Validation of the Italian version of the Stanford Childhood Health Assessment Questionnaire for measuring functional status in children with chronic arthritis

F. Fantini, G. Corvaglia, P. Bergomi, M. Gattinara, V. Gerloni, C. Lomater, J. Mazzotti, G. Calori[1], L. Galli[1]

Chair of Rheumatology of the University of Milan, Pediatric Rheumatology Section, Gaetano Pini Institute, Milan; [1]Epidemiology Unit, S. Raffaele Hospital, Milan, Italy.

ABSTRACT. *The Stanford Health Assessment Questionnaire developed by Singh et al. to measure functional status in children with chronic arthritis (CHAQ) was translated into Italian (I-CHAQ), with minor modifications to obtain cross-cultural equivalence. This version was evaluated in a series of 96 patients with juvenile rheumatoid arthritis (JRA), both males and females ranging in age from 3 to 19 years (mean age 9.9 years). All three onset subtypes and all four classes of disability were represented in the sample.*

The questionnaire was filled in by the parents if the children were less than 8 years of age (23 cases), and by the children themselves in all other cases; a health professional was always present to provide assistance. As expected, JRA patients with a systemic or polyarticular disease onset had higher scores than those with a pauciarticular onset, and there were statistically significant differences in disability index values between patients from different Steinbrocker functional classes.

The instrument showed good reproducibility in a test-retest over a two-week period, a high correlation between the child and parent scores, excellent internal reliability, and good convergent and discriminant validity.

Key words: children, juvenile rheumatoid arthritis, disability, functional status, questionnaire, self-assessment, translation, validation.

Introduction

In the late 1980s Singh *et al.* (1, 2) developed a new parent and/or patient-administered questionnaire designed to assess the health status in children with JRA: the Childhood Health Assessment Questionnaire (CHAQ) adapted from the Stanford HAQ. Following our validation of the Italian version of the Functional Disability Index of the Stanford HAQ for adults with rheumatoid arthritis (3), we decided to translate the CHAQ into Italian, and to repeat the study reported by Singh's group in order to test its reliability and validity in an Italian setting. Singh *et al.* (4) recently published the results of a further study (4), and our findings in a larger Italian population were very similar to theirs, confirming that CHAQ is a rapid, reliable and valid instrument to assess functional status in children with chronic arthritis.

F. Fantini, MD, Associate Professor of Rheumatology, University of Milan; G. Corvaglia, MD; P. Bergomi, MD; M. Gattinara, MD; V. Gerloni, MD; C. Lomater, MD; G. Calori, MD; L. Galli, Statistician.

Please address correspondence and reprint requests to: Prof. Flavio Fantini, Cattedra di Reumatologia, Università di Milano, Piazza Cardinale Ferrari 1, 20122 Milan, Italy.

Table I. I-CHAQ, the Italian version of the CHAQ.

A) Dressing & grooming = Vestirsi e lavarsi

Il vostro bambino è capace di:

1 - Vestirsi, compreso allacciarsi le stringhe delle scarpe, abbottonarsi e sbottonarsi?
2 - Lavarsi i capelli?
3 - Togliersi le calze?
4 - Tagliarsi le unghie delle mani?

B) Arising = Alzarsi

Il vostro bambino è capace di:

5 - Alzarsi da una seggiolina o dal pavimento?
6 - Entrare e uscire dal letto o alzarsi in piedi nel lettino?

C) Eating = Mangiare

Il vostro bambino è capace di:

7 - Tagliarsi la carne?
8 - Portare alla bocca una tazza o un bicchiere?
9 - Scartare una merendina?

D) Walking = Camminare

Il vostro bambino è capace di:

10 - Camminare all'aperto su un terreno piano?
11 - Salire cinque gradini?

E) Hygiene = Igiene

Il vostro bambino è capace di:

12 - Lavare e asciugare ogni parte del suo corpo?
13 - Fare il bagno nella vasca, compreso entrare e uscire?
14 - Sedersi ed alzarsi dal water o dal vasino?
15 - Lavarsi i denti?
16 - Pettinarsi?

F) Reach = Prendere oggetti

Il vostro bambino è capace di:

17 - Prendere e tirar giù un oggetto pesante da un ripiano sopra la sua testa?
18 - Piegarsi a raccogliere un indumento o un pezzo di carta dal pavimento?
19 - Infilarsi una maglietta o un maglione dalla testa?
20 - Girare il collo per guardare indietro?

