Rheumatol Int (2006)
DOI 10.1007/s00296-006-0108-5

ORIGINAL ARTICLE

Tim Takken · F. den Eijkhof · H. Hoijtink
P.J.M. Helders · J. der Net

Examining the psychometric characteristics of the Dutch childhood health assessment questionnaire: room for improvement?

Received: 15 April 2005 / Accepted: 25 December 2005

Abstract The aim of this study was to examine the psychometric characteristics of the childhood health assessment questionnaire-disability index (CHAQ-DI). Seventy-six patients with juvenile idiopathic arthritis (JIA), age range 4.8–15.8 years, completed a CHAQ questionnaire one or more times. In total, 321 CHAQ questionnaires were available for analysis. Factor analysis and correlation were used to analyse the data. The analysis indicated that 12 items could be removed from the original 30 items of the CHAQ-DI. Also the addition of "aids and assistance" to the overall scoring method of the CHAQ-DI did not contribute to the overall measuring concept of the CHAQ-DI. The psychometric characteristics of the CHAQ-DI could be improved by removing 12 items from the original 30 items. Moreover, a simple scoring method, without the addition of aids and assistance to the total CHAQ-DI improves sensitivity to change of the CHAQ-DI.

Keywords Disability evaluation · Outcome assessment (Health care) · Juvenile rheumatoid Arthritis · Functional ability · Children

Introduction

In 1994 Singh and collaborators developed the childhood health assessment questionnaire [CHAQ (1)]. The CHAQ consists of 30 items in 8 domains, the disability index, combined with two visual analogue scales for pain and severity, forming the discomfort scale [1]. The CHAQ disability index [CHAQ-DI] was developed from the template of the Stanford HAQ-DI [2]. There were new questions added to the CHAQ-DI while others were modified to adapt the instrument for the use in children between the ages of 1–19 years. Over the years the CHAQ-DI has been translated in over a dozen languages and was utilized in more than 40 populations of different language/cultural background [3, 4]. In 1997 the CHAQ-DI became the functional outcome instrument of choice in the PRInTO core set of outcome measures [5, 6]. Therefore, the CHAQ-DI is to be considered one of the best-implemented functional outcome measures in paediatric rheumatology. The CHAQ-DI in studies has now has been administered to children with spina bifida [7], diabetes [4], JIA and JDM and has been validated in children with JIA and JDM [1, 8]. The reported psychometric characteristics of the CHAQ are very good [4], involving all the classical measures of reliability and validity as well as determining the smallest relevant clinical difference [9, 10]. So far only very few publications have dealt with the CHAQ-DI in a longitudinal nature [11, 12]. The scoring rules applied on the CHAQ-DI are a matter of another concern. When applied properly the rule is: in each of the eight domains, the highest score is taken, while the use of aids or assistance may add to the score if the difficulty completing a task is only 'limited'. However, in this way the patient is 'penalized' for utilizing help or a device. Also a patient shows no progression in his score domain when the worst score in a domain remains unchanged while other tasks or activities may have been improved considerably. This scoring rule might very well result in lack of sensitivity to change since aids and devices, once they are available in the household, patients will still use them even when they can perform these tasks without them as well. One can question if 'aids and devices' are important for the content validity of CHAQ-DI.

From a historical perspective, the current patient population with rheumatic diseases seems to be considerably less impaired as a result of new medication

T. Takken (✉) · P.J.M. Helders · J. der Net
Department of Pediatric Physical Therapy & Exercise Physiology, Wilhelmina Children's Hospital, University Medical Center Utrecht, Room KB2.056.0, 85090,
3508 AB Utrecht, The Netherlands
E-mail: t.takken@umcutrecht.nl
Tel.: +31-30-2504030
Fax: +31-30-2505333

F. den Eijkhof · H. Hoijtink
Department of Methods and Statistics, Faculty of Social Sciences, Utrecht University, Utrecht, The Netherlands

2 9 6 | 1 0 8 | B
Journal ID | Article ID
Dispatch: 13.1.06 Journal: 296 No. of pages: 5
Author's disk received ☑ Used ☑ Corrupted ☐ Mismatch ☐ Keyed ☐

regimes (e.g., MTX. and biologicals) [13]. This makes us question the face validity of several CHAQ-DI items, as some items score less and less positively and might turn out to be redundant. Especially 'aids and devices' become less prominent in these patient groups.

