091855]

58. le Coq EM, Boeke AJ, Bezemer PD, Colland VT, van Eijk JT. Which source should we use to measure quality of life in children with asthma: the children themselves or their parents? Qual Life Res. 2000; 9:625–636. [PubMed: 11236853]

59. Matza LS, Swensen AR, Flood EM, Secnik K, Leidy NK. Assessment of health-related quality of life in children: a review of conceptual, methodological, and regulatory issues. Value Health. 2004; 7:79–92. [PubMed: 14720133]

60. Cella D, Gershon R, Lai JS, Choi S. The future of outcomes measurement: item banking, tailored short-forms, and computerized adaptive assessment. Qual Life Res. 2007; 16:133–141. [PubMed: 17401637]

61. Reeve BB, Burke LB, Chiang YP, Clauser SB, Colpe LJ, Elias JW, et al. Enhancing measurement in health outcomes research supported by agencies within the US department of health and human services. Qual Life Res. 2007; 16:175–186. [PubMed: 17530449]

62. Eiser C, Morse R. The measurement of quality of life in children: past and future perspectives. J Dev Behav Pediatr. 2001; 22:248–256. [PubMed: 11530898]

63. Berg-Nielsen TS, Vika A, Dahl AA. When adolescents disagree with their mothers: CBCL-YSR discrepancies related to maternal depression and adolescent self-esteem. Child Care Health Dev. 2003; 29:207–213. [PubMed: 12752611]

64. Tarullo LB, Richardson DT, Radkeyarrow M, Martinez PE. Multiple sources in child diagnosis—parent-child concordance in affectively ill and well families. J Clin Child Psychol. 1995; 24:173–183.

65. Waters E, Doyle J, Wolfe R, Wright M, Wake M, Salmon L. Influence of parental gender and self-reported health and illness on parent-reported child health. Pediatrics. 2000; 106:1422–1428. [PubMed: 11099598]

What is new?

1. Key findings:

   - Agreements between the ratings of adolescents' health-related quality of life (HRQOL) by parents and adolescents were poor.

   - Adolescents reported significantly higher HRQOL scores than parents in the domains of physical, emotional, social, and school functioning.

2. What this adds to the existing literature:

   - Standard statistical methods (e.g., *t*-test or linear regression) used to compare adolescents' HRQOL as reported by parents and adolescents are limited because they cannot assure whether measurements between the dyads are comparable (i.e., measurement noninvariance). This lack of comparability between adolescents and parents might be due to differences in perceived constructs of HRQOL or item meanings.

   - This study demonstrated the use of multi-group confirmatory factor analysis to detect inconsistent measurement properties in HRQOL ratings and provided solutions to adjust this measurement issue. These adjustments will allow for meaningful comparisons of adolescents' HRQOL by parents and adolescents.

