

Measure of Function in Rheumatoid Arthritis: Individualized or Classical Scales?

Raphaèle Seror ¹, Florence Tubach ¹, Gabriel Baron ¹, Francis Guillemin ², and Philippe Ravaud ¹

(1) AP-HP, Hôpital Bichat, Département d'Epidémiologie, Biostatistique et Recherche Clinique, Paris, France; INSERM U738, Paris, France; Université Paris 7-Denis Diderot, UFR de médecine, Paris, France

(2) INSERM, Centre d'Epidémiologie Clinique CIE6, Nancy ; CHU de Nancy, Service d'épidémiologie et évaluation cliniques, Nancy ; Nancy-Université, EA 4003, Vandoeuvre-les-Nancy, France

Correspondence and reprint requests to:

Dr Raphaèle Seror

Département d'Epidémiologie, Biostatistique et Recherche Clinique, INSERM U738, Hôpital Bichat, 46 rue Henri Huchard, 75018 Paris, France.

Email : raphaele.se@gmail.com

KEY WORDS: Health assessment questionnaire, Rheumatoid arthritis, Patient-specific index, Individualization, Patient reported outcome, Patient's perspective.

WORD COUNT: 1999

ABSTRACT (250 words)

Objective. The Health Assessment Questionnaire Disability Index (HAQ-DI) is the most widely used measure of function in rheumatoid arthritis (RA). To enhance the incorporation of patients' view in outcome assessment, this study aimed to evaluate individualized forms of the HAQ-DI.

Patients and methods. HAQ-DI data were prospectively obtained from 370 RA outpatients treated with leflunomide over a 6-month period. At baseline and final visits, patients had to rate the importance they attached to each activity addressed by the 20 HAQ-DI items, and to select the 5 activities they considered the most important. Different individualized scales were evaluated: scales preserving all domains, in which the score for each item is multiplied by or added to its importance; and scales involving for each patient only the 5 most important items. The psychometric properties of these scales were compared to those of the HAQ-DI.

Results. For each HAQ-DI item, severity and importance scores were weakly correlated. Scores for all individualized scales were highly correlated with the HAQ-DI score ($\rho > 0.75$). All scales had a good internal consistency (Cronbach's alpha: 0.87 to 0.88). Compared to the HAQ-DI, individualized scales did not have better sensitivity to change (standardized response mean: 0.64 to 0.69 vs. 0.74).

Conclusion. Individualized scales have similar properties as the HAQ-DI. However, individualized questionnaires measuring importance gave complementary information to the measure of disability. Therefore, even if individualization is probably not needed for group assessment in all randomized controlled trials, the use of individualized questionnaires could be clinically relevant for individual RA patients.

Outcome assessment in rheumatoid arthritis (RA) was primarily based on physician-reported outcomes such as measures of disease activity [1-3] or structural progression [4, 5]. However, symptoms such as pain, fatigue and disability are relevant and their assessment is recommended for evaluation of response to treatment.[6] The Health Assessment Questionnaire Disability Index (HAQ-DI) is a valid, reliable, and widely used tool to measure functional status in RA patients.[7-10] It is a standardized questionnaire, identical for all patients, investigating patients' ability to perform some activities, as well as the importance of being able to perform each activity. However, these issues may vary widely among individuals. Furthermore, enhancing patients' perspectives in RA outcome assessment was noted as a major focus for research.[6, 11]

For these reasons, patient-specific, or individualized, scales have been developed to address each patient's priorities in outcome assessment in RA.[12-15] Some have been shown to be more sensitive to change than classical instruments.[16-18] However, most of them involve semi-structured interviews requiring 15 to 40 minutes to complete.[19] Hewlett et al. developed an individualized questionnaire derived from the modified HAQ, [20, 21] where each item is weighted by the importance it has for each patient. However, responsiveness of the scale has not been evaluated.

We designed this prospective study to develop, validate and evaluate psychometric properties of patient-specific scales, assessing the level of disability in RA patients, derived from the HAQ-DI according to several methods of individualization.

PATIENTS AND METHODS

Study population

Data were obtained from a 6-month, prospective, open-label study involving 378 RA outpatients treated with leflunomide.[22] Patients included had RA as defined by the American Rheumatism Association revised criteria,[23] with active disease defined by a physician global assessment of disease activity ≥ 2 on a 5-point Likert scale (from 1 to 5). Two visits were scheduled, one at baseline before initiation of leflunomide and one 6 months later.

