

Evaluation of the Multidimensional Health Assessment Questionnaire in Finnish Patients with Rheumatoid Arthritis

MARJA ARKELA-KAUTIAINEN, HANNU KAUTIAINEN, TOINI UUTELA, KARI LAIHO, HARRI BLÅFIELD, MARJATTA LEIRISALO-REPO, and MARKKU HAKALA

ABSTRACT. Objective. To study the characteristics of the Multidimensional Health Assessment Questionnaire (MDHAQ) in Finnish patients with rheumatoid arthritis.

Methods. The reliability of the questionnaire was tested by test-retest procedure. Construct validity was studied by factor analysis and convergent validity by calculating correlations between the Finnish MDHAQ (Finn-MDHAQ) scales and the Finnish Health Assessment Questionnaire (HAQ) and the Finnish Arthritis Impact Measurement Scales (Finn-AIMS2). Correlations between Finn-MDHAQ and measures of clinical characteristics, disease activity, and functional class were also measured. An item analysis was made of the Finn-MDHAQ scales Function (FN) and Psychological (PS).

Results. Internal consistency on the FN scale was 0.92 (95% lower limit 0.89) and 0.66 (0.56) on the PS scale. Reproducibility (95% CI) on FN was 0.93 (0.82 to 0.97) and on PS 0.84 (0.70 to 0.92). Factor analysis identified 2 factors, mobility of upper extremities and trunk, and mobility of lower extremities. Strong correlations were found between the FN scale and HAQ and physical subscales of Finn-AIMS2 and between PS and the psychological subscales of Finn-AIMS2. In item analysis corrected item correlation was high on the Finn-MDHAQ scales, except in one item on the PS scale.

Conclusion. The Finn-MDHAQ is an applicable, reliable, and valid instrument for the part of the FN scale measuring functional ability in Finnish rheumatic patients. The incongruity in the PS scale structure that produced moderate internal consistency can be overcome with minor modifications. (J Rheumatol 2005;32:1426–31)

Key Indexing Terms:

EVALUATION VALIDITY RELIABILITY RHEUMATOID ARTHRITIS
MULTIDIMENSIONAL HEALTH ASSESSMENT QUESTIONNAIRE

The Multidimensional Health Assessment Questionnaire (MDHAQ)¹ was published by Pincus and coworkers in 1999 as an assessment tool for clinical care of patients with rheumatic diseases. The MDHAQ is derived from the Health Assessment Questionnaire (HAQ)² and its modified version, MHAQ³. The HAQ was published in 1980 and has been translated or culturally adapted into over 60 different languages⁴, and has also been part of the core set of physical function measures for clinical care of patients with rheuma-

toid arthritis (RA) from 1993⁵. The Finnish version of HAQ was published in 1993⁶. The initial HAQ was further developed to include questions concerning perceived patient satisfaction and perceived change in degree of difficulty, termed the Modified Health Assessment Questionnaire³. Pincus and coworkers reported¹ that this version was further modified to compensate for the phenomenon of floor effect, and additionally to assess the physical status and also the psychological aspects of patient outcomes. Our objective was to translate the MDHAQ into Finnish, and evaluate its characteristics in Finnish patients with RA.

MATERIALS AND METHODS

Patients. Participants for the study were recruited from 3 different institutions treating rheumatic patients, 2 of them central hospitals (Central Hospitals of Lapland, Rovaniemi, and South Ostrobothnia, Seinäjoki, outpatient departments) and the third the Rheumatism Foundation Hospital (patients treated on the ward). As inclusion criteria, all participants fulfilled the 1987 American Rheumatism Association (ARA) criteria⁷ for RA. We included patients with a definite and stable diagnosis, with arthritis diagnosed ≥ 3 years previously. Participants were ≥ 16 years old at time of diagnosis, and had had no changes in their disease modifying antirheumatic medication for the previous 2 months.