G) Grip = Afferrare e aprire oggetti

Il vostro bambino è capace di:

21 - Scrivere o scarabocchiare con una penna o una matita?
22 - Aprire la portiera dell'automobile?
23 - Svitare il coperchio di un barattolo già aperto in precedenza?
24 - Aprire e chiudere i rubinetti?
25 - Aprire una porta girando il pomello?

H) Activities = Attività varie

Il vostro bambino è capace di:

26 - Fare degli acquisti da solo?
27 - Entrare e uscire dall'automobile o dall'autobus della scuola o dalla macchinina?
28 - Andare in bicicletta o sul triciclo?
29 - Sbrigare faccende domestiche?
30 - Correre e giocare?

Patients and methods

Translation. The most recent version of the CHAQ was translated into Italian using the same method as for the adult instrument (3). Table I presents the Italian version of the CHAQ. The same set of questions, using the same wording, appeared in both the parent- and the child-administered questionnaires.

Scoring. The instrument is divided into 8 functional areas or categories. For each area 3 separate components are evaluated: the difficulty of performing a given task, the need for special aids or devices, and the need for assistance by another person. A total of 30 questions must be answered, each question being scored from 0 to 3 [0: without any difficulty ("senza alcuna difficoltà"); 1: with some difficulty ("con poca difficoltà"); 2: with much difficulty ("con molta difficoltà"); 3: unable to do ("non riesce")], as well as the response "not applicable" ("non applicabile") if the child is still too young to be able to perform a given activity. The highest score for any component question determines the score for that category. If aids or devices are used or help is needed to complete tasks in a certain area, a minimum score of 2 is recorded for the corresponding functional area.

The Disability Index (D.I.) is calculated by adding the scores for each of the categories and dividing by the number of categories answered. This yields a score in the 0 to 3.0 range.

Study sample. Ninety-six children, ranging in age from 1.8 to 19 years, affected by JRA according to the ARA criteria, were studied. Both sexes, all three onset forms of JRA, and the four Steinbroker classes of disability were represented in the sample. The main characteristics of the study subjects are reported in Table II.

Questionnaire administration. The Italian version of the CHAQ (I-CHAQ) was administered directly to those children 8 years of age or older, and to their parents if the patients were less than 8 years of age. In both situations the questionnaires were completed with the assistance of a trained health professional. On the same day a doctor, without knowledge of the questionnaire results, completed another form in which the main clinical fea-

Table II. Characteristics of the study subjects.

Sex:	Male	32
	Female	64
Age:	Less than 8 yrs. of age	23
	8 or more yrs. of age	73
	(mean age 9.9 years)	
JRA subtype:	Systemic	33
	Polyarticular	27
	Pauciarticular	36
Functional class:	I	46
	II	25
	III	17
	IV	8

tures of the patient were registered, including joint involvement (number and type of joint involvement), the Steinbrocker functional class, the Steinbrocker radiologic stage of the most severely affected joint, laboratory parameters and present drug therapy, in order to evaluate the spreading, severity and activity of the disease. A total of 79 joints were evaluated (TMJ 2, neck 1, upper limbs 19 x 2, lower limbs 19 x 2). Involved joints were defined as those with at least one of the following: pain or tenderness, heat, swelling, limited function, or deformation.

The questionnaire was also completed separately on the same day by one parent of each of 13 children aged 8 or more. In 26 cases the questionnaires were re-administered after a 2-week interval to the children (19 cases) or the parents (7 cases).

Reliability. The reliability of the questionnaire was estimated by three methods: the test-retest method after a 2-week interval both in children and parents; the inter-observer method by comparing the scores given by the children with those of their parents; and by evaluating the internal reliability of the instrument.

Validity. The validity of the questionnaire was estimated in three ways: the construct validity by evaluating the correspondence between the D.I. score (the functional status attributed by the child or his/her parent) and the Steinbrocker's functional class (physician-attributed functional status); the convergent validity by examining the correlations with other parameters related to the severity of the disease; and the discriminant validity by evaluating how well the D.I. was able to distinguish between patients with different degrees of disability.

Statistical analysis. All of the data from the completed forms was entered into a DB III database, and subjected to statistical analysis using the BMDP package. The Mann-Whitney test was applied for the comparison of groups of patients. The increase in the mean D.I. values as a function of the Steinbrocker class was evaluated by the test for linear trends with orthogonal polynomials. Spearman's correlation coefficients were used to evaluate the test-retest reliability, inter-observer correlation and convergent validity. Cronbach's alpha was calculated to test the internal reliability. Stepwise logistic analysis was used to test the discriminant validity.

