

Determining Rheumatologists' Accuracy at Assessing Functional Disability in Rheumatoid Arthritis Patients Using the Health Assessment