Patient questionnaire. The Internet version of the MDHAQ questionnaire⁸ used in this study consists of 8 items from MHAQ³ with questions on activ-

From the Rheumatism Foundation Hospital, Heinola; Central Hospital of Lapland, Rovaniemi; Central Hospital of South Ostrobothnia, Seinäjoki; and Department of Medicine, Division of Rheumatology, Helsinki University Central Hospital, Helsinki, Finland.

Supported by the Rheumatism Foundation Hospital Research Fund, Heinola, Finland.

M. Arkela-Kautiainen, MSc, PT; H. Kautiainen, BA, Statistician; K. Laiho, MD, PhD; M. Hakala, MD, PhD, Rheumatism Foundation Hospital, Heinola; T. Uutela, MD, Central Hospital of Lapland, Rovaniemi; H. Blåfield, MD, Central Hospital of South Ostrobothnia, Seinäjoki; M. Leirisalo-Repo, MD, PhD, Professor, Helsinki University Central Hospital.

Address reprint requests to M. Arkela-Kautiainen, Likolahdenkatu 43, FIN-44150 Äänekoski, Finland.

E-mail: marja.arkela-kautiainen@medcare.fi

Accepted for publication February 7, 2005.

ities of daily living (ADL) and 2 new items concerning advanced function (here termed the Function scale, FN), and 3 items on Psychological stress (Psychological scale, PS). The translation of the initial questionnaire into Finnish was first undertaken independently by 2 persons, an experienced rheumatologist (MH) and an experienced physiotherapist (MAK), following the standardized guidelines for the process of cross-cultural adaptation of self-report measures^{9,10}. A synthesis of the 2 translations was made after discussion by the initial translators and one experienced rheumatologist (MLR). Back-translation of the translated questionnaire into English was made blindly by a professional translator. Then the same group who made the synthesis reached agreement on the final version of the Finnish MDHAQ (Finn-MDHAQ).

Evaluation methods. Patients completed the Finn-MDHAQ questionnaire and were also asked to complete the Finnish HAQ^{6,11} and the Finnish version of the Arthritis Impact Measurement Scales 2 (Finn-AIMS2)¹². Raw responses to the Finn-MDHAQ scales were recoded from 1–4 to 0–3, 0 representing good functional ability and 3 the poorest. A physician clinically evaluated the patient by completing the physician's overall assessment, using a visual analog scale (VAS) and ARA functional classification¹³. Disease activity was measured by erythrocyte sedimentation rate (ESR, mm/h) and assessment of the number of swollen and tender joints. The reproducibility of the questionnaire was tested on patients, who had no change in therapy and received no intraarticular injections with glucocorticoids, during the visit. These patients were also asked to complete the Finn-MDHAQ questionnaire 2 weeks later and return it by mail.

Statistical methods. Results are expressed as mean or median, standard deviation (SD), or interquartile range (IQR), with 95% confidence intervals (95% CI). The “floor value” is defined in this study as the worst possible value of the item or as the minimum total value of the scale, and the “ceiling value” is the best possible value of the item or the maximum total value of the scale. The reliability of the 2 scales in the questionnaire was evaluated by calculating the intraclass correlation coefficient (ICC) and coefficient of repeatability, with bias corrected and accelerated bootstrapping (5000 replications) confidence intervals. Internal consistency was estimated by calculating Cronbach's alpha internal consistency with 95% one-sided CI. Convergent validity was assessed by calculating the relationship of the Finn-MDHAQ scales and HAQ and the Finn-AIMS2 scales. Construct validity was studied by MINRES factor analysis with promax-rotation for the Finn-MDHAQ items matrix of polychoric correlations. Item analysis of the Finn-MDHAQ scales was performed by analyzing item discriminating power (corrected item correlation) and item difficulty (item mean) depicted by explanatory data analysis. Corrected item correlation was estimated using polychoric or polyserial correlations, and correlation coefficients were calculated by the Spearman method, using Sidak-adjusted probabilities.

RESULTS

A total of 123 patients, 103 women (84%) and 20 men (16%), completed the questionnaires. Table 1 shows the demographic and clinical characteristics of the study group. The mean age of the patients was 56 years, mean duration of disease 20 years.