The aim of this study was to examine the psychometric characteristics of the childhood health assessment questionnaire-disability index [CHAQ-DI], and selecting those items that support the CHAQ-DI's underlying construct of functional ability. The second aim was to explore different scoring methods on the psychometric characteristics of the CHAQ-DI to increase sensitivity to change of the CHAQ-DI.

Methods

Patients

Seventy-six patients with JIA were included in this analysis; their age range was from 4.8 to 15.8 years (mean ± SD 9.19 ± 2.54 years). Twenty boys and 56 girls were included in this analysis. Thirty-three patients had an oligoarticular JIA, 43 patients had a polyarticular JIA including systemic JIA. In total, 321 CHAQ questionnaires were filled out by the parents, and were available for analysis. The children in this convenience sample participated between 1999 and 2002 in one of our previous conducted studies in our department [11, 14], or visited our outpatient clinic. Parents and patients gave their informed consent for participating in the study. The Institutional Review Board approved all procedures.

The Dutch translation of the CHAQ was used as a self-administered pencil and paper questionnaire for the parents (proxy), as an index of functional ability. The CHAQ [1] has been adapted from the Stanford health assessment questionnaire, so that at least one question in each domain was relevant to children of 6–19 years. The CHAQ was recently cross-culturally adapted and validated for the Dutch language [3]. The question with the highest score within each domain (range 0–3; able to do with no difficulty = 0; able to do with some difficulty = 1; able to do with much difficulty = 2; unable to do = 3; time frame for scoring responses was past week) determined the score for that domain, unless aids or assistance was required (raising the score for that domain to a minimum of two). The mean of the scores on the eight domains provided the CHAQ-DI (range 0–3). The lower the score the better the child's functional ability.

Statistics

Data were entered and analyzed using SPSS 11.5 for Windows. Factor analyses using the Oblimin method with Kaiser Normalization was used to analyze the data.

In the original scoring of the CHAQ, in each of the eight domains, the highest score is taken, while the use of aids or assistance may add to the score. Addition of the resulting eight highest scores into a sum-score is only reasonable if the eight scores are indicative of one underlying latent trait (functional ability), and, if the sum-score gives an indication of this trait (in this paper, the smaller the score the better a child's functional ability). The latter can be investigated using classical test theory [15] and factor analysis [16]. A large reliability implies that the sum-score is a reliable indicator of one underlying latent trait. The same implication holds if the proportion of variance explained by the first principal component resulting from a factor analysis is large, and the other eigenvalues are small.

Validity Statistics

The validity of the sum-scores presented in the previous section will be investigated using correlations across children using the visual analogue scales for pain and severity. Furthermore, sum-scores are only useful if they can be used to track a child's functional ability across time. Most of the children in the sample were evaluated with the CHAQ more than once. For each of the sum-scores the average partial correlation with pain and severity within children will be computed. The larger these are within correlations, the better the sum-score can track changes in discomfort as represented by pain and severity within children. Alpha level was set at $P < 0.05$ for all analyses.

Results

As can be seen in Table 1, analysis of the original scoring of the CHAQ-DI rendered a reliability of 0.88, and the first principal component explained 56% of the variation of the 8 domain-scores. All the other components explained rather small proportions of variance (that is, eigenvalues smaller than 1). Furthermore, the factor loading were around 0.8 for each of the eight domain-scores. The conclusion is that the sum-score resulting from the original scoring of the CHAQ-DI renders a reliable indication of one underlying latent trait, that is, a child's functional ability. Note that, if the factor loading had been unequal, a weighted sum-score might have been more appropriate. Weighting is achieved via multiplication of each standardized score with the corresponding factor loading. Another name for such a weighted sum-score is factor-score. Note also that in practice use of the sum-score is much easier than use of the weighted sum-score.

Another way to score the CHAQ-DI is to base the sum-score on each of the 30 items contained in the questionnaire. Such a sum-score has the advantage to make the CHAQ-DI more sensitive to changes in a

Table 1 Reliability of the different scoring methods of CHAQ-DI

Scoring	Reliability cronbach's alfa	Percentage of variance explained	Correlation with pain	Correlation with severity	Partial correlation with pain	Partial correlation with severity
Original	0.88	56	0.60	0.64	0.43	0.45
29 item	0.93	39 + 9 + 5	0.62	0.64	0.54	0.54
18 item	0.93	n/a	0.68	0.67	0.57	0.57

Note: On the 18-item model no factor analysis is performed, hence the percentage of variance explained is not applicable.

single item. The item 'clipping finger-nails' was excluded because it was not scored in almost half of the respondents. The reason is that most parents clip the nails of their children themselves. All subsequent analyses concern the 29 remaining items. Note that the information with respect to assistance is not included in this scoring. As can be seen in Table 1, the reliability of the 29 item-scores is 0.93. The first principal component explained 39% of the variation in the 29 item-scores, the second and third added 9 and 5%, respectively. The implication is that the sum-score is a reliable indicator of a dominant latent trait, but that the existence of other latent traits cannot be ruled out completely. Since no factor analysis is performed upon the remaining 18 items, the equivalent percentage of explained variance is not relevant or computed, as can be seen in Table 1.