Measurement

At baseline and final visits, patients completed:

- The French version of the HAQ-DI,[24] termed the *severity questionnaire*. It comprises 20 items, referring to ability in performing daily life activities, involving 8 domains and includes a "help and device" component. For each item, score ranged from 0 to 3. The score of each domain is the highest item score of that domain, unless help or a device is required. Dependence on help and device increases a lower score to 2. The overall score is the mean of the 8 domain scores and varies from 0 to 3.

- The *importance questionnaire*: patients had to rate the importance of each HAQ-DI item, by answering the question "In your daily life, the following activity is...". Importance was rated by a 4-point Likert scale, ranging from 0="not important" to 3="extremely important".

- The *preference questionnaire*: patients had to select and order the 5 activities (i.e. HAQ-DI items) they considered the most important in their daily lives.

The swollen and tender joint counts, erythrocyte sedimentation rate (ESR), level of C-reactive protein (CRP) were also collected to obtain the disease activity score (DAS28).

Methods of individualization

Several individualized scales were obtained:

(1) Scales based on the importance questionnaire

- With multiplicative, or additive, methods: for each item, the severity score was multiplied by, or added to, the importance score.

The score for each domain was the highest score obtained after combining the severity and importance scores of items of that domain. The total score was the mean score of all 8 domains. Individualized scale, with multiplicative method, preserving all 20 items was also obtained.[25]

(2) Scales based on the preference questionnaire.

- (a) Top 5 HAQ: involving only the 5 most important activities for each patient.
- (b) Weighted top 5 HAQ: involving only the 5 most important activities, whereby the severity score is weighted by the rank of preference for this activity (highest weight referred to the most important activity).

In these scales, the resulting items are not the same for all patients. The total score was the sum of the scores, weighted or not, of the 5 most important items.

Initial scores were obtained by combining data from the baseline visit. Final scores were obtained by combining baseline importance or preference ratings and final severity scores. Because the rating of importance has not been applied to the help and device component, individualized scales were developed without this component. Each scale was linearly transformed to a scale of 0 to 3; 0 indicating no disability and 3 indicating maximal possible disability.

Statistical analysis

All patients who had completed the original HAQ-DI with no missing data at baseline were involved in the development of individualized forms. Psychometric properties of each individualized scale were evaluated, and compared to that of the HAQ-DI:

- Construct validity was assessed by Spearman rho correlation between the score of each individualized scale and that of the HAQ-DI. We also examined divergent validity using Spearman correlation between scores of individualized scales and other measures of disease activity (swollen and tender joint count, ESR, CRP level and DAS28).
- Internal consistency was assessed, when estimable (i.e. for scales with identical items), by Cronbach's alpha coefficient.[26] Estimation of confidence intervals and comparisons of Cronbach's alpha coefficients involved the bootstrap method, with 1000 replications.[27]

- Responsiveness was assessed by the standardized response mean (SRM), which is the mean change in score between baseline and final visit divided by the standard deviation of the change in score.[28] SRM values can be considered large (>0.8), moderate (0.5-0.8) or small (<0.5).[29, 30] SRM confidence interval estimations and SRM comparisons involved the bootstrap method with 1000 replications.[27]

Statistical analysis involved use of SAS, release 9.1 (SAS Inst., Cary, NC), and R, release 2.2.1 (The R Foundation for Statistical Computing, Vienna, Austria).

RESULTS

Of the 378 outpatients with RA enrolled in the study, 370 (98%) completed the HAQ-DI without missing data at baseline (table 1). Among them, 360 (95%) and 281 (74%) patients correctly completed the importance and preference questionnaires, respectively.

For each item, severity and importance scores were weakly or not significantly correlated (rho ranging from 0.01 to 0.29) (table 2). Plots of severity by importance scores revealed that patients with the highest severity scores for items did not necessarily grade the items as important and patients with the lowest severity score for items did not necessarily consider the items as not important (figure 1). As for item “get on and off the toilet”, 71% (124/175) of the not disabled patients (severity score of 0) considered this activity as important or extremely important. At the baseline, based on the preference questionnaire, the 3 most important items were “dress yourself” (86%), “wash and dry entire body” (63%), and “walk outdoors on flat ground” (55%). Data collected by the importance questionnaire also showed that each of these 3 activities was considered important or extremely important by more than 75% of the patients. At final visit, no change in selection of the 5 most important items was observed for 17% of patients, whereas 34% (61/182) and 66% (121/182) of patients selected at least 4 and 3, respectively, of the initial 5 items they chose.