Results

Disability Index. The means and standard deviations of the scores for each of the 8 categories and of the D.I. in the patients divided according to sex are reported in Table III. The pattern of disability was slightly different for the two sexes, with the males experiencing greater difficulty in dressing and grooming, eating, and reach, and the females in grip. These differences were not statistically significant, however.

Table IV reports the means and standard deviations of the scores for each of the 8 categories and of the D.I. for the patients divided according to onset type. Patients in the systemic group generally showed higher D.I. scores than those in the polyarticular group, although the difference

Table III. Mean scores (± SD) for the categories and the D.I., divided by sex.

	Males (n = 32)	Females (n = 64)
Dressing and grooming	1.47 ± 1.34	1.17 ± 1.11
Arising	0.84 ± 1.25	0.87 ± 1.02
Eating	0.72 ± 1.02	0.41 ± 0.77
Walking	0.69 ± 1.12	0.73 ± 1.03
Hygiene	1.06 ± 1.19	1.05 ± 1.16
Reach	1.25 ± 1.19	1.06 ± 1.10
Grip	0.88 ± 1.04	1.00 ± 1.09
Activities	1.03 ± 1.23	1.06 ± 1.16
D.I.	1.02 ± 0.97	0.94 ± 0.86

was not statistically significant, while there was a marked and statistically significant difference between the systemic and polyarticular groups, and the pauciarticular group (p < 0.01). The visualization of the mean scores for the categories in the three onset groups (Fig. 1) shows that the pattern of disability was similar in the systemic and polyarticular onset groups, while the pauciarticular onset group showed a different pattern.

Test-retest reliability. Test-retest reliability was evaluated in 26 cases: 19 children and 7 parents completed the questionnaire a second time after a two-week interval. There was a high and statistically significant correlation between the D.I. values in the two groups (Spearman's correlation coefficient = 0.94 for the children, p < 0.001; and 0.99 for the parents, p = 0.002).

Table IV. Mean scores (± SD) for the categories and the D.I., divided by onset type.

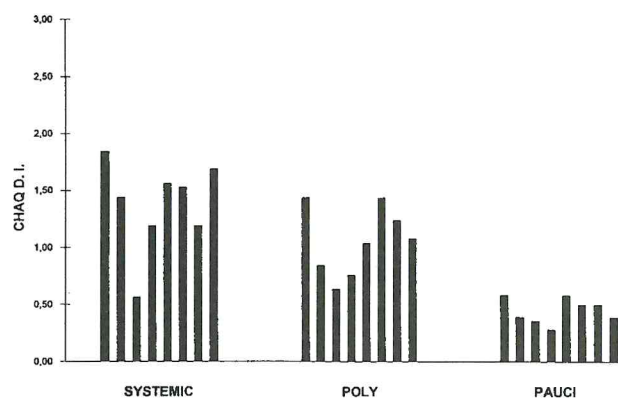|	Systemic (n = 32)	Polyarticular (n = 25)	Pauciarticular (n = 36)
Dressing & grooming	1.84 ± 1.27	1.44 ± 1.08	0.58 ± 0.84
Arising	1.44 ± 1.24	0.84 ± 1.07	0.39 ± 0.73
Eating	0.56 ± 0.95	0.64 ± 0.99	0.36 ± 0.72
Walking	1.19 ± 1.31	0.76 ± 1.09	0.28 ± 0.51
Hygiene	1.56 ± 1.24	1.04 ± 1.17	0.58 ± 0.91
Reach	1.53 ± 1.29	1.44 ± 1.04	0.50 ± 0.70
Grip	1.19 ± 1.20	1.24 ± 1.09	0.50 ± 0.77
Activities	1.69 ± 1.23	1.08 ± 1.22	0.39 ± 0.64
D.I.	1.42 ± 0.94	1.06 ± 0.89	0.45 ± 0.58

Fig. 1. Mean scores for the 8 categories of the CHAQ in the patients, divided by onset type: systemic, polyarticular and pauciarticular (each bar corresponds to a category).

Inter-observer reliability. When the D.I. values for 13 patients who completed the questionnaire themselves were compared to the D.I. calculated on the basis of their parent's responses, a high and statistically significantly correlation was found (Spearman's coefficient 0.91, $p < 0.001$).

Internal reliability. The internal reliability as measured by Cronbach's alpha for the 73 questionnaires completed by the children was 0.92, while for the 23 questionnaires answered by the parents it was 0.93. Both alpha coefficients were well above the 0.70 value generally accepted as the threshold value above which those items forming a scale reflect a unitary concept (5).