3. What is the implication:

   - Given the poor agreement in adolescents' HRQOL as reported by parents and adolescents, this study suggests that dyadic ratings should be used rather than parents' or adolescents' ratings alone because both may provide different, yet valuable information.
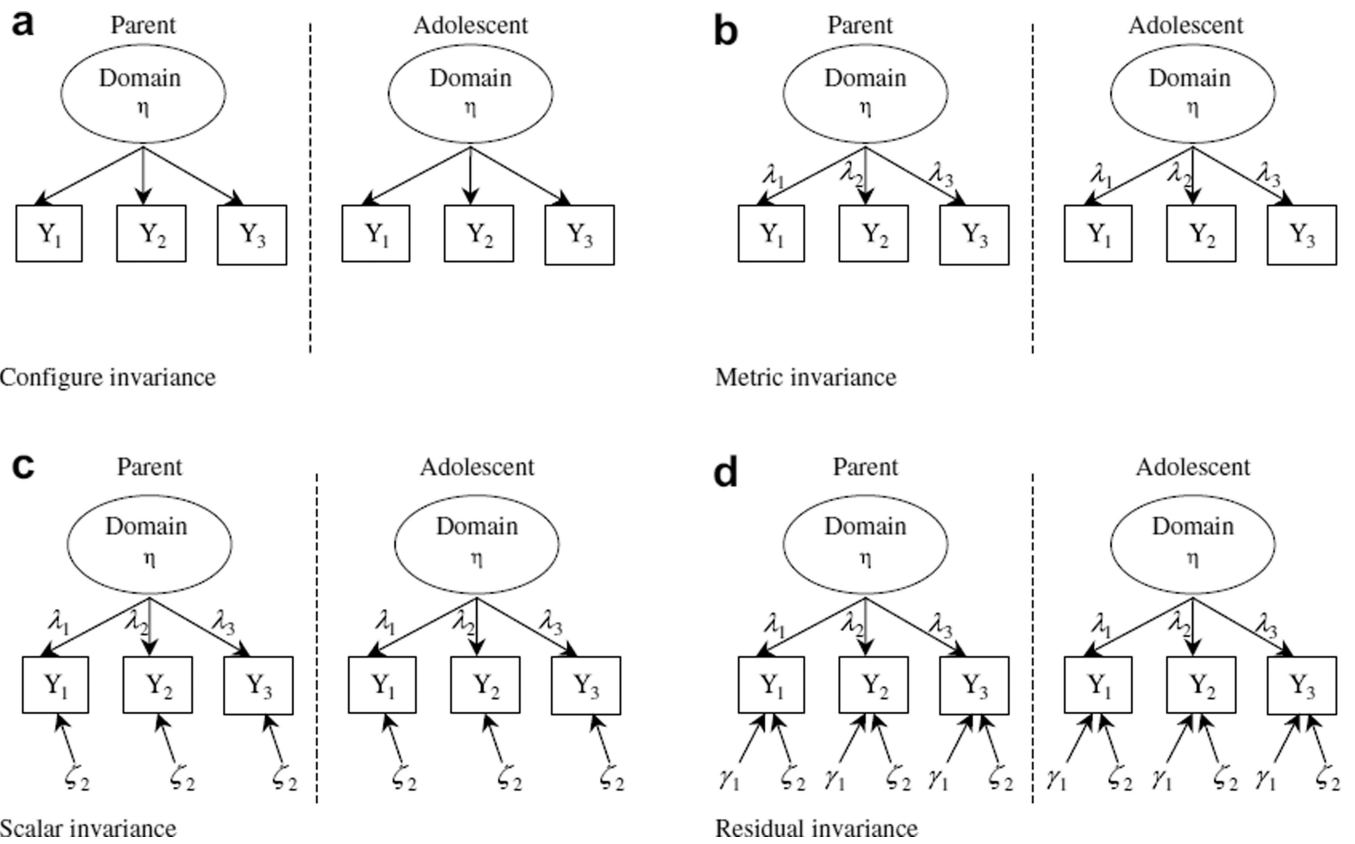
Fig. 1.
Tests of measurement invariance in the framework of multi-group confirmatory factor analysis. (a) Configure invariance, (b) Metric invariance, (c) Scalar invariance, and (d) Residual invariance. Where, Y, item score; , domain (or factor) score; , factor loading; , intercept of item score; , residuals.

Fig. 2.
Discrepancy in adolescents' HRQOL rated by the dyads vs. level of health status. Effect size: negligible (< 0.2), small (0.2~0.49), moderate (0.5~0.79), and large (0.8).