Construct validity

All individualized scale scores were highly correlated with the HAQ-DI score (rho \geq 0.75; table 3). Lower correlations were observed with measures of disease activity: tender and swollen joint counts (rho ranging from 0.21 to 0.39) and DAS28 (rho ranging from 0.38 to 0.47). The lowest correlations were observed with biological features of disease activity, such as ESR and CRP level (rho ranging from 0.10 to 0.18).

Internal consistency

For scales with 8 domains, Cronbach’s alpha coefficient ranged from 0.87 to 0.88 (table 3). Scales preserving all 20 items had higher Cronbach’s alpha coefficients (alpha=0.94).

Responsiveness

The SRM of the HAQ-DI was 0.74, compared to 0.69 without help and device component (p $<$ 0.05). SRMs of individualized scales ranged from 0.64 to 0.69 (table 3). SRM of scales preserving all 20 items was 0.66.

Table 1. Baseline characteristics of patients

	<i>Patients with complete data for HAQ-DI N=370</i>	<i>Patients with complete data for all scores N=281</i>
Age (years)	58.37 ± 11.93	58.13 ± 11.87
Sex (female)	292 (78.92%)	225 (80.07%)
Rheumatoid factor positive	248 (66.85%)	174 (61.92%)
Disease duration (years)	9.84 ± 8.49	9.34 ± 8.25
Prior DMARD intake	325 (87.84%)	243 (86.48%)
Patient global assessment of disease activity	52.41 ± 19.08	52.82 ± 19.53
Physician global assessment of disease activity	51.58 ± 17.11	51.63 ± 17.12
Swollen joint count (/28)	9.78 ± 5.70	9.51 ± 5.76
Tender joint count (/28)	12.02 ± 6.58	11.98 ± 6.66
ESR (mm)	30.14 ± 22.25	30.46 ± 21.81
CRP level (mg/L)	23.43 ± 30.80	23.64 ± 31.03
DAS28	5.62 ± 1.10	5.62 ± 1.11
HAQ-DI	1.40 ± 0.62	1.41 ± 0.61
	Range [0.00–2.87]	Range [0.12–2.75]

Data are presented as mean ± SD or n (%)

DAS: disease activity score; DMARD: disease-modifying anti-rheumatic drug; ESR: erythrocyte sedimentation rate; HAQ-DI: health assessment questionnaire disability index; CRP: C-reactive protein

Table 2. Importance of each activity of the HAQ-DI from the patient's perspective

Domain	Item	Severity Questionnaire	Importance Questionnaire	Correlation between severity and importance	Preference Questionnaire		
		Mean baseline disability (n=370)	Mean baseline importance (n=360)		Items selected at baseline visit (n=274)	Items selected at both baseline and final visits (n=182)	Identical items selected at baseline and at final visit (n=182)
Dressing and grooming	Dress yourself	1.07 ± 0.76	2.14 ± 0.80	0.15	234 (86%)	124 (68%)	142 (78%)
	Shampoo your hair	0.95 ± 0.89	1.89 ± 0.91	0.24	37 (14%)	7 (4%)	156 (86%)
Arising	Stand from a chair	0.68 ± 0.67	1.95 ± 0.90	0.11	73 (27%)	26 (14%)	143 (79%)
	Get in and out of bed	0.74 ± 0.71	2.01 ± 0.90	0.10	74 (27%)	24 (13%)	140 (77%)
Eating	Cut your meat	0.95 ± 0.84	1.95 ± 0.88	0.22	86 (31%)	26 (14%)	132 (73%)
	Lift a full cup or glass to mouth	0.60 ± 0.72	1.92 ± 0.92	0.14	57 (21%)	16 (9%)	145 (80%)
	Open a new carton of milk	1.41 ± 0.93	1.83 ± 0.87	0.20	16 (6%)	3 (2%)	167 (92%)
Walking	Walk outdoors on flat ground	0.74 ± 0.69	2.01 ± 0.93	0.22	150 (55%)	66 (36%)	123 (68%)
	Climb 5 steps	0.84 ± 0.75	2.04 ± 0.88	0.17	34 (12%)	6 (3%)	154 (85%)
Hygiene	Wash and dry entire body	0.85 ± 0.77	2.18 ± 0.91	0.14	172 (63%)	97 (53%)	135 (74%)
	Take a bath	1.16 ± 1.05	1.87 ± 1.01	0.01	43 (16%)	13 (7%)	148 (81%)
	Get on and off the toilet	0.59 ± 0.67	2.08 ± 0.99	0.07	75 (27%)	32 (18%)	133 (73%)
Reach	Reach and get down a 5-lb object	1.68 ± 0.97	1.77 ± 0.88	0.13	45 (16%)	13 (7%)	148 (81%)
	Bend down and pick up clothing	0.82 ± 0.76	1.79 ± 0.89	0.23	18 (7%)	6 (3%)	166 (91%)
Grip	Open car doors	0.75 ± 0.74	1.70 ± 0.89	0.17	24 (9%)	4 (2%)	165 (91%)
	Open jars (previously opened)	1.13 ± 0.87	1.82 ± 0.83	0.25	26 (9%)	8 (4%)	165 (91%)
	Turn taps on and off	0.94 ± 0.75	2.04 ± 0.85	0.29	34 (12%)	8 (4%)	155 (85%)
Other activities	Run errands and shop	0.86 ± 0.85	1.87 ± 0.91	0.14	89 (32%)	36 (20%)	114 (63%)
	Get in and out of car	0.73 ± 0.69	1.83 ± 0.92	0.21	31 (11%)	4 (2%)	158 (87%)
	Do chores	1.31 ± 0.88	1.87 ± 0.89	0.09	87 (32%)	41 (23%)	151 (83%)

