

Research

Open Access

## Evaluating score distributions in the revised Dutch version of the Childhood Health Assessment Questionnaire

Jessika W Ouwerkerk<sup>1</sup>, Philomine A van Pelt<sup>2</sup>, Tim Takken<sup>3</sup>, Paul JM Helders<sup>3</sup> and Janjaap van der Net<sup>\*3</sup>

Address: <sup>1</sup>Utrecht School of Health Sciences, Utrecht University, Utrecht, The Netherlands, <sup>2</sup>Department of Pediatric Rheumatology, Wilhelmina Children's Hospital, University Medical Center Utrecht, Utrecht, The Netherlands and <sup>3</sup>Department of Pediatric Physical Therapy & Exercise Physiology, Wilhelmina Children's Hospital, University Medical Center Utrecht, Utrecht, The Netherlands

Email: Jessika W Ouwerkerk - [jessikaouwerkerk@hotmail.com](mailto:jessikaouwerkerk@hotmail.com); Philomine A van Pelt - [p.a.vanpelt@umcutrecht.nl](mailto:p.a.vanpelt@umcutrecht.nl);

Tim Takken - [t.takken@umcutrecht.nl](mailto:t.takken@umcutrecht.nl); Paul JM Helders - [p.j.m.helders@umcutrecht.nl](mailto:p.j.m.helders@umcutrecht.nl); Janjaap van der Net\* - [j.vandernet@umcutrecht.nl](mailto:j.vandernet@umcutrecht.nl)

\* Corresponding author

Published: 11 September 2008

Received: 3 April 2008

*Pediatric Rheumatology* 2008, **6**:14 doi:10.1186/1546-0096-6-14

Accepted: 11 September 2008

This article is available from: <http://www.ped-rheum.com/content/6/1/14>

© 2008 Ouwerkerk et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

### Abstract

**Objectives:** Evaluating the original, and the revised version of the Dutch Childhood Health Assessment Questionnaire (CHAQ). To explore the effect of different score calculation methods and eight more challenging items as proposed by Lam et al. (2004) on the score distribution in a population of patients with Juvenile Idiopathic Arthritis (JIA).

**Methods:** Two convenience samples of 59 and 31 children with JIA were studied. Box-and-whisker plots and the Kolmogorov-Smirnov (K-S) one-sample test of normality were used, to explore the score distributions.

**Results:** The results of this study confirm a ceiling effect when using the original CHAQ-30 with either score calculation method. The original CHAQ with the added eight more challenging items and the "mean" score calculation method, as well as the revised CHAQ showed less ceiling effect.

**Conclusion:** The original CHAQ-38 with the "mean" score calculation method as well as the revised CHAQ are a possible alternative for future studies. However, there is a need for further prospective studies to improve the CHAQ and to support our findings.

### Introduction

The Childhood Health Assessment Questionnaire (CHAQ) is the most widely utilized functional status measure in paediatric rheumatology today. The CHAQ consists of a disability index (30 items; 8 domains), and a discomfort scale (two visual analogue scales) and can be completed by children as well as their parents/guardians. The CHAQ has shown to be a valid, reliable, and sensitive functional status measure in children with Juvenile Idiopathic Arthritis (JIA) [1]. Over the years the use of the

CHAQ has also broadened to other childhood rheumatic conditions [2-5].

Despite its advantages and wide use, the CHAQ suffers from a ceiling effect [6]. Therefore it is difficult to discriminate distinct levels of function at the mild end of the disability continuum and to assess improvement in health for less impaired patients [6]. Lam et al. (2004) tried to influence this ceiling effect by testing different response options (visual analogue scale (CHAQ<sub>VAS</sub>), categorical

(CHAQ<sub>Cat</sub>), and choice (CHAQ<sub>Choice</sub>) and by adding eight more challenging items [6]. Respondents were instructed to compare their capabilities to that of their age peers over the last week. The different response options made it possible to assess not only patient's limitations (original CHAQ), but also the patient's strengths. Lam et al. added more challenging items so as to allow less impaired patients to score below the ceiling [4]. The results showed greater sensitivity, a more normal distribution, and a diminished ceiling effect for all three response options. The CHAQ<sub>Cat</sub> showed best concordance as a proxy report and might be easiest to complete.