The loadings of the first unrotated principal component were all around 0.6. The implication is that each item is equally indicative of the dominant latent trait, and that there is no reason to replace the sum-score by a weighted sum-score. The oblimin-rotated factor loadings of the first two and the first three principal components are displayed in Table 2. As can be seen, 18 items are selected that have factor loadings larger than about 0.30 in both rotations on the first rotated factor (one loading was 0.278). The reliability of the sum-score of these 18 items is 0.93, that is, equal to the reliability of the whole set of 29 items

As can be appreciated from Table 1, the sum-score obtained for 18 items has the largest correlation both with pain and severity, that is, larger than the original score, the factor-score for the first principal component, and the sum-score for 29 items. Both pain and severity were also predicted using the scores on the three oblimin-rotated factors displayed in Table 2 (note that the sum-score for 18 items was derived from the first oblimin-rotated factor). With respect to pain, the multiple correlations for the three factors were 0.71, with respect to severity 0.70. Although these numbers are slightly larger than the corresponding numbers for the 18 item sum-score, the improvement is so small that it is fair to say that both have the same validity.

The partial correlation (the average within correlation within children) between the sum-scores and pain and severity, can also be found in Table 1. Children, for whom the CHAQ was scored only once, were excluded during the computation of the partial correlations. Remaining were 63 children for who the CHAQ was scored at least three times (in total for these 63 children the CHAQ was scored 295 times). As can be seen, the 18 item sum-score has the highest partial correlations, that is, has the best performance in tracking pain and severity within children.

Discussion

This study showed that the psychometric characteristics of the Dutch CHAQ could be improved for the use in children with JIA by removing twelve redundant items from the original 30 items of this questionnaire. The removed 12 items do not contribute significantly to the final CHAQ-DI score. The addition of these items can even lead to a decrease in psychometric characteristics of the CHAQ-DI. The 18-item short form performance is superior to the traditional CHAQ-DI.

Moreover, our results indicate that the sensitivity to change of the CHAQ-DI could be improved using the total sum of the scores on individual items rather than the scoring original rule of the summation of domain scores. The addition of 'aids and assistance' information to the final score, does not contribute to the reliability and does not increase the correlation with pain and disease severity in patients with JIA. We feel that this part in the scoring procedure can be omitted from the CHAQ-DI scoring method without clinical and psychometrical consequences. In a previous study it was found that this new scoring method changed the sensitivity and distribution of the CHAQ scores by only a very small degree [17]. Moreover, this scoring method has also been used very recently by Lam et al. [18], who studied three revised versions of the CHAQ in a group of children with variety musculoskeletal disorders to increase sensitivity to change.

Looking at the oblimin-rotated factor loadings of the first two and three principal components, one can hypothesize that component #1 in Table 2 represents most likely lower extremity performance while component #2 could well represent upper extremity performance (dexterity) of the item-constructs. Only the constructs of item positions 17, 41 and 47 show contradictory results. Two mayor constructs may explain the third component in the analysis; item-positions 7, 34, 36, 37 and 38 are self-care/hygiene activities, while 47, 50, 54 and 58 are school-based activities. Note that in most Dutch primary schools turning faucets is a