Data are presented as mean ± SD or n (%)

Table 3. Psychometric properties of classical HAQ-DI and individualized HAQ scales

	<i>HAQ-DI</i>	Importance questionnaire		Preference questionnaire	
		<i>Individualized HAQ multiplicative</i>	<i>Individualized HAQ additive</i>	<i>5-item HAQ</i>	<i>Weighted 5-item HAQ</i>
Mean baseline score¹	1.40 ± 0.62	0.92 ± 0.55	1.70 ± 0.46	1.01 ± 0.64	1.03 ± 0.66
Mean change in score¹	-0.41 ± 0.55	-0.30 ± 0.44	-0.19 ± 0.28	-0.38 ± 0.59	-0.39 ± 0.61
Construct validity					
Spearman's correlation with HAQ-DI score	1	0.88	0.77	0.77	0.76
Internal consistency					
Cronbach's alpha	0.87	0.88	0.88	NE	NE
[95% CI]	[0.85–0.89]	[0.85–0.90]	[0.86–0.90]		
Sensitivity to change					
SRM	0.74	0.69	0.68	0.65	0.64
[95% CI]	[0.64–0.86]	[0.58–0.79]	[0.58–0.80]	[0.55–0.77]	[0.54–0.76]

¹Data are presented as mean ± SD.

HAQ-DI: health assessment questionnaire disability index; NE: not estimable; SRM: standardized response mean.

Each score was linearly transformed to a score from 0 to 3.

Mean changes in scores were calculated between the baseline visit at week 0 (W0) and the final visit at week 24 (W24): W24-W0.

For Spearman's correlations, all p-values were <0.0001.

[95% CI]: 95% confidence intervals were estimated by the bootstrap method with 1000 replications.

DISCUSSION

This study evaluated the interest of individualizing the HAQ-DI to measure disability in RA patients. Individualized scales have similar psychometric properties as the parent HAQ-DI, whatever the individualization method used. However, adding a measure of “importance” to a measure of “severity” for each activity provided complementary information that could be clinically useful in decision making for individual patients.

Many studies have shown divergent views of a patient’s health between patients and clinicians.[31-33] So far, the use of patient-reported outcomes and inclusion of patients’ opinions in their development has gradually increased.[34] One usual approach to incorporate patients’ opinion in development of patient-reported outcomes is to determine, in a group of patient, the importance of each item and to weight that item with the mean importance score at group level. However, such a group approach assumes concordance of all patients’ views. To facilitate patient input into the assessment process, individualized measured has been proposed. Two key practices are recommended for individualization :[35] (i) to incorporate patients’ rating of importance in addition to severity for each item, and/or (ii) to augment the instrument with supplemental items that patients can add. Our study relies on the first recommendation, and used 2 methods. The first method combined, for each item, the rating of “severity” and “importance”. In the second method, each patient selected a limited number of items he considered the most important. Moreover, both methods were combined, by ranking the importance of these most important items. Regarding the second recommendation, our study unfortunately, did not offer patients the opportunity to provide supplemental relevant items, not contained in the HAQ-DI.[31]

Arguing that importance of an activity is closely related to level of disability in performing that activity, some authors wonder whether assessment of disability alone could be sufficient.[36, 37] Against this hypothesis, our resulted showed that for most items, rating of importance and level of disability were weakly correlated. Moreover, as previously shown, ratings of the importance were extremely variable among patients with similar levels of disability, but were also similar among patients with diverse levels of disability. [20, 38] So, individualized approach may be useful in clinical practice to help physician to prioritize care on what is really meaningful to each patient.