The scoring rules applied to the original CHAQ are rather complex. The thirty items of the disability index assess eight domains of physical function. Of each domain three components are evaluated: (a) difficulty to perform each activity (0 = no difficulty, 1 = some difficulty, 2 = much difficulty, 3 = unable to do), (b) use of special aids or devices, and (c) required need for assistance of another person. The total score is the average of the highest score in each domain. If component b or c is scored, the minimum score for that domain is two. Takken et al. (2006) questioned the importance of these rules and explored the use of less complex score calculation methods [7]. Their results indicated that calculating the average of the thirty items improves sensitivity to change.

The main goal of this study was to evaluate the score distributions of the original [8] and the revised version of the Dutch CHAQ<sub>Cat</sub>. Furthermore, we wanted to explore the individual influence of different score calculation methods and the eight more challenging items as proposed by Lam et al. (2004) on these score distributions. We hypothesized 1: The revised CHAQ to have a more normal score distribution than the original CHAQ, 2: A less complex score calculation method to improve the score distribution of the original CHAQ, and 3: The eight more challenging items to have a positive influence on the score distributions of both questionnaires.

## Methods

### Questionnaires

The revised CHAQ with the categorical response option was translated and adapted following the absolutist approach with forward translations, consensus meetings, panel review, back translation, and, finally, authorisation by the developers [9].

Varying with score calculation methods and the eight more challenging items resulted in seven different (score-) versions of the CHAQ (Figure 1).

### Participants

Two convenience samples of children with JIA were studied. As we were studying the ceiling of the CHAQ, patients with remitting disease were also included. Sample A included 59 children from a transition of care study in patients between 12 and 25 years of age. They completed the original CHAQ and the eight more challenging items. Sample B were 31 children who attended the outpatient physical therapy department of the Wilhelmina Children's Hospital (WKZ) and Medical Center Utrecht, for a regular check-up between May 2005 and December 2005. They completed the translated version of the revised CHAQ.

### Statistical analysis

Data were entered and analysed using SPSS 11.5 for Windows. To evaluate the score distributions of the seven (score-) versions of the CHAQ, box-and-whisker plots and the Kolmogorov-Smirnov (K-S) one-sample test of normality was used. Box-and-whisker plots give a visual representation of the median, the quartiles, and the smallest and greatest value in the distribution. The K-S test statistic represents the largest absolute difference between the observed distribution and theoretical cumulative distribution functions. A p-value less than .05, was considered statistically significant. To analyse the differences between the groups on anthropometric parameters a Student t test was used.

## Results

### Participants

The 59 children from sample A had a mean age of 14.85 (8–25) years and a mean duration of joint complaints of 8.30 (0–20) years. They scored a median of .25 on the original CHAQ with a score range of 0 to 2.86. The 31 patients from sample B had a mean age of 10.81 (4–18) years and a mean duration of joint complaints of 2.90 (0–11) years. They scored a median of -.22 on the revised CHAQ with a score range of -1.55 to +1.05. The characteristics of the participants from both samples are summarized in Table 1. Group B was significantly younger ( $p < .0000$ ) and had significantly shorter disease ( $p < .0000$ ) compared to group A.

### Hypothesis 1

As hypothesized, the box-and-whisker plot of the revised CHAQ shows a more normal score distribution than that of the original CHAQ (Figure 2). This is also reflected in the K-S one-sample test of normality, with a p-value of .008 for the original CHAQ and .136 (statistically significantly normal) for the revised CHAQ.