Table 1
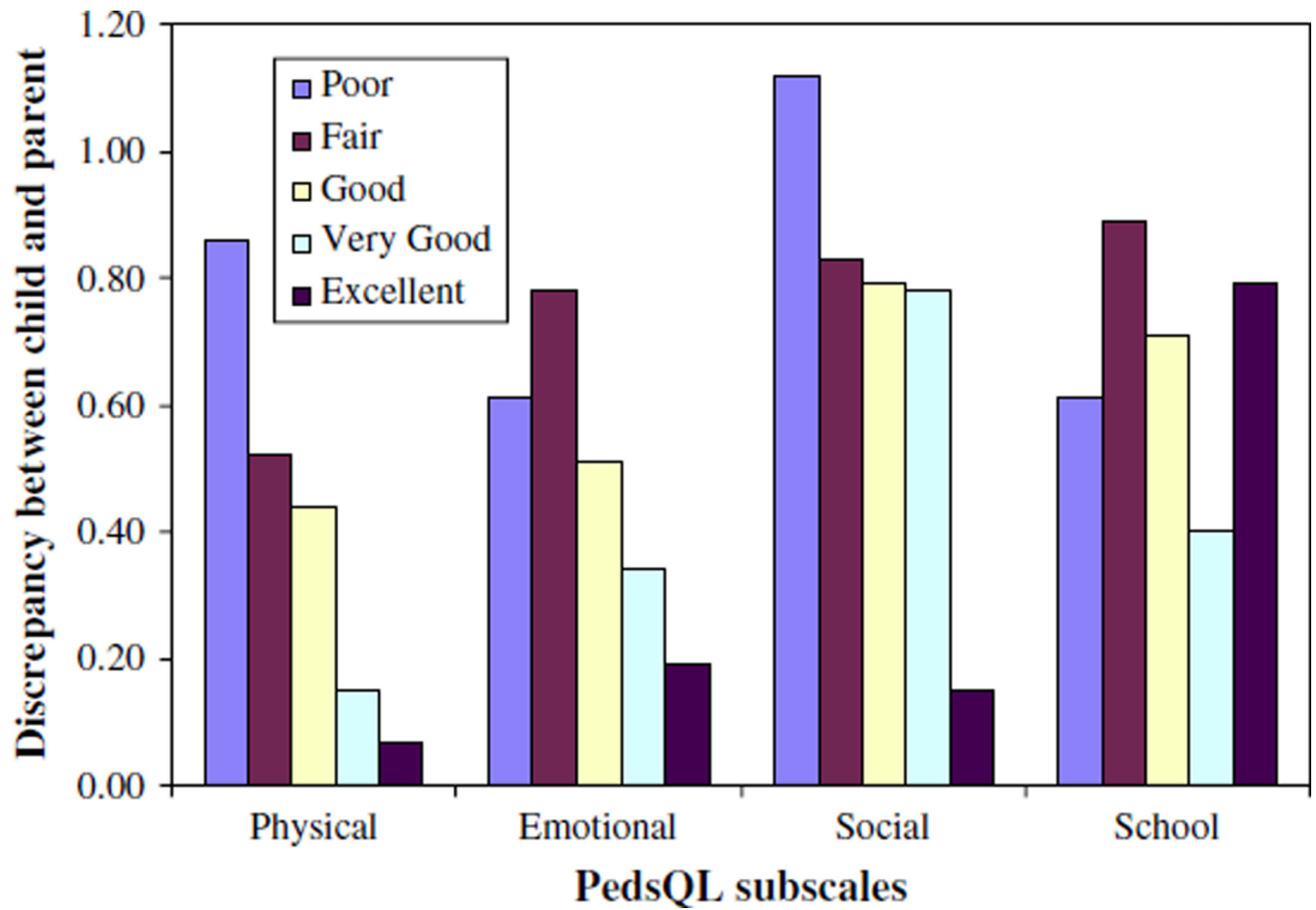
Taxonomy of measurement invariance

Measurement invariance	Interpretations	Statistical tests in multiple-group confirmatory factor analysis
Configural invariance	There is an equivalent HRQOL construct between two groups	Whether item clusters are identical between two groups?
Metric invariance	Given configure invariance, there are equivalent associations between items and underlying HRQOL between two groups (or the same unit of measurement)	Given configure invariance, whether factor loadings of specific items are identical between two groups?
Scalar invariance	Given metric invariance, there are equivalent starting values of item ratings between two groups	Given metric invariance, whether intercepts of specific items are identical between two groups?
Residual invariance	Given scalar invariance, there are equivalent nonsystematic measurement errors in the item responses between two groups	Given scalar invariance, whether residuals of specific items are identical between two groups?