In our previous study evaluating individualized forms of the WOMAC function subscale, ratings of the importance and severity of each item were strongly correlated;[38] furthermore, the individualized scale retaining the 5 most important activities was more sensitive to change than the parent scale. In that study, patients were asked to define activities that were the most in need of improvement. In the present study, the wording was slightly different, and patients had to define importance of activities of their daily lives, even though they had no difficulty performing them. As shown for item “get on and of the toilet”, with this wording, it is likely that important items for a patient may be those that are relevant to him, but perhaps easier to do. Therefore, if these items are initially close to normal, improvement is not possible. Items are therefore less sensitive to change. This difference might explain the discrepancies between results of these studies.

As previously proposed,[25, 39] we also used the HAQ-DI without the help and devices component, but found it less sensitive to change than the parent HAQ-DI. Because the rating of importance was not applied to the help and devices component, individualized scales

did not include this component. This omission could have explained in part why the individualized scales had lower responsiveness than the parent HAQ-DI.

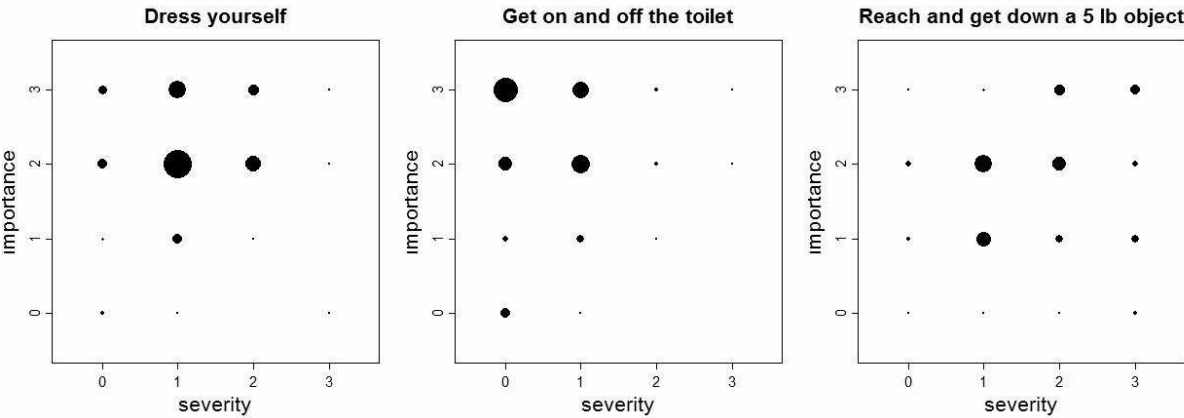
Individualized scales derived from the HAQ-DI did not have better responsiveness than the parent HAQ-DI. This could be explained by the wording of the individualized questionnaire that did not focus on activities in need of improvement, and by the scoring method of the HAQ-DI, somewhat already individualized considering for each domain the activity for which the patient is most disabled. However, the measure of importance of an activity gave complementary information to the measure of disability. Therefore, even if individualization is probably not needed for group assessment in all randomized controlled trials, the use of individualized questionnaires could be clinically relevant in decision making for individual patients.

Acknowledgements: This study was funded Aventis Pharma, Paris, France. We thank all the physicians who participated to this study.

Competing interests: None.

The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, an exclusive licence (or non-exclusive for government employees) on a worldwide basis to the BMJ Publishing Group Ltd and its Licensees to permit this article to be published in Annals of the Rheumatic Diseases editions and any other BMJ PGL products to exploit all subsidiary rights, as set out in our licence <http://ard.bmjournals.com/ifora/licence.pdf>

Figure 1. Patterns of plots of scores of severity by importance for 3 of the 20 HAQ-DI items.