Construct validity. Table V shows the D.I. values (mean ± SD) for the patients divided by Steinbrocker functional class, as attributed by the physician at the moment the questionnaire was compiled. A highly significant linear trend in the mean D.I. values was observed along the functional classes.
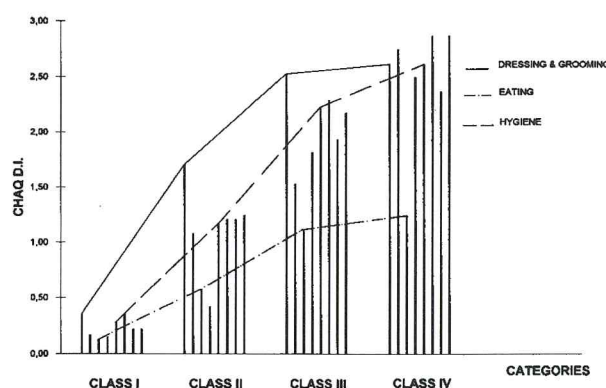

Fig. 2. Mean scores for the 8 categories of the CHAQ in the patients, divided according to Steinbrocker functional class (each bar corresponds to one category).

Table V. Disability index according to functional class.

Class	No. cases	Mean ± SD
I	46	0.24 ± 0.32
II	25	1.13 ± 0.46
III	17	1.96 ± 0.56
IV	8	2.51 ± 0.35

$p < 0.001$

A very good correlation was also observed between the D.I. based on the children's or their parents' evaluations, and the functional class (Spearman's coefficient 0.90 for the children and 0.75 for the parents, $p < 0.001$).

As shown in Figure 2, the mean values for the 8 categories varied somewhat differently with the degree of impairment as evaluated by Steinbrocker's functional class: the earliest and most severely impaired category was category 1 (dressing and grooming), while the latest and least severly impaired was category 3 (eating); the other categories, such as category 5 (hygiene) showed an intermediate pattern.

Convergent validity. Table VI shows the correlation coefficients and the statistical significance of the correlations between the D.I. values as assessed by the children and by the parents respectively, and other parameters of disease severity recorded at the time the questionnaire was administered. In both groups there was a statistically significant correlation between the D.I. and the number of involved joints, the hemoglobin level (negative correlation), and the ESR. The correlation between the D.I. derived from the children's questionnaires and the Steinbrocker's radiologic stage of the worst joint was significant, while there was no significant correlation between the D.I. and disease duration.

In the patients on drug treatment and therefore presumably with more active disease, the D.I. values were

Table VI. Correlation between the D.I. based on the children's and the parents' evaluations, and other parameters of disease severity.

	Children (n = 73)	Parents (n = 23)
No. of involved joints	0.60 **	0.74 **
Hemoglobin	-0.60 **	-0.60 *
ESR	0.52 **	0.56 *
Radiologic stage	0.40 **	0.18
Disease duration	0.03	0.36

Spearman's correlation coefficient: * $p < 0.01$; ** $p < 0.001$.

Table VII. Stepwise logistic analysis of the D.I. according to the functional class.

Functional class	No.	Exp. (coeff.) = r.r.	Coeff./SE Coeff.	p
Class I	29			
> Class I	44	759	3.77	< 0.05
Class II	22			
> Class II	22	45.8	3.61	< 0.05
Class III	14			
Class IV	8	11.9	1.74	0.08

r.r. = relative risk

higher than in those not on drug therapy (1.09 ± 0.89 versus 0.37 ± 0.72 for the children who completed the questionnaire; 1.16 ± 0.92 versus 0.25 ± 0.21 for the parents who completed the questionnaire). This difference was statistically significant only for the children ($p = 0.0012$), probably due to the small sample of parents evaluating patients not under drug treatment ($n = 6$).

Discriminant validity. As shown in Table VII, the discriminant validity for a unit increase in the D.I. was clearly demonstrated by stepwise logistic analysis for functional class I versus classes II-IV (relative risk = 759; $p < 0.05$), and for class II versus classes III and IV (r.r. = 45.8; $p < 0.05$). Probably due to the small number of patients in classes III and IV, the discriminant validity between these latter classes was poor (r.r. = 11.9; $p = 0.08$).

Discussion

Although several validated function and outcome measures exist for the adult rheumatic diseases, pediatric patients are usually not considered when these measures undergo reliability testing. One pediatric trial of the physical disability and pain dimensions of the AIMS yielded disappointing results (6).

