

Appropriate statistical testing of quality of life scores from children with asthma and their caregivers

John A Myers¹
Robert William Prasaad
Steiner¹
James Legleiter²
Yu-Ting Chen¹
Robert J Esterhay¹

¹University of Louisville School of Public Health and Information Sciences, ²Asthma Disease Management, Passport Health Plan, Louisville, KY, USA

Objective: Scores on the Mini-Pediatric Asthma Quality of Life Questionnaire (mPAQLQ) and Pediatric Asthma Caregiver's Quality of Life Questionnaire (PACQLQ) assess disease-specific quality of life (QoL). Paired *t*-tests are a commonly used analytic method to examine differences between these QoL scores. This report examines whether nonparametric approaches may be more appropriate.

Study design and setting: Responses on the mPAQLQ and PACQLQ were collected in a cross-sectional design with 68 child-caregiver dyads attending a community asthma prevention program – Healthy Hoops[®] Kentucky. Scores and differences in scores were compared using both parametric and nonparametric approaches.

Results: Scores on mPAQLQ subscales and overall QoL were not normally distributed, indicating that parametric tests may not be appropriate. Results of statistically significant differences in overall mean QoL scores between caregiver and children differed with parametric methods (paired *t*-tests: $P \leq 0.013$) and nonparametric approaches (Wilcoxon signed-rank tests: $P \leq 0.15$). Achieved statistical power with nonparametric methods was 97%–99% for differences in subscale and overall QoL scores. Probability plots for the tests for normality showed skewness at extreme values.

Conclusion: The currently advocated method for using parametric methods to compare scores from the mPAQLQ and PACQLQ may be incorrect. Using parametric methods with non-normal distributions can lead to errors. This study supports the nonparametric approach (Wilcoxon signed-rank test) for testing for differences between the mPAQLQ and PACQLQ. Instructions for scoring the mPAQLQ and PACQLQ may need to be revised and updated.

Keywords: pediatric asthma, quality of life, Mini-Pediatric Asthma Quality of Life Questionnaire, Caregiver's Pediatric Asthma Quality of Life Questionnaire, nonparametric, Wilcoxon signed-rank test

Key findings

The use of parametric statistical methods to compare differences in quality of life (QoL) scores from the mPAQLQ and PACQLQ may lead to incorrect results, since using parametric methods on data with non-normal distributions can lead to errors.

The results from paired *t*-tests were statistically significantly different when comparing scores from the non-normal distributions from the mPAQLQ and PACQLQ. However, the indicated Wilcoxon signed-rank test on the same data indicated no statistically significant differences between overall QoL scores, despite adequate statistical power.

The parametric approach to QoL assessment in this study appears to be too liberal, yielding statistically significant results when none may actually be present.

Correspondence: Rob Steiner
Health Management and Systems
Sciences, University of Louisville School
of Public Health and Information Sciences,
485 East Gray Street, Suite 115, Louisville,
KY 40202, USA
Tel +1 502 852 3006
Fax +1 502 852 3294
Email r.steiner@louisville.edu

This methodologic study supports the nonparametric approach (Wilcoxon signed-rank test) to test for differences between scores on the mPAQLQ and PACQLQ. Further research is needed to see if similar results are present in other datasets. Instructions for scoring the mPAQLQ and PACQLQ may need to be revised.

Introduction

Multi-domain quality of life (QoL) assessment questionnaires for children are known to be influenced by disease status, including asthma.¹ Chronic pediatric asthma can also have significant impact on family functioning.² Several valid and reliable disease-specific measures for asthma and QoL are available, some measuring QoL of the child, and others measuring QoL of the parents.^{1,3,4} The Mini-Pediatric Asthma Quality of Life Questionnaire⁵ (mPAQLQ) and the Pediatric Asthma Caregiver's Quality of Life Questionnaire⁶ (PACQLQ) provide comparative information from the child and their caregiver about the impact of the child's asthma. Reducing severity of asthma symptoms can improve QoL among children with asthma and their families.⁷

Proper techniques must be used to analyze data collected from these questionnaires. Analysts traditionally use parametric tests, such as *t*-tests, to evaluate differences between Likert scale scores from children and their caregivers. Paired *t*-tests are advocated since children with caregivers are matched pairs, rather than independent groups. However, there are two main assumptions required for proper use of the paired *t*-test: the outcome variable should be continuous; and the outcome variable should be normally distributed. When the assumptions of the *t*-test are violated, nonparametric methods are indicated.