These graphics represent patterns of plots of severity scores against importance scores. The area of each point is proportional to the number of patients. Plots showed that patients with the highest severity scores for items did not necessarily grade the items as important (item 13: “Reach and get down a 5 lb object”) and patients with the lowest severity score for items did not necessarily consider the items as not important (item 1 and 12: “Dress yourself” and “Get on and off the toilet”).

References

- 1 Felson DT, Anderson JJ, Boers M, Bombardier C, Furst D, Goldsmith C, *et al.* American College of Rheumatology. Preliminary definition of improvement in rheumatoid arthritis. *Arthritis Rheum* 1995;**38**:727-35.
- 2 van Gestel AM, Prevoo ML, van 't Hof MA, van Rijswijk MH, van de Putte LB, van Riel PL. Development and validation of the European League Against Rheumatism response criteria for rheumatoid arthritis. Comparison with the preliminary American College of Rheumatology and the World Health Organization/International League Against Rheumatism Criteria. *Arthritis Rheum* 1996;**39**:34-40.
- 3 Prevoo ML, van 't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL. Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis Rheum* 1995;**38**:44-8.
- 4 Sharp JT. Radiologic assessment as an outcome measure in rheumatoid arthritis. *Arthritis Rheum* 1989;**32**:221-9.
- 5 van der Heijde DM. Joint erosions and patients with early rheumatoid arthritis. *Br J Rheumatol* 1995;**34 Suppl 2**:74-8.
- 6 Aletaha D, Landewe R, Karonitsch T, Bathon J, Boers M, Bombardier C, *et al.* Reporting disease activity in clinical trials of patients with rheumatoid arthritis: EULAR/ACR collaborative recommendations. *Ann Rheum Dis* 2008;**67**:1360-4.
- 7 Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. *Arthritis Rheum* 1980;**23**:137-45.
- 8 Fries JF, Spitz PW, Young DY. The dimensions of health outcomes: the health assessment questionnaire, disability and pain scales. *J Rheumatol* 1982;**9**:789-93.
- 9 Bruce B, Fries JF. The Stanford Health Assessment Questionnaire: a review of its history, issues, progress, and documentation. *J Rheumatol* 2003;**30**:167-78.
- 10 Ramey D, Fries JF, Sing G. The Health Assessment Questionnaire - Status and review. In: In: Spikler B e, editor. *Quality of life and pharmacoeconomics in clinical trials. 2nd ed.* Philadelphia: Lippincott-Raven, 1996:227-37.
- 11 Kirwan J, Heiberg T, Hewlett S, Hughes R, Kvien T, Ahlmen M, *et al.* Outcomes from the Patient Perspective Workshop at OMERACT 6. *J Rheumatol* 2003;**30**:868-72.
- 12 Tugwell P, Bombardier C, Buchanan WW, Goldsmith CH, Grace E, Hanna B. The MACTAR Patient Preference Disability Questionnaire--an individualized functional priority approach for assessing improvement in physical disability in clinical trials in rheumatoid arthritis. *J Rheumatol* 1987;**14**:446-51.
- 13 Bell MJ, Bombardier C, Tugwell P. Measurement of functional status, quality of life, and utility in rheumatoid arthritis. *Arthritis Rheum* 1990;**33**:591-601.
- 14 Meenan RF, Gertman PM, Mason JH, Dunaif R. The arthritis impact measurement scales. Further investigations of a health status measure. *Arthritis Rheum* 1982;**25**:1048-53.
- 15 Ripat J, Etcheverry E, Cooper J, Tate RB. A comparison of the Canadian Occupational Performance Measure and the Health Assessment Questionnaire. *Can J Occup Ther* 2001;**68**:247-53.
- 16 Tugwell P, Wells G, Strand V, Maetzel A, Bombardier C, Crawford B, *et al.* Clinical improvement as reflected in measures of function and health-related quality of life following treatment with leflunomide compared with methotrexate in patients with rheumatoid arthritis: sensitivity and relative efficiency to detect a treatment effect in a twelve-month, placebo-controlled trial. Leflunomide Rheumatoid Arthritis Investigators Group. *Arthritis Rheum* 2000;**43**:506-14.
- 17 Verhoeven AC, Boers M, van der Liden S. Validity of the MACTAR questionnaire as a functional index in a rheumatoid arthritis clinical trial. The McMaster Toronto Arthritis. *J Rheumatol* 2000;**27**:2801-9.
- 18 Tugwell P, Bombardier C, Buchanan WW, Goldsmith C, Grace E, Bennett KJ, *et al.* Methotrexate in rheumatoid arthritis. Impact on quality of life assessed by traditional standard-item and individualized patient preference health status questionnaires. *Arch Intern Med* 1990;**150**:59-62.
- 19 Jolles BM, Buchbinder R, Beaton DE. A study compared nine patient-specific indices for musculoskeletal disorders. *J Clin Epidemiol* 2005;**58**:791-801.
- 20 Hewlett S, Smith AP, Kirwan JR. Measuring the meaning of disability in rheumatoid arthritis: the Personal Impact Health Assessment Questionnaire (PI HAQ). *Ann Rheum Dis* 2002;**61**:986-93.
- 21 Pincus T, Summey JA, Soraci SA, Jr., Wallston KA, Hummon NP. Assessment of patient satisfaction in activities of daily living using a modified Stanford Health Assessment Questionnaire. *Arthritis Rheum* 1983;**26**:1346-53.
- 22 Nguyen M, Kabir M, Ravaud P. Short-term efficacy and safety of leflunomide in the treatment of active rheumatoid arthritis in everyday clinical use : open-label, prospective study. *Clin Drug Investig* 2004;**24**:103-12.
- 23 Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, *et al.* The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988;**31**:315-24.