The first normalized disability assessment tool for JRA patients was the JAFAS (Juvenile Arthritis Functional Assessment Scale), which showed excellent internal and convergent validity and internal reliability (7). However, the JAFAS had the disadvantage of requiring professional personnel and special equipment for its administration.

From this instrument two questionnaires were developed, one based on patient reports (JAFAR-C) and one on parent reports (JAFAR-P): the respondents had to rate the patient's current ability to perform 23 activities. The patient reports and parent reports were found to correlate highly with each other, and with an objective assessment performed by therapists (8). The JAFAR-C questionnaire is administered to JRA patients by a social worker; the minimum age for self-reporting was established at 7 years, although the minimum age at which a child is able to self-administer a questionnaire remains undefined.

The Stanford Health Assessment Questionnaire (HAQ) is one of the most widely used self-administered questionnaires for quantifying disability in terms that represent the patient's perceptions of this entity. The Stanford-Philadelphia Group led by Singh developed an adaptation for children of the HAQ (the CHAQ). In their questionnaire Singh *et al.* followed the design characteristics of the adult instrument, in order to maintain compatibility with the ARAMIS databases, to facilitate the comparison of functional status in various rheumatic diseases, and to allow clinicians to follow patients into adulthood using the same conceptual and questionnaire structure (4). The CHAQ was adapted for use in children by adding several new questions, so that for each functional area there is at least one question that is relevant to children of all ages.

To eliminate the discrepancies introduced by growth and development, only those difficulties caused by arthritis are to be noted, and each child's score is based on the performance of age-appropriate activities. In fact, the question with the highest score determines the score for a given category and therefore the older children are scored on a more complex activity than the younger ones and their scores may be considered analogous.

The CHAQ was first administered to parents of JRA patients, and was shown to be a valid and reliable instrument (1). In another study in which it was administered to parents of JRA patients on at least two occasions, the CHAQ Disability Index was found to be a more sensitive measure than either morning stiffness or the joint count (2). A high concordance was also demonstrated between assessments performed by children aged 8 - 19 years and their parents during the same visit (4).

The performance of the earliest version of the CHAQ (eight-category disability dimension assessed using a total of 21 questions rather than the present total of 30) was investigated in a population-based prospective follow-up study of patients with JCA carried out in southwestern Sweden from 1987 to 1992 (9). In this study the parents, and the patients if they were 9 years of age or older, completed the Swedish version of CHAQ independently. Test-retest and inter-observer correlations, as well as the internal reliability, were excellent. Convergent validity was demonstrated by strong correlations of the disability index with disease activity and the Steinbrocker functional class. Discriminant validity was demonstrated by the capacity of the instrument to evaluate patients as being active or in

remission. In the Swedish study the mean scores calculated from the patients' questionnaires were consistently higher, i.e. worse, than the scores based on the parents' responses, indicating that the data were not completely interchangeable. A recent study by an Irish Group using the same instrument confirmed a high level of agreement between children aged 8 to 16 years, and their mothers with respect to disability (not including pain) (10).

In the early 1990's an Italian group of rheumatologists and biostatisticians translated into Italian the Functional Disability Index of the HAQ, and validated this version in a series of adults with rheumatoid arthritis (3). The same methodology was applied to the translation into Italian of the latest version of the CHAQ.

This study of the reliability and validity of I-CHAQ has been conducted along the same lines as those of Singh's study. Unlike the Swedish study, which was carried out on two consecutive series of relatively older children (mean age 17.9 and 13 years, respectively, in the two studies) with an under-representation of systemic JRA and only a few subjects in functional classes III and IV, our study sample was carefully selected to represent both sexes (M/F ratio = 1/2), all three subtypes of JRA (about 1/3 each), and all 4 classes of disability. All the patients were less than 20 years of age; the mean age was about 10 years and the minimum age 1 year and 9 months. Therefore, the I-CHAQ was tested in a true juvenile setting and covered the entire spectrum of disease and disability.

In our study the questionnaires were filled in by the children themselves if they were 8 years of age or older, with the assistance of a trained health professional. In general, once the child becomes acquainted with the procedure he requires less and less assistance. For children less than 8 years of age the parents, in general their mothers, completed the questionnaire. In our experience, parents of very young children tended to overestimate their disabilities. Nevertheless, the correlation between the D.I. values derived from the responses of the children and those of their parents was high and statistically significant in our study.