### Hypothesis 2

We also hypothesized that replacing the original score calculation method for a less complex one would improve

	CHAQ-30	CHAQ-38	CHAQ-revised
<b>SCORE CALCULATION METHOD</b>	<b>Original</b>		<b>Revised</b>
	<ul style="list-style-type: none"> <li>• Score range: 0-3</li> <li>• The average of the highest scores of each domain (8)</li> <li>• Minimum domain score of 2 when components b or c (aid and assistance) are scored.</li> </ul>		<ul style="list-style-type: none"> <li>• Score range: -2 to +2</li> <li>• The average of the 38 individual items</li> </ul>
	<b>Alternative</b>	<b>Alternative</b>	<b>Revised-30</b>
	<ul style="list-style-type: none"> <li>• Score range: 0-3</li> <li>• The average of the highest scores of each domain (8)</li> </ul>	<ul style="list-style-type: none"> <li>• Score range: 0-3</li> <li>• The more challenging items form an extra domain</li> <li>• The average of the highest scores of each domain (9)</li> </ul>	<ul style="list-style-type: none"> <li>• Score range: -2 to +2</li> <li>• The average of the 30 individual items</li> </ul>
	<b>Mean</b>	<b>Mean</b>	
	<ul style="list-style-type: none"> <li>• Score range: 0-3</li> <li>• The average of the 30 individual items</li> </ul>	<ul style="list-style-type: none"> <li>• Score range: 0-3</li> <li>• The average of the 38 individual items</li> </ul>	

**Figure 1**  
**Score calculation methods of the seven different versions of the Childhood Health Assessment Questionnaire (CHAQ).**

**Table 1: Participant characteristics of sample A and B.**

	Sample A (n = 59; 18 ♂)	Sample B (n = 31; 16 ♀)
Age, mean (range)	14.85 (8–25)	10.81* (4–18)
Duration of joint complaints, mean (range)	8.34 (0–20)	2.90* (0–11)
Original CHAQ, median (range)	.25 (0–2.86)	-
Revised CHAQ, median (range)	-	-.22 (-1.55 – +1.05)

Sample A = patients from the transition of care study in patients with JIA; Sample B = patients who attended the outpatient physical therapy department of the Wilhelmina's Children Hospital and Medical Center Utrecht, for a regular check-up between May 2005 and December 2005. \*)  $p < .0000$

the score distribution of the original CHAQ. The box-and-whisker plots in the first column of Figure 2 however show the opposite. After omitting the questions about aids and assistance (Alternative) the ceiling effect raised. The ceiling effect rose even further when the average scores of thirty items were calculated (Mean).

**Hypothesis 3**

Finally, the eight more challenging items indeed have a positive influence on the score distributions of both questionnaires. The middle column of Figure 2 shows that adding the eight more demanding items to the Alternative and Mean version of the original CHAQ fully compensates for the raise in ceiling effect as seen in the first column. Removing the eight more challenging items from the revised CHAQ (Third column, Figure 2), shows a similar change in the score distribution as in the original CHAQ, raising the ceiling effect.

**Discussion**

We have shown that the revised version of the Dutch CHAQ<sub>Cat</sub> shows a significantly normal distribution statistically compared to the original Dutch CHAQ. However, opposed to our hypothesis, a less complex score calculation method of the original CHAQ did not improve the score distribution. Thirdly, the eight more demanding items as proposed by Lam et al. (2004) did have a positive influence on the score distribution of the original CHAQ as well as the revised CHAQ.

Our results are in agreement with the findings of Lam et al. (2004), but also show a crucial difference. The score distribution of the revised CHAQ was significantly normal statistically, but the score range was very narrow. This limits the applicability of the questionnaire in clinical research and clinical settings, because of a decreased ability to detect changes over time. This difference could be explained by the homogeneous groups in this study with only JIA patients. Lam et al. (2004) included patients with JIA, as well as other rheumatic disorders, injuries, fractures, spina bifida, and hemophilia with a history of haemarthroses. A second possible explanation is seen in the cultural differences. Dutch children seem to underesti-

mate their capabilities compared to Canadian children. Even though they could score between -2 and +2 (+1 and +2 perform better than peers) on the CHAQ<sub>Cat</sub>, almost all respondents scored below 0.