The current study is designed to evaluate whether the results from the indicated Wilcoxon signed-rank test are comparable to results from parametric methods, when comparing concurrent scores on the mPAQLQ (children scores) and the PACQLQ (caregiver scores).

Quality of life assessments

Two disease-specific instruments have been advocated for comparing QoL of children with asthma and their parents/guardians. The Mini-Pediatric Asthma Quality of Life Questionnaire (mPAQLQ) is a 13-item instrument derived from the longer 23-item Pediatric Asthma Quality of Life Questionnaire (PAQLQ).⁸ The shorter and simpler mPAQLQ was developed to meet the need for greater efficiency in large clinical trials, monitoring groups of patients, and large surveys.¹ The mPAQLQ is psychometrically valid

and reliable.¹ Three subscales are evaluated in the mPAQLQ: "activity limitations," "emotional function", and "symptoms". The response options to all items are on a 7-point Likert scale, where 1 indicates maximum impairment and 7 indicates no impairment. All items are equally weighted. The results are expressed as the mean score per item for each of the three subscales and for overall QoL, as per the instructions for scoring.⁹

The Caregiver's Pediatric Asthma Quality of Life Questionnaire (PACQLQ) has 13 items designed to examine the impact of living with an asthmatic child on the caregiver's QoL.⁶ The questionnaire measures the problems that are most troublesome to the parents/caregivers of children with asthma over the course of 1 week. The PACQLQ includes the same subscales, except for the omission of "Symptoms". Scoring is the same as with mPAQLQ.⁹ The PACQLQ is psychometrically valid and reliable.⁶ Permission to use both questionnaires in the Healthy Hoops[®] Kentucky project was granted directly from the offices of Professor Elizabeth Juniper, the primary author of the two QoL questionnaires.

Methods

Students in Louisville, KY ranging from 9–13 years of age, each with a verified diagnosis of asthma, were recruited through mass media to attend a community program sponsored by a regional Medicaid managed care organization in collaboration with the University of Louisville, doing business as Passport Health Plan, during Fall, 2008. Healthy Hoops[®] Kentucky provides health screenings and education about asthma prevention and medical advice from health experts, while using basketball drills conducted by professional coaches as an incentive for participation. Trained volunteers collected data, including vital signs, pulmonary function testing, medication lists, and QoL assessments. Health experts provided medical advice about proper medication use, nutrition, and exercise for children with asthma and their caregiver. Each child left the program with an asthma action plan, including referral to a physician for select cases.

The QoL questionnaires were administered by trained data collectors to 72 children and their respective caregivers, typically mothers or grandmothers, each at separate stations to minimize interactions between child and caregiver during the collection of responses. Cue cards with a list of response options were provided to each participant. Four caregivers did not fully complete all items on the questionnaire, providing an analytic cohort of 68 child-caregiver dyads. IBM SPSS software (v 19.0; SPSS Inc, Chicago, IL) was used for statistical

data analyses. All data were de-identified. The Institutional Review Board at the University of Louisville provided approval for analyses of data for this project (IRB Tracking No 11.0108, approved April 20, 2011).

Analytical techniques

Juniper and associates advocate using paired *t*-tests to examine differences in overall QoL and subscale scores between children and their caregivers on their respective questionnaires. The Shapiro–Wilk test was used to determine if the QoL outcome variables were normally distributed.¹⁰ Subsequently, the differences between scores from the child and their caregiver were analyzed using both the advocated parametric approach as well as the nonparametric approach (Wilcoxon signed-rank test).¹⁰ The Wilcoxon signed-rank test is indicated for matched pairs, when data is non-normally distributed and the level of the data is ordinal or higher. Finally, we evaluated whether the different statistical approaches produced similar or different results with the same data in simulations.

Results

Most of the child participants were male (62%), African American (68%), with a mean age of 10.5 years old, whose home zip codes were in metropolitan areas with high poverty rates.

Measures of internal consistency measured by Cronbach's alpha, were excellent for overall QoL assessments on both mPAQLQ ($\alpha = 0.88$) and PACQLQ ($\alpha = 0.89$). Reliability coefficients for the subscales ranged from $\alpha = 0.66$ for emotional subscales for the child, to $\alpha = 0.82$ for emotional subscales from the caregivers (see Table 1).