Table 2 shows the response rates and mean scores (SD) on every Finn-MDHAQ item and also valid percentages of the floor and ceiling values on both scales of the instrument. The response rate on the Finn-MDHAQ scales varied from 95% to 100%. The mean (SD) score on FN was 1.03 (0.63) and on PS 0.80 (0.52). Floor and ceiling values on FN were 4% and 1%, and on PS 13% and 1%, respectively. In the HAQ, which was also completed by the respondents, the floor and ceiling values were 4% and 2%, respectively. We

Table 1. Demographic and clinical characteristics of 123 RA patients.

Variables	Measures
Female/male, n (%)	103 (84)/20 (16)
Age, yrs, mean (SD)	56 (11)
Duration of disease, yrs, mean (range)	20 (3–53)
Rheumatoid factor present, n (%)	103 (84)
ESR (mm/h), median (IQR)	21 (13, 33)
No. of swollen joints, median (IQR)	5 (2, 16)
No. of tender joints, median (IQR)	7 (2, 17)
Patient overall assessment, VAS, median (IQR)	49 (28, 67)
Physician overall assessment, VAS, median (IQR)	46 (30, 58)
Patient assessment of pain, VAS, median (IQR)	48 (25, 69)
DAS28, mean (SD)	4.41 (1.33)
HAQ score, median (IQR)	1.25 (0.75, 1.75)
Functional class, n (%)	
I	6 (5)
II	50 (41)
III	61 (49)
IV	1 (1)
Not known	5 (4)
Education level, n (%)	
Primary school	46 (37)
Secondary school	33 (27)
College	31 (25)
Graduate school	7 (6)
Not known	6 (5)

ESR: erythrocyte sedimentation rate, IQR: interquartile range, VAS: visual analog scale, HAQ: Health Assessment Questionnaire, DAS: Disease Activity Score.

found correlation sized 0.37 (95% CI 0.20 to 0.51) between the FN and PS scales of the Finn-MDHAQ.

Internal consistency values (95% CI lower limit) were 0.92 (0.89) and 0.66 (0.56) on FN and on PS, respectively. Thirty-two patients completed the Finn-MDHAQ retest questionnaire. The reproducibility (95% CI) of the Finn-MDHAQ on FN was 0.93 (0.82 to 0.97) and on PS was 0.84 (0.70 to 0.92) (Table 3).

Factor analysis performed for construct validity showed the FN scale was loaded on 2 factors: mobility of upper extremities and trunk and mobility of lower extremities. These factors explained 61% of the total variance (Table 4).

Item analysis of the FN scale (Figure 1) showed that all items had high corrected item correlation, but one (“Participate in sports and games as you would like”) had a higher item mean than the others. Item analysis of the PS (Figure 2) showed that 2 items had high corrected item correlation and one (“Get a good night's sleep”) had low correlation.

For convergent validity, we found strong correlations between the FN scale and HAQ and between FN and all the subscales of the Finn-AIMS2 except 2 social subscales (Table 5). In view of the moderate internal consistency in the PS scale and clear differences in item analysis between PS items we studied the relationship of PS and HAQ, Finn-AIMS2, and clinical characteristics in 2 different approach-

Table 2. Characteristics for the Finn-MDHAQ items and scales.