Our results did not concur with Takken et al. (2006), who concluded that the original score calculation method of the CHAQ could be replaced with a less complex one without clinical and psychometrical consequences.

A limitation of this study was that the samples we used were convenience samples consisting of different patients. Sample A was an existing historical data-set of children from a transition of care study in children with JIA between 12 and 25 years of age, and the data of sample B were gathered retrospectively from patient files during January and February of 2006. Both samples did not complete the original as well as the revised CHAQ, but only one or the other (of the two). Therefore the observed differences could in part be caused by differences in patient characteristics, such as age and duration of joint complaints.

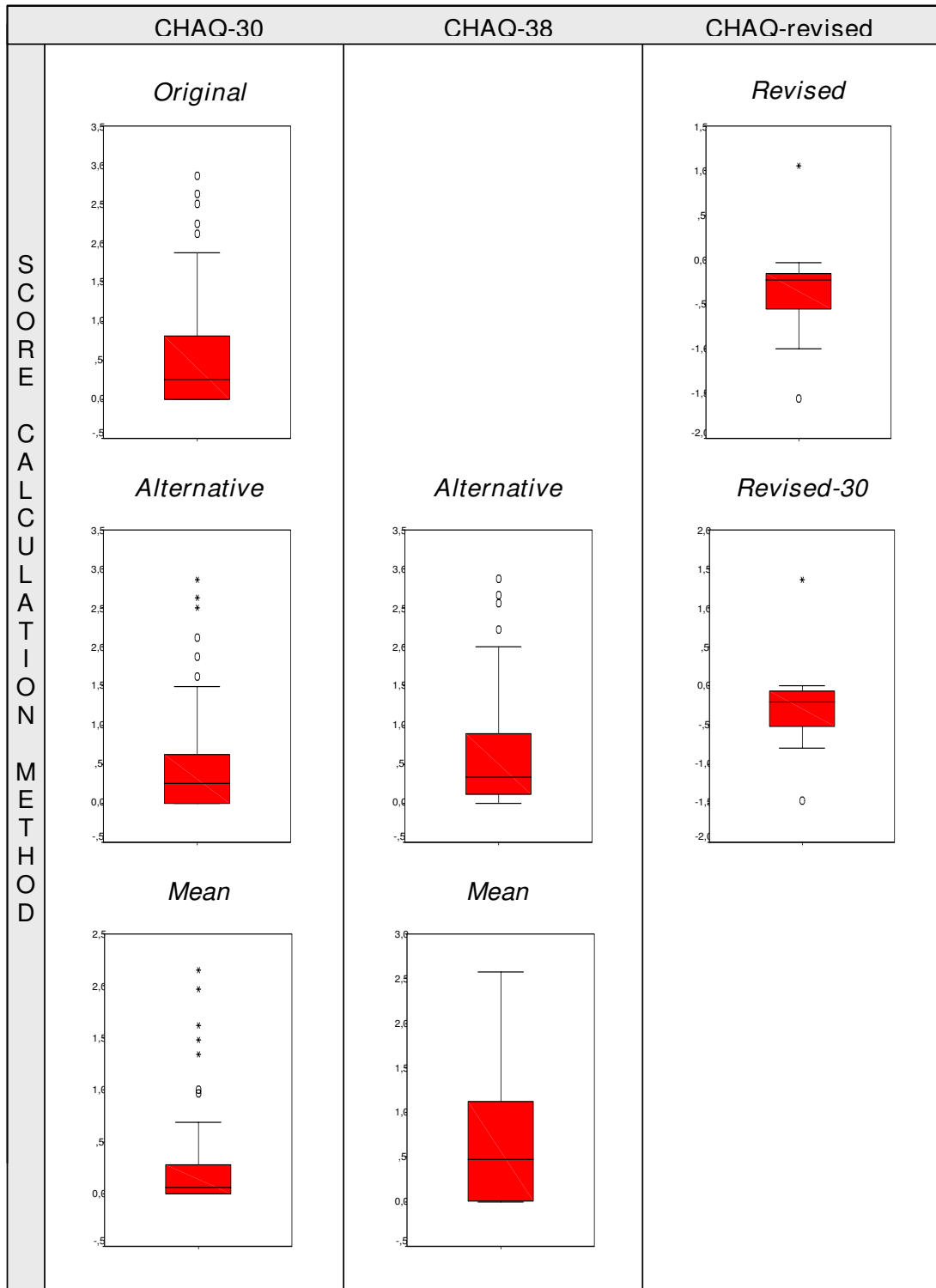
The results of this study confirm a ceiling effect using the original CHAQ-30 with either score calculation method. This emphasises the need for further exploration to improve the CHAQ. The original CHAQ-38, with the "mean" score calculation method as well as the revised CHAQ, show less ceiling and therefore are a better alternative for future studies in paediatric rheumatology. However, there is a need for further prospective studies to support our findings.