Differences between various QoL scores from parents and their children were examined. Statistically significant differences were present in overall QoL scores between mPAQLQ and PACQLQ when using the paired *t*-tests approach (5.24 versus 4.79; $\Delta = 0.46$; $P_{\text{Paired } t\text{-tests}} = 0.013$),

Table 1 Comparing Cronbach's alpha coefficients for mPAQLQ and PACQLQ

	Reliability coefficients (Cronbach's α)	
	mPAQLQ	PACQLQ
Questionnaire		
Overall quality of life	0.879	0.889
Sub-scale		
Activity limitations	0.729	0.768
Symptoms	0.756	N/A
Emotional function	0.662	0.826

Abbreviations: mPAQLQ, Mini-Pediatric Asthma Quality of Life Questionnaire; PACQLQ, Pediatric Asthma Caregiver's Quality of Life Questionnaire.

while differences between children's and caregivers' scores for "activities limitation" were seen using both statistical approaches ($P_{\text{Paired } t\text{-tests}} \leq 0.0001$ and $P_{\text{Wilcoxon}} \leq 0.0001$, see Table 2). However, the Shapiro–Wilk tests indicate the scores for subscales and overall QoL for the mPAQLQ are not normally distributed, whereas scores for the PACQLQ are normally distributed. Consequently, the Wilcoxon signed-rank test is indicated as a preferred test in this situation.

Results for the differences between mPAQLQ and PACQLQ scores for overall QoL by Wilcoxon signed-rank test were not statistically significant ($P_{\text{Wilcoxon}} = 0.150$), although statistically significant differences remained for "Activities Limitation" ($P_{\text{Wilcoxon}} \leq 0.0001$). Differences in subscale scores "Emotional Function" did not reach statistical significance by either method ($P_{\text{Paired } t\text{-tests}} = 0.093$; $P_{\text{Wilcoxon}} = 0.142$).

The results using the Wilcoxon signed-rank method and paired *t*-tests provided different levels of statistical significance for the differences in overall QoL scores between the mPAQLQ and PACQLQ (see Table 2).

Despite lack of statistical significance for differences between scores, estimates of achieved statistical power for the Wilcoxon method were high: 97% for emotional functioning, 99% for activity limitations, and 99% for overall QoL.¹¹ This indicates that insufficient statistical power of the study does not threaten the statistically significant differences with nonparametric methods.

Null-hypothesis tests for normality are notoriously unreliable. Hence, we used probability plots to further illustrate the data we collected. Figure 1 shows the probability plot of mean scores for the emotional function subscale for children with asthma. This plot is included to represent our concerns about deviations from normal distributions among the various measures and subscales for QoL scores. From the plot, the data deviates from the normal distribution at both high and low extremes of the range. These findings further support our conclusions about rationale for not using *t*-tests, and preferring the nonparametric Wilcoxon signed-rank test.

A Monte Carlo study was conducted to further evaluate the effects of extreme mean scores for QoL while using *t*-tests for differences in means among the non-normal data distributions. The simulations were based on the data we obtained from our sample, which had an overall average mean QoL score for the $n = 72$ sample mean of 5.02. For the simulations, the mean quality values of QoL scores were categorized at values 3.5, 4.0, 4.5, 5.0, 5.5, and 6.0. One thousand random samples were generated. Simulation results were evaluated for bias, precision, and significance testing against the results of the original sample.

Table 2 Quality of life scores for subscale and overall quality of life, for mPAQLQ scores for children and PACQLQ scores for their caregivers

QoL measures	mPAQLQ scores child n = 72		Distribution for children's scores		PACQLQ scores caregiver n = 68		Distribution for caregivers' scores		Difference between QoL Scores n = 68		Effect size for differences between mPAQLQ and PACQLQ scores		Parametric statistics		Nonparametric statistics	
	Mean (SD)	Median	Shapiro-Wilks P value	Shapiro-Wilks P value	Mean (SD)	Median	Shapiro-Wilks P value	Shapiro-Wilks P value	Difference in means	Difference in medians			Paired t-test P value	Wilcoxon signed-rank test P-value		
Overall QoL	5.24 (1.23)	5.30	0.0026***	0.0026***	4.79 (1.19)	4.92	0.1244	0.1244	0.46	0.38	0.37	0.013***	0.150			
Subscale: "Activity limitation"	5.19 (1.29)	5.33	0.0006***	0.0006***	4.40 (1.49)	4.50	0.1181	0.1181	0.79	0.83	0.56	<0.001***	<0.001***			
Subscale: "Symptoms"	5.16 (1.35)	5.50	0.0002***	0.0002***	N/A	N/A	N/A	N/A				N/A	N/A			
Subscale: "Emotional function"	5.39 (1.61)	5.48	<0.0001***	<0.0001***	4.97 (1.19)	5.11	0.1611	0.1611	0.42	0.37	0.27	0.093	0.142			

Notes: All scores were based on a 1-7 ordinal scale, with "1" representing maximal limitations, and "7" representing no limitations. Differences between scores from a child and their caregiver were tested using paired t-tests and Wilcoxon signed-rank test. The significance level was set at $P \leq 0.05$; *** $P \leq 0.001$.