Item/Scale	Score, mean (SD)	Range	Response Rate, %	Floor*, %	Ceiling **, %
Function (FN)					
a. Dressing yourself, including tying shoelaces and doing buttons?	1.02 (0.71)	0–3	100	20	4
b. Get in or out of bed?	0.59 (0.72)	0–3	100	53	2
c. Lift a full cup or glass to your mouth?	0.72 (0.75)	0–3	100	44	2
d. Walk outdoors on flat ground?	0.59 (0.68)	0–3	100	50	1
e. Wash and dry your entire body?	0.78 (0.73)	0–3	99	38	2
f. Bend down to pick up clothing from the floor?	0.79 (0.83)	0–3	100	42	5
g. Turn regular faucets on and off?	1.00 (0.90)	0–3	98	35	6
h. Get in or out of a car, bus, train, or airplane?	1.03 (0.88)	0–3	99	30	7
i. Walk 2 miles?	1.56 (1.09)	0–3	98	19	28
j. Participate in sports and games as you would like?	2.30 (0.90)	0–3	95	2	72
Mean 10-item FN score	1.03 (0.63)	0–3		4	1
Psychological (PS)					
k. Get a good night's sleep?	1.04 (0.71)	0–3	99	18	5
l. Deal with the feelings of anxiety or being nervous?	0.71 (0.66)	0–3	100	41	0
m. Deal with the feelings of depression or feeling blue?	0.67 (0.62)	0–3	100	41	0
Mean 3 item PS score	0.80 (0.52)	0–3		13	1

* Worst possible value of the item or minimum total value of the scale. ** Best possible value of the item or maximum total value of the scale.

Table 3. Reproducibility of MDHAQ scores.

Score	N	First Measurement, mean (SD)	Change from First to Second Measurement, mean (95% CI)	Reproducibility	
				ICC (95% CI)	CR (95% CI)
Function (FN)	32	0.96 (0.74)	0.11 (0.02 to 0.21)	0.93 (0.82 to 0.97)	0.53 (0.42 to 0.70)
Psychological (PS)	32	0.77 (0.56)	0.01 (–0.09 to 0.12)	0.84 (0.70 to 0.92)	0.62 (0.50 to 0.83)

ICC: Intraclass correlation coefficient. CR: Coefficient of repeatability. Express the expected maximum size of 95% of the absolute differences between paired observations. 95% CI obtained by bias corrected and accelerated bootstrapping.

Table 4. Explanatory factor analysis with promax-rotated factor loadings of the Finn-MDHAQ function items. Coefficients with values < 0.50 not shown.

Item	Factor 1	Factor 2
a. Dressing yourself, including tying shoelaces and doing buttons?	0.88	
b. Get in or out of bed?	0.60	
c. Lift a full cup or glass to your mouth?	0.62	
d. Walk outdoors on flat ground?		0.61
e. Wash and dry your entire body?	0.82	
f. Bend down to pick up clothing from the floor?	0.82	
g. Turn regular faucets on and off?	0.81	
h. Get in or out of a car, bus, train, or airplane?		0.66
i. Walk 2 miles?		0.99
j. Participate in sports and games as you would like?		0.67

es, first for the initial 3-item scale and then for the 2-item scale and for the sleep item separately. The PS scale in both approaches revealed no underlying significant correlation with HAQ. The 2-item PS scale showed strong correlations with 2 physical subscales of Finn-AIMS2 (Mobility level and Household tasks) and 2 psychological subscales. The sleep item correlated with some physical subscales of Finn-AIMS2 and the subscale of Arthritis Pain. As shown in Table 6, FN had strong correlations with almost all clinical characteristics studied. In both PS scale relationship

approaches we found strong correlations with physician's and patient's overall assessment. The sleep item had a strong correlation with the patient's assessment of pain, but the 2-item PS scale did not show this relationship.

DISCUSSION

This study is the first translation and cross-cultural adaptation research carried out on the MDHAQ questionnaire. This instrument is not yet in widespread use but it unifies 3 important features of a rheumatologic assessment instru-