Table 2

Tests of measurement invariance in four domains of the PedsQL 4.0

Models of MI of testing	Reference model	SB ^{2}	P-value of SB ^{2}	RMSEA	CFI	CFI	No (%) of items with MI
Physical							
M1: Configural	—	—	—	0.06	0.97	—	8 (100%)
M2: Metric	M1	3.48	0.75	0.05	0.97	0	8 (100%)
M3: Scalar	M2	99.24	<0.05	0.08	0.92	0.05	—
M4: Partial scalar	M2	0.20	0.90	0.05	0.97	0	3 (38%)
M5: Residual	M4	54.38	<0.05	0.08	0.93	0.04	—
M6: Partial residual	M4	3.13	0.37	0.05	0.97	0	4 (50%)
Emotional							
M1: Configural	—	—	—	0.05	0.99	—	5 (100%)
M2: Metric	M1	6.10	0.19	0.04	0.99	0	5 (100%)
M3: Scalar	M2	13.74	<0.05	0.05	0.98	0.01	—
M4: Partial scalar	M2	2.26	0.32	0.04	0.99	0	3 (60%)
M5: Residual	M4	12.4	<0.05	0.05	0.98	0.01	—
M6: Partial residual	M4	3.55	0.47	0.04	0.99	0	4 (80%)
Social							
M1: Configural	—	—	—	0.06	0.97	—	5 (100%)
M2: Metric	M1	4.59	0.33	0.06	0.97	0	5 (100%)
M3: Scalar	M2	30.76	<0.05	0.08	0.95	0.02	—
M4: Partial scalar	M2	0.61	0.44	0.06	0.97	0	2 (40%)
M5: Residual	M4	39.23	<0.05	0.09	0.92	0.05	—
M6: Partial residual	M4	5.28	0.15	0.06	0.97	0	3 (60%)
School							
M1: Configural	—	—	—	0.02	0.99	—	5 (100%)
M2: Metric	M1	1.31	0.96	0.01	1.00	−0.01	5 (100%)
M3: Scalar	M2	18.19	<0.05	0.04	0.98	0.02	—
M4: Partial scalar	M2	4.13	0.13	0.01	1.00	0	3 (60%)
M5: Residual	M4	0.32	0.99	0.01	1.00	0	5 (100%)

Abbreviation: MI, measurement invariance.

Poor model fits were flagged in italic.

Table 3

Mean (SD) and the discrepancy in adolescents' HRQOL rated by the dyads[a]

	Parent	Adolescent	Difference[b]	Effect size[c]
Before adjusting noninvarianc[d]				
Physical	66.5 (26.5)	77.9 (21.3)	11.5 (22.9)	0.54
Emotional	66.7 (22.9)	74.1 (22.4)	7.4 (25.5)	0.33
Social	66.1 (25.7)	81.0 (20.7)	14.7 (25.6)	0.72
School	59.9 (22.9)	65.9 (19.3)	5.7 (23.6)	0.31
After adjusting noninvariance[e]				
Physical	73.7 (27.9)	82.3 (21.9)	8.6 (23.2)	0.40
Emotional	77.2 (15.7)	85.3 (15.5)	8.1 (17.6)	0.53
Social	70.5 (20.0)	83.0 (16.2)	12.4 (20.8)	0.77
School	60.5 (27.1)	74.1 (19.5)	13.6 (26.8)	0.70

[a]Unadjustment for covariates.

[b]All discrepancy in HRQOL ratings between the dyads were statistically significance ($P < 0.001$).

[c]Magnitude of the effect size: negligible (<0.2), small (0.2~0.49), moderate (0.5~0.79), and large (0.8).

[d]HRQOL ratings unadjusted for measurement noninvariance.

[e]HRQOL ratings adjusted for measurement noninvariance.

Table 4

Discrepancy in adolescents' HRQOL rated by the dyads after adjusting for covariates

	Physical functioning	Emotional functioning	Social functioning	School functioning
Group[a]	8.12 *** (0.38)[b]	7.71 *** (0.50)[b]	11.84 *** (0.73)[b]	12.94 *** (0.66)[b]
Age of child	−0.86	−0.44	0.62	1.43
Gender of child[c]	6.48 **	4.43 **	4.02 **	−1.97
Race of child				
Black	7.38 **	2.40	4.45 **	2.84
Others	3.84	0.26	−1.47	2.15
Health status report by parents[d]				
Very good	−1.21	−3.07	−2.67	−4.22
Good	−9.77 **	−6.16 ***	−6.04 *	−7.89 *
Fair	−13.93 ***	−9.58 ***	−6.42 *	−9.69 *
Poor	−18.51 ***	−15.12 ***	−8.29 *	−10.98 *
Health status by CRGs[e]				
Acute	3.18	3.89	4.86	8.33
Mild chronic	2.59	−3.26	−5.53	−9.41 *
Moderate chronic	−3 13	−1.56	−2.86	−2.59
Severe chronic	−11.76 ***	−2.05	−1.37	−2.55

*
$P < 0.05$;

**
$P < 0.01$;

***
$P < 0.001$.

[a]Parent is the reference group.

[b]Value in parenthesis is the effect size: negligible (<0.2), small (0.2~0.49), moderate (0.5~0.79), and large (0.8).

[c]Girl is the reference group.

[d]Excellent health is the reference group.

[e]Healthy is the reference group.