**Competing interests**

The authors declare that they have no competing interests.

**Authors' contributions**

JWO summarized the existing literature, prepared and analyzed data of samples A and B and prepared the draft of the manuscript. PhP offered the data set of sample A and supervised data analysis of this sample. TT and JN developed the methodology and supervised the statistical



**Figure 2**  
**Box-and-whisker plots of the seven different versions of the Childhood Health Assessment Questionnaire (CHAQ) Score calculations as described in Fig 1.**

procedures. PJMH facilitated the logistics, the use of the two datasets and senior read the manuscript.

## Acknowledgements

The authors would like to thank Dr. Brian M. Feldman, from the Department of Paediatric Rheumatology of SickKids, Toronto, Canada, for his contribution to this study.

## References

1. Singh G, Athreya BH, Fries JF, Goldsmith DP: **Measurement of health status in children with juvenile rheumatoid arthritis.** *Arthritis Rheum* 1994, **37(12)**:1761-1769.
2. Huber AM, Hicks JE, Lachenbruch PA, Perez MD, Zemel LS, Rennebohm RM, Juvenile Dermatomyositis Disease Activity Collaborative Study Group, et al.: **Validation of the Childhood Health Assessment Questionnaire in the juvenile idiopathic myopathies.** *J Rheumatol* 2001, **28**:1106-11.
3. Flato B, Aasland A, Vandvik IH, Forre O: **Outcome and predictive factors in children with chronic idiopathic musculoskeletal pain.** *Clin Exp Rheumatol* 1997, **15**:569-77.
4. Alman BA, Bhandari M, Wright JG: **Function of dislocated hips in children with lower level spina bifida.** *J Bone Joint Surg Br* 1996, **78**:294-8.
5. Bruce B, Fries JF: **The stanford health assessment questionnaire: a review of its history, issues, progress, and documentation.** *J Rheumatol* 2003, **30**:167-78.
6. Lam C, Young N, Marwaha J, McLimont M, Feldman BM: **Revised versions of the Childhood Health Assessment Questionnaire (CHAQ) are more sensitive and suffer less from a ceiling effect.** *Arthritis Rheum* 2004, **51(6)**:881-889.
7. Takken T, Eijkhof F van den, Hoijtink H, Helders PJM, Net J van der: **Examining the psychometric characteristics of the Dutch childhood health assessment questionnaire: room for improvement?** *Rheumatol Int* 2006, **26(11)**:979-983.
8. Wulffraat N, Net J van der, Ruperto N, Kamphuis S, Prakken BJ, Ten Cate R, van Soesbergen RM, van Rossum MA, Raat H, Landgraf JM, Kuis W: **The Dutch version of the Childhood Health Assessment Questionnaire (CHAQ) and the Child Health Questionnaire (CHQ).** *Clin Exp Rheumatol* 2001, **19**:S111-5.
9. Guillemin F, Bombardier C, Beaton D: **Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines.** *J Clin Epidemiol* 1993, **46(12)**:1417-1432.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:  
[http://www.biomedcentral.com/info/publishing\\_adv.asp](http://www.biomedcentral.com/info/publishing_adv.asp)