Abbreviations: QoL, quality of life; mPAQLQ, Mini-Pediatric Asthma Quality of Life Questionnaire; PACQLQ, Pediatric Asthma Caregiver's Quality of Life Questionnaire; SD, standard deviation.

The results from the simulations show that mean QoL of scores less than 4.5 show no major differences when compared to the original data. However, for mean QoL scores greater than 4.5, the simulated results showed several bias. For example, the simulations produced more projected scores with variance estimates smaller than those in the observed data, the 95% confidence limits about the projected values were not symmetric relative to the mean, and the number of paradoxical associations in the simulation (ie, significance in the wrong direction) increased. These findings indicate that the higher projected QoL scores on the questionnaires studied are associated with problems when examined with *t*-tests. In light of the findings from the simulation, and since the Likert scales are ordinal in nature but include responses with non-normal distributions, the indicated nonparametric Wilcoxon signed-rank tests are the better selection for statistical tests to analyze these QoL scores.

Discussion

The parametric and nonparametric methods of analyses applied to differences in mPAQLQ and PACQLQ scores for differences in overall QoL assessments between children and their caregivers provide different results in terms of statistical significance. The results for differences in overall QoL scores using paired *t*-tests are highly statistically significant, whereas the results on the same data using Wilcoxon signed-rank test are not statistically significant. Nonparametric methods are indicated when the scores for the QoL assessments are not normally distributed. Even though the nonparametric methods to evaluate differences between children's and caregivers' overall QoL scores were not statistically significant, post-hoc achieved power calculations showed there was sufficient statistical power present in the current study to detect the differences. Simulations provided evidence that multiple biases were present at the extremes of QoL scores. Even though *t*-tests are known to be robust against deviations from normality, the differences present in the face of adequate statistical power suggest that nonparametric analyses should be preferred for analyzing differences between scores from these specific QoL instruments.

The magnitudes of children's scores are always greater than scores from their respective caregivers for overall QoL and the two asthma subscales. These findings are similar to results from other studies,¹² but differ from experimental designs using similar questionnaires on rural children with asthma and their caregivers.¹³ Whether the differences between QoL scores for children and their caregivers in previous studies, as well as our study, represents differences

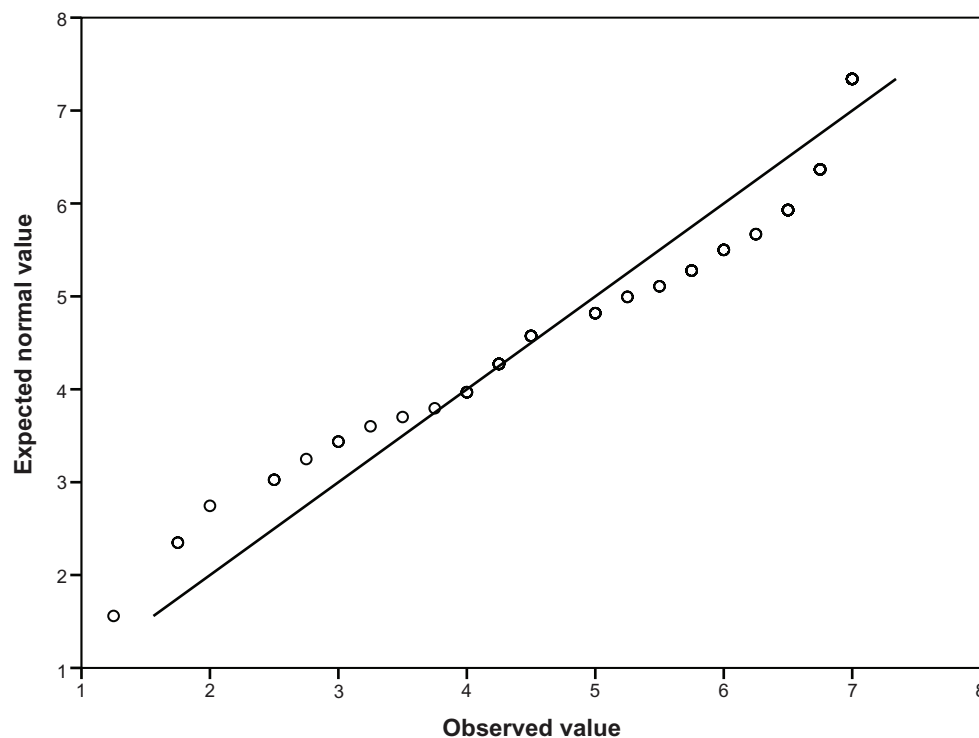


Figure 1 Probability plots to assess normality of clinical data.

in the characteristics of the study participants or are attributed to other factors is uncertain. Further studies are needed to evaluate the current results in comparison to results previously reported in the literature.