Table 2 Factor analysis of the CHAQ-DI items

CHAQ positions	Item	Principal components		Principal components		
		1	2	1	2	3
6	*Dress, including tying shoelaces and doing buttons**	*0.336*	*0.498*	*0.385*	*0.453*	*−0.081*
7	Shampoo his/her hair	−0.218	0.834	−0.047	0.688	−0.376
8	*Remove socks*	*0.455*	*0.328*	*0.590*	*0.202*	*−0.218*
12	*Stand up from a low chair or floor*	*0.742*	*0.070*	*0.700*	*0.097*	*0.146*
13	*Get in/out bed or stand up in a crib*	*0.755*	*−0.031*	*0.826*	*−0.107*	*−0.053*
16	Cut his/her meat	−0.060	0.703	−0.162	0.800	0.142
17	*Lift up a cup or glass to mouth*	*0.343*	*0.197*	*0.368*	*0.170*	*−0.023*
18	Open a new cereal box	−0.038	0.722	−0.042	0.731	−0.035
21	*Walk outdoors on flat ground*	*0.867*	*−0.124*	*0.813*	*−0.090*	*0.190*
22	*Climb up five stairs*	*0.868*	*−0.042*	*0.865*	*−0.053*	*0.094*
34	Wash and dry entire body	0.038	0.747	0.193	0.612	−0.316
35	*Take a tub bath (get in and out of tub)*	*0.421*	*0.295*	*0.474*	*0.244*	*−0.069*
36	*Get on and off the toilet or potty chair*	*0.540*	*0.124*	*0.771*	*−0.092*	*−0.374*
37	Brush teeth	0.166	0.392	0.335	0.240	−0.310
38	Comb or brush hair	−0.056	0.715	0.061	0.615	−0.255
41	*Reach and get down a heavy object such as a large game or books from just above his/her head*	*0.556*	*0.285*	*0.491*	*0.336*	*0.0157*
42	*Bend down to pick up clothing or a piece of paper from the floor*	*0.719*	*0.131*	*0.838*	*0.013*	*−0.152*
43	*Pull on a sweater over his/her head*	*0.278*	*0.468*	*0.402*	*0.355*	*−0.222*
44	*Turn neck to look back over shoulder*	*0.403*	*0.187*	*0.474*	*0.119*	*−0.098*
47	Write or scribble with pen or pencil	0.349	0.210	0.170	0.368	0.350
48	*Open car door*	*0.388*	*0.420*	*0.323*	*0.476*	*0.136*
49	Open jars which have been previously opened	0.140	0.646	0.051	0.728	0.141
50	Turn faucets on and off	0.250	0.451	0.073	0.610	0.324
51	Open a door turning a key	0.319	0.509	0.182	0.631	0.255
54	*Run and play*	*0.912*	*−0.226*	*0.750*	*−0.096*	*0.396*
55	*Get in and out of a car, a toy car or school bus*	*0.571*	*0.107*	*0.437*	*0.220*	*0.294*
56	*Ride a bike or tricycle*	*0.747*	*−0.008*	*0.668*	*0.051*	*0.218*
57	Do household chores (e.g., wash dishes, take out trash, vacuuming, yard work, make bed, clean room)	0.037	0.535	0.045	0.531	−0.038
58	*Participate in physical education*	*0.688*	*−0.130*	*0.449*	*0.074*	*0.509*

Note: The CHAQ positions refer to the position of the items on the PRINTO version of the CHAQ as published in [19].
Note: The 18 items selected for the shorter form are given in italics.

common facility, while in most households a more expensive ''one lever'' faucet is more frequent.

Thus, the factor analysis showed that there seems to be two principal components of the CHAQ-DI. As previously suggested by Jackson et al. [19] the principal components of the CHAQ-DI are a lower extremity function component and an upper extremity function component. We can just speculate so far for determining the underlying third component of the CHAQ-DI in the second Factor analysis. It could consist of two constructs such as 'self care/hygiene' and 'school-based activities'. Future studies are needed to clarify this matter.

A recent study from Canada added eight new items to the CHAQ, seven items regarding sports activities and one item on dexterity [18]. These items also gave a better differentiation between controls and patients. These more physically challenging items can probably fit into component #1 or component #2, or maybe they measure a new component of the CHAQ, ''aerobic physical fitness''. In a previous study we found moderate to low correlations between CHAQ-DI scores and aerobic

physical fitness [20]. The addition of items on sports activities might improve its association with aerobic fitness.

A limitation in this study is that data are stemming from just one centre and of a relatively small cohort of JIA patients. This CHAQ-DI -short form should be obviously cross-culturally validated in the other languages in which the CHAQ-DI is already available, since items may have different loadings in different cultures. Moreover, the CHAQ-DI is not only used in JIA, but also in several other chronic disabling diseases. These diseases might have their own characteristics, and this probably could result in another short form. Thus, a cross-validation of the CHAQ-DI short form is needed in a larger and a more diverse population.

In conclusion, in this paper we examined the psychometric characteristics of the CHAQ-DI. From our analysis only 18 items could be selected that measure the CHAQ-DI's underlying construct of functional ability. We propose a CHAQ-ID consisting of only these 18 items, which has superior psychometric characteristics over the original 30-item CHAQ-DI. We also propose a simple scoring method of the sum of the scores on the 18 items, without the addition of aids and assistance to the total CHAQ-DI score to improve sensitivity to change of the CHAQ-DI.