The D.I. values did not vary significantly with the sex of the patient, although the pattern of disability based on the 8 functional areas was somewhat different between the two sexes. A more definite picture may emerge from studies on larger populations of patients.

As expected, higher disability scores were recorded for the systemic and polyarticular subtypes than for the pauci-articular subtype. Moreover, the disability pattern was fairly similar for the two more severe forms, while in the patients with pauciarticular JCA some categories such as dressing, eating, walking, and miscellaneous activities were only minimally impaired (Fig. 1).

Our study also clearly underlined how the impairment of the 8 functional areas differed in pattern and weight in the 4 functional classes, the most sensitive being dressing and grooming, and the least eating (Fig. 2).

The performance of I-CHAQ in our study was very similar to that of Singh's and the Swedish studies: its reliability was demonstrated by excellent test-retest and inter-observer correlations and high Cronbach's alpha values. Construct, convergent and discriminant validity were also confirmed in our series of patients. Studies are now underway to assess the sensitivity of I-CHAQ in children with chronic arthritis who are on long-term treatment with methotrexate or cyclosporin.

In conclusion, the Italian version of the CHAQ appears to be a rapid, reliable and valid instrument to assess the functional status of children of any age with chronic arthritis. It can be self-administered in children aged 8 or older, while parents can serve as reliable proxy-reporters for younger children. Its principal advantages are that it takes only 10 - 20 minutes to complete and less than 2 minutes to code, and that it maintains the same conceptual and questionnaire structure as the instument for adult patients. We would recommend that the CHAQ be translated into other languages, as this would facilitate international multicenter studies on JRA outcome and JRA-bound disability in adulthood.

Acknowledgements

References

1. SINGH G, ATHREYA B, FRIES J, GOLDSMITH D, OSTROY B: Measurement of functional status in Juvenile Rheumatoid Arthritis. *Arthritis Rheum* 1990; 33 (Suppl. 9): 515.
2. SINGH G, BROWN B, ATHREYA B, GOLDSMITH D, RETTIG P, BLOCH D, FRIES J: Functional status in Juvenile Rheumatoid Arthritis: sensitivity to change of the childhood health assessment questionnaire. *Arthritis Rheum* 1991; 34 (Suppl. 9): 581.
3. RANZA R, MARCHESONI A, CALORI G, BIANCHI G, BRAGA M, CANAZZA S, CANESI B, FUMAGALLI M, MASTAGLIO C, MATHIEU A, FANTINI F: The Italian version of the Functional Disability Index of the Health Assessment Questionnaire. A reliable instrument for multicenter studies on rheumatoid arthritis. *Clin Exp Rheumatol* 1993; 11: 123-8.
4. SINGH G, ATHREYA BH, FRIES JF, GOLDSMITH DP: Measurement of health status in children with Juvenile Rheumatoid Arthritis. *Arthritis Rheum* 1994; 37: 1761-9.
5. CARMINES EG, ZELLER RA: Reliability and validity assessment. *Sage University Paper Series on Quantitative Applications in the Social Sciences, Series n. 07-017.* Sage Publications, Beverly Hills and London, 1979.

6. COULTON CJ, ZBOROWSKY E, LIPTON J, NEWMAN AJ: Assessment of the reliability and validity of the Arthritis Impact Measurement Scales for children with juvenile arthritis. *Arthritis Rheum* 1987; 30: 819-24.

7. LOVELL DJ, HOWE S, SHEAR E, HARTNER S, McGIRR G, SCHULTE M, LEVINSON J: Development of a disability measurement tool for Juvenile Rheumatoid Arthritis. The Juvenile Arthritis Functional Assessment Scale. *Arthritis Rheum* 1989; 32: 1390-5.

8. HOWE S, LEVINSON J, SHEAR E, HARTNER S, McGIRR G, SCHULTE M, LOVELL D: Development of a disability measurement tool for juvenile rheumatoid arthritis. The Juvenile Arthritis Functional Assessment Report for Children and their Parents. *Arthritis Rheum* 1991; 34: 873-80.

9. ANDERSSON GARE B, FASTH A, WIKLUND I: Measurement of functional status in juvenile chronic arthritis: evaluation of a Swedish version of the Childhood Health Assessment Questionnaire. *Clin Exp Rheumatol* 1993; 11: 569-76.

10. DOHERTY E, YANNI G, CONROY RM, BRESNIHAN B: A comparison of child and parent ratings of disability and pain in Juvenile Chronic Arthritis. *J Rheumatol* 1993; 20: 1563-6.