Finally, the magnitude of all P values from the paired t -tests were smaller than the P values obtained from the Wilcoxon signed-rank test on the same data, suggesting that the results from the parametric approach (paired t -tests) appear to be too liberal, yielding statistically significant results when none may actually be present. Thus, the advocated parametric approach for testing for differences between scores on the mPAQLQ and PACQLQ may not be appropriate, and may lead to spurious results.

Conclusion

The nonparametric approach using the Wilcoxon signed-rank test to analyze differences in QoL scores from dyads of asthmatic children with mPAQLQ scores and caregivers with PACQLQ scores is supported by the current results. Analyses of differences in QoL scores from mPAQLQ and PACQLQ may best be accomplished using nonparametric methods, such as the Wilcoxon signed-rank test, instead of the currently recommended paired t -tests. Further research is suggested to replicate the findings of this report using additional datasets with concurrent mPAQLQ and PACQLQ scores,

and comparing results from parametric and nonparametric methods of analyses. Instructions for scoring the mPAQLQ and PACQLQ may need to be revised and updated.⁹

Disclosure

The authors report no conflicts of interest in this work.

References

1. Juniper EF, Guyatt GH, Feeny DH, Ferrie PJ, Griffith LE, Townsend M. Measuring quality of life in children with asthma. *Qual Life Res.* 1996; 5(1):35–46.
2. Townsend M, Feeny DH, Guyatt GH, Furlong WJ, Seip AE, Dolovich J. Evaluation of the burden of illness for pediatric asthmatic patients and their parents. *Ann Allergy.* 1991;67(4):403–408.
3. Bender BG. Measurement of quality of life in pediatric asthma trials. *Ann Allergy Asthma Immunol.* 1996;77:438–447.
4. Juniper EF. Health-related quality of life in asthma. *Curr Opin Pulm Med.* 1999;5(2):105–110.
5. Annett RD. Assessment of health status and quality of life outcomes for children with asthma. *J Allergy Clin Immunol.* 2001;107(Suppl 5): S473–S481.
6. Juniper EF, Guyatt GH, Feeny DH, Ferrie PJ, Griffith LE, Townsend M. Measuring quality of life in the parents of children with asthma. *Qual Life Res.* 1996;5(1):27–34.
7. Everhart RS, Fiese BH. Asthma severity and child quality of life in pediatric asthma: a systematic review. *Patient Educ Couns.* 2009;75(2):162–168.
8. Juniper EF, Guyatt GH, Feeny DH, Griffith LE, Ferrie PJ. Minimum skills required by children to complete health-related quality of life instruments for asthma: comparison of measurement properties. *Eur Respir J.* 1997;10:2285–2294.

9. Juniper E. Paediatric Asthma Quality of Life Questionnaires (PAQLQ, PAQLQ(S), MiniPAQLQ and PACQLQ): Background, Administration and Analysis, QOL Technologies LTD, Apr 2006.
10. Zar, JH. *Biostatistical Analysis, 3rd edition*. Upper Saddle River, NJ: Prentice Hall Publishing; 1996:163–177.
11. Rosner BE. *Fundamentals of Biostatistics, 5th edition*. Pacific Grove, CA: Duxbury Thomson Learning; 2000:388–395.
12. Townsend M, Feeny DH, Guyatt GH, Furlong WJ, Seip AE, Dolovich J. Evaluation of the burden of illness for pediatric asthmatic patients and their parents. *Ann Allergy*. 1991;67(4):403–408.
13. Walker J, Winkelstein M, Land C, et al. Factors that influence quality of life in rural children with asthma and their parents. *J Pediatr Health Care*. 2008;22(6):343–350.

Open Access Medical Statistics

Dovepress

Publish your work in this journal

Open Access Medical Statistics is an international, peer-reviewed, open access journal publishing original research, reports, reviews and commentaries on all areas of medical statistics. The manuscript management system is completely online and includes a very quick and fair

peer-review system. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <http://www.dovepress.com/open-access-medical-statistics-journal>