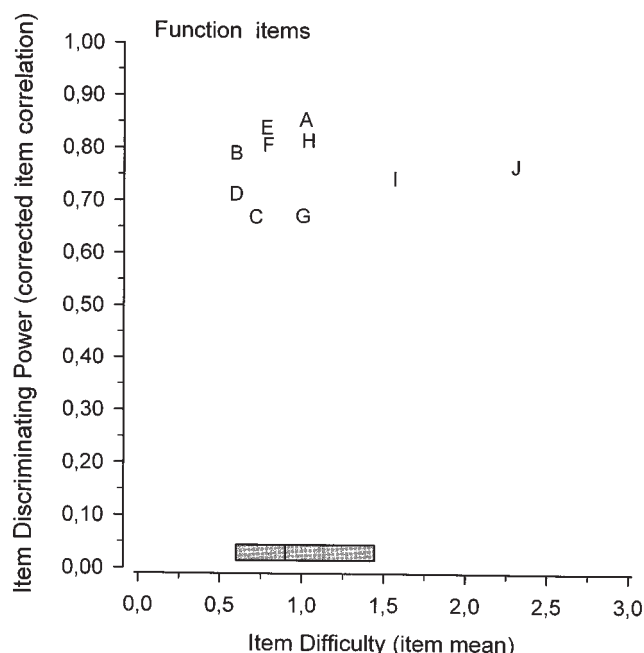


Figure 1. Item analysis for Function items. The bar denotes median and interquartile range of total FN score. Letters indicate corresponding items in the Function scale.

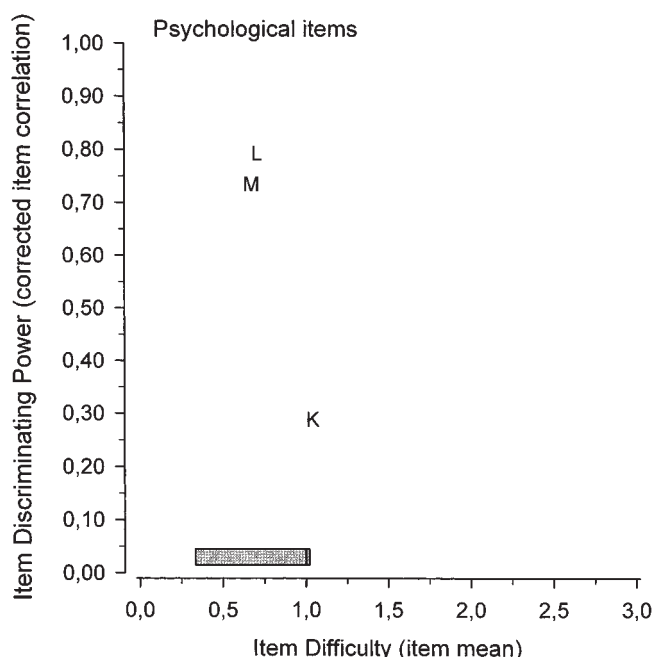


Figure 2. Item analysis for Psychological items. The bar denotes median and interquartile range of total PS score. Letters indicate corresponding items in the Psychological scale.

ment: the physical and mental aspects of the patient's functioning¹⁴, a patient oriented perspective¹⁵, and brief, patient-friendly format¹⁶.

The evaluation process was carried out on patients with a stable diagnosis of RA and representing a wide spectrum of disease to cover a wide range of RA patients, as the mean

duration of the disease may indicate. We found high response rates on all items of both Finn-MDHAQ scales, from which it may be concluded that the instrument was easy to complete. Previously, in the case of MHAQ^{17,18} and the initial MDHAQ¹, the floor and ceiling effect was discussed, but we did not find any signs of the floor effect in the Finn-MDHAQ, as the floor and also the ceiling values of the 2 total scales were low. Thus, the Finn-MDHAQ can detect the whole spectrum of physical and psychological aspects of RA patients' outcomes. However, 2 or 3 items as in the PS scale can only provide a useful screening tool for psychological distress, as previously stated¹. The reproducibility of Finn-MDHAQ can be considered good on both the FN and PS scales.

Internal consistency was good in the FN scale and moderate in the PS. This finding was supported by the results of item analysis. In the 3-item PS scale the corrected item correlation of the sleep question was clearly lower than the others. To ask the patient about his/her quality of sleep is clinically very important, but it does not here support the item structure in the PS scale and does not measure the same attribute as the other items. For future use of the MDHAQ it would be appropriate to consider leaving the sleep item as a separate part, as was done in the CLINHAQ¹⁹. Perhaps also altering the format of the question from the Likert scale to a VAS would give patients more freedom to rate their quality of sleep. In the item analysis of the FN there was one item with high item difficulty ("Participating in sports and games as you would like"). As this item does not indicate the games and sports the respondents would participate in, the question leaves the respondent a wide range of activities to choose from and, for example, the wish to play darts or badminton sets very different standards for one's mobility.