References

1. Singh G, Athreya BH, Fries JF, Goldsmith DP (1994) Measurement of health status in children with juvenile rheumatoid arthritis. Arthritis Rheum 37:1761–9
2. Fries JF, Spitz P, Kraines RG, Holman HR (1980) Measurement of patient outcome in arthritis. Arthritis Rheum 23:137–45
3. Wulffraat N, van der Net JJ, Ruperto N, Kamphuis S, Prakken BJ, Ten Cate R, et al. (2001) The Dutch version of the Childhood Health Assessment Questionnaire (CHAQ) and the Child Health Questionnaire (CHQ). Clin Exp Rheumatol 19:S111–5
4. Bruce B, Fries JF (2003) The stanford health assessment questionnaire: a review of its history, issues, progress, and documentation. J Rheumatol 30:167–78
5. Giannini EH, Ruperto N, Ravelli A, Lovell DJ, Felson DT, Martini A (1997) Preliminary definition of improvement in juvenile arthritis. Arthritis Rheum 40:1202–9
6. Ruperto N, Martini A (2004) International research networks in pediatric rheumatology: the PRINTO perspective. Curr Opin Rheumatol 16:566–70
7. Alman BA, Bhandari M, Wright JG (1996) Function of dislocated hips in children with lower level spina bifida. J Bone Joint Surg Br 78:294–8
8. Huber AM, Hicks JE, Lachenbruch PA, Perez MD, Zemel LS, Rennebohm RM, et al. (2001) Validation of the Childhood Health Assessment Questionnaire in the juvenile idiopathic myopathies. Juvenile Dermatomyositis Disease Activity Collaborative Study Group. J Rheumatol 28:1106–11
9. Dempster H, Porepa M, Young N, Feldman BM (2001) The clinical meaning of functional outcome scores in children with juvenile arthritis. Arthritis Rheum 44:1768–74
10. Pouchot J, Ecosse E, Coste J, Guillemin F (2004) Validity of the childhood health assessment questionnaire is independent of age in juvenile idiopathic arthritis. Arthritis Rheum 51:519–26
11. Takken T, Van der Net J, Kuis W, Helders PJ (2003) Aquatic fitness training for children with juvenile idiopathic arthritis. Rheumatology 42:1408–1414
12. Lovell DJ, Giannini EH, Reiff A, Cawkwell GD, Silverman ED, Nocton JJ, et al. (2000) Etanercept in children with polyarticular juvenile rheumatoid arthritis Pediatric Rheumatology. Collaborative Study Group. N Engl J Med 342:763–9
13. Flato B, Aasland A, Vinje O, Forre O (1998) Outcome and predictive factors in juvenile rheumatoid arthritis and juvenile spondyloarthropathy. J Rheumatol 25:366–75
14. Takken T, van der Net J, Helders PJ (2001) Do juvenile idiopathic arthritis patients benefit from an exercise program? A pilot study. Arthritis Rheum 45:81–5
15. Carmines EG, Zeller RA (1979) Reliability and Validity Assessment. Sage, London
16. Tabachnick BG, Fidell LS (1996) Using Multivariate Statistics. Harper Collins College Publishers, New York
17. Whitney-Mahoney KJ, Young NL, Feldman BM, Badley EM (2001) Alternative scoring methods improve the responsiveness of the Childhood Health Assessment Questionnaire (CHAQ) [abstract]. Arthritis Rheum 44:S381
18. Lam C, Young N, Marwaha J, McLimont M, Feldman BM (2004) Revised versions of the Childhood Health Assessment Questionnaire (CHAQ) are more sensitive and suffer less from a ceiling effect. Arthritis Rheum 51:881–9
19. Jackson TR, Foley SP, Whitney-Mahoney K, Barron AC, Feldman BM, Brunner HI, et al. (2003) Measurement Properties of the Childhood Health Assessment Questionnaire (CHAQ) in the Evaluation of Physical Functioning in Children with Juvenile Rheumatoid Arthritis (JRA). Arthritis Rheum
20. Takken T, Van der Net J, Helders PJ (2003) Relationship between functional ability and physical fitness in juvenile idiopathic arthritis (JIA) patients. Scand J Rheumatol 32:174–178

			2	9	6			1	0	8	B	Dispatch: 13.1.06	Journal: 296	No. of pages: 5
	Journal ID					Article ID						Author's disk received ☑ Used ☑ Corrupted ☐ Mismatch ☐ Keyed ☐		