- 24 Guillemin F, Braincon S, Poure J. [Measurement of the functional capacity in rheumatoid polyarthritis: a French adaptation of the Health Assessment Questionnaire (HAQ)]. *Rev Rhum Mal Osteoartic* 1991;**58**:459-65.
- 25 Bruce B, Fries JF. The Health Assessment Questionnaire (HAQ). *Clin Exp Rheumatol* 2005;**23**:S14-8.
- 26 Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;**16**:297-334.
- 27 Efron B, Tibshirani RJ. An Introduction to the Bootstrap. New York: Chapman and Hall, 1993.
- 28 Liang MH, Fossel AH, Larson MG. Comparisons of five health status instruments for orthopedic evaluation. *Med Care* 1990;**28**:632-42.
- 29 Husted JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness: a critical review and recommendations. *J Clin Epidemiol* 2000;**53**:459-68.
- 30 Streiner DL, Norman GR. Health Measurement Scales: A Practical Guide to Their Development and Use. 2nd ed ed. New York: Oxford University Press, 1996.
- 31 Hewlett S, Smith AP, Kirwan JR. Values for function in rheumatoid arthritis: patients, professionals, and public. *Ann Rheum Dis* 2001;**60**:928-33.
- 32 Kwok CK, O'Connor GT, Regan-Smith MG, Olmstead EM, Brown LA, Burnett JB, *et al*. Concordance between clinician and patient assessment of physical and mental health status. *J Rheumatol* 1992;**19**:1031-7.
- 33 Berkanovic E, Hurwicz ML, Lachenbruch PA. Concordant and discrepant views of patients' physical functioning. *Arthritis Care Res* 1995;**8**:94-101.
- 34 Gossec L, Dougados M, Rincheval N, Balanescu A, Boumpas DT, Canadelo S, *et al*. The elaboration of the preliminary Rheumatoid Arthritis Impact of Disease (RAID) score: a EULAR initiative. *Ann Rheum Dis* 2008.
- 35 Gill TM, Feinstein AR. A critical appraisal of the quality of quality-of-life measurements. *Jama* 1994;**272**:619-26.
- 36 Kvien TK, Heiberg T. Patient perspective in outcome assessments--perceptions or something more? *J Rheumatol* 2003;**30**:873-6.
- 37 Heiberg T, Kvien TK. Preferences for improved health examined in 1,024 patients with rheumatoid arthritis: pain has highest priority. *Arthritis Rheum* 2002;**47**:391-7.
- 38 Seror R, Tubach F, Baron G, Falissard B, Logeart I, Dougados M, *et al*. Individualising the Western Ontario and McMaster Universities osteoarthritis index (WOMAC) function subscale: incorporating patient priorities for improvement to measure functional impairment in hip or knee osteoarthritis. *Ann Rheum Dis* 2008;**67**:494-9.
- 39 Wolfe F. Which HAQ is best? A comparison of the HAQ, MHAQ and RA-HAQ, a difficult 8 item HAQ (DHAQ), and a rescored 20 item HAQ (HAQ20): analyses in 2,491 rheumatoid arthritis patients following leflunomide initiation. *J Rheumatol* 2001;**28**:982-9.