Two factors — mobility of upper extremities and trunk and mobility of lower extremities — detected on the FN scale were to be expected by reason of the physical character of the initial HAQ from which this scale is derived. Thus the FN scale covers a wide spectrum of mobility in the subject.

We compared the Finn-MDHAQ to another multidimensional questionnaire, Finn-AIMS2: the strongest correlations were between the FN scale and Finn-AIMS2 physical subscales and between the PS and Finn-AIMS2 psychological subscales, indicating convergent validity. As seen in Tables 5 and 6, sleep is strongly correlated with measures of pain. This result is in agreement with the findings of Houssien and coworkers²⁰. A subject of consideration was the lack of correlation between FN score and number of swollen joints (Table 6), and we found no explanation for this. In contrast to this result, Pincus, *et al*²¹ found strong correlation between total joint count (including joint swelling, joint tenderness, and joint limited scores) and MHAQ score.

We have translated and evaluated the Multidimensional

Table 5. Correlations between Finn-MDHAQ scores and HAQ and Finn-AIMS2 subscales.

	Multidimensional Health Assessment Questionnaire			
	Function FN score	PS score	Psychological PS Score, 2-item [†]	PS Score, Sleep ^{††}
HAQ	0.91***	0.25	0.18	0.28
AIMS2				
Mobility level	0.74***	0.43***	0.38***	0.32**
Walking and bending	0.73***	0.29*	0.17	0.31**
Hand and finger function	0.57***	0.26*	0.29*	0.14
Arm function	0.60***	0.32**	0.20	0.35**
Self-care	0.61***	0.29*	0.22	0.27*
Household tasks	0.57***	0.33**	0.37***	0.20
Social activities	0.19	0.21	0.23	0.10
Support from family and friends	0.06	0.21	0.26*	0.08
Arthritis pain	0.35**	0.40***	0.23	0.43***
Level of tension	0.33**	0.67***	0.69***	0.25
Mood	0.33**	0.59***	0.66***	0.19

[†] Only items "Deal with the feelings of anxiety or being nervous?" and "Deal with the feelings of depression or feeling blue?". ^{††} Only item "Get a good night's sleep?". * p < 0.05, ** p < 0.01, *** p < 0.001; Sidak-adjusted probabilities.

Table 6. Correlations between Finn-MDHAQ scores and demographic and clinical characteristics.

	Multidimensional Health Assessment Questionnaire			
	Function FN Score	PS Score	Psychological PS Score, 2-item [†]	PS Score, Sleep ^{††}
Age	0.32**	0.14	0.09	0.18
Duration of disease	0.21	0.15	0.05	0.21
ESR	0.18	0.01	0.02	0.01
No. of swollen joints	0.08	0.17	0.24	-0.04
No. of tender joints	0.31**	0.24	0.29*	0.09
Physician overall assessment	0.39***	0.35***	0.26*	0.30**
Patient overall assessment	0.48***	0.40***	0.32**	0.30**
Patient assessment of pain	0.40***	0.30**	0.20	0.28**
DAS28	0.28*	0.19	0.20	0.06
Functional class	0.49***	0.22	0.20	0.14

[†] Only items "Deal with the feelings of anxiety or being nervous?" and "Deal with the feelings of depression or feeling blue?". ^{††} Only item "Get a good night's sleep?". * p < 0.05, ** p < 0.01, *** p < 0.001; Sidak-adjusted probabilities.

Health Assessment Questionnaire, MDHAQ. The instrument proved to be applicable, reliable, and also valid for the part of the FN scale measuring the functional ability of Finnish rheumatic patients. However, the incongruity in the PS scale structure produced moderate internal consistency in this scale. With the minor modifications suggested, this weakness can be overcome.

ACKNOWLEDGMENT

We thank Niina Helminen, Leena Leppänen, Merja Hakala, and Varpu Järveläinen for kindly helping with patient recruitment, and Drs. Mikko Vasala, Kai Immonen, Markku Kauppi, and Jorma Vuotila for effective cooperation.

REFERENCES

- Pincus T, Swearingen C, Wolfe F. Toward a multidimensional assessment questionnaire (MDHAQ). *Arthritis Rheum* 1999;42:2220-30.
- Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. *Arthritis Rheum* 1980;23:137-45.
- Pincus T, Summey JA, Soraci SA, et al. Assessment of patient satisfaction in activities of daily living using a modified Stanford Health Assessment Questionnaire. *Arthritis Rheum* 1983;26:1346-53.
- Bruce B, Fries J. The Stanford Health Assessment Questionnaire: A review of its history, issues, progress, and documentation. *J Rheumatol* 2003;30:167-78.
- Felson DT, Anderson JJ, Boers M, et al. The American College of Rheumatology preliminary core set of disease activity measures for rheumatoid arthritis clinical trials. *Arthritis Rheum* 1993;36:729-40.
- Hakala M, Nieminen P, Manelius J. Joint impairment is strongly correlated with disability measured by self-report questionnaires. Functional status assessment of individuals with rheumatoid arthritis in a population based series. *J Rheumatol* 1993;21:64-9.
- Arnett FC, Edworthy SM, Bloch DA, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988;31:315-24.
- Internet. Available from: <http://www.mdhaq.com>. Accessed April 14, 2005.
- 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;46:1417-32.

10. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine* 2000;25:3186-91.
11. Hakkinen A, Kautiainen H, Hannonen P, et al. Pain and joint mobility explain individual subdimensions of the Health Assessment Questionnaire (HAQ) disability index in patients with rheumatoid arthritis. *Ann Rheum Dis* 2005;64:59-63. Epub 2004 May 6.
12. Arkela-Kautiainen M, Kauppi M, Heikkilä S, et al. Evaluation of the Arthritis Impact Measurement Scales (AIMS2) in Finnish patients with rheumatoid arthritis. *Scand J Rheumatol* 2003;32:300-5.
13. Steinbocker O, Traeger CH, Batterman RC. Therapeutic criteria in rheumatoid arthritis. *JAMA* 1949;140:659-62.
14. Piro MH, Kwok CK. Update on measurement of relevant outcomes in rheumatology. *Curr Opin Rheumatol* 1996;8:101-5.
15. Kwok CK, Ibrahim SA. Rheumatology patient and physician concordance with respect to important health and symptom status outcomes. *Arthritis Rheum* 2001;45:372-7.
16. Katz JN, Larson MG, Phillips CB, Fossel AH, Liang MH. Comparative measurement sensitivity of short and longer health status instruments. *Med Care* 1992;30:917-25.
17. Stucki G, Stucki S, Brühlmann P, Michel BA. Ceiling effects of the Health Assessment Questionnaire and its modified version in some ambulatory rheumatoid arthritis patients. *Ann Rheum Dis* 1995;54:461-5.
18. Serrano MAB, Fabregat JB, Garzón JO. Should the MHAQ ever be used? *Ann Rheum Dis* 1996;55:271-2.
19. Wolfe F. A brief health status instrument: CLINHAQ [abstract]. *Arthritis Rheum* 1989;32 Suppl:S99.
20. Houssien DA, McKenna SP, Scott DL. The Nottingham Health Profile as a measure of disease activity and outcome in rheumatoid arthritis. *Br J Rheumatol* 1997;36:69-73.
21. Pincus T, Callahan LF, Brooks RH, Fuchs HA, Olsen NJ, Kaye JJ. Self-report questionnaire scores in rheumatoid arthritis compared with traditional physical, radiographic, and laboratory measures. *Ann Intern Med* 1989;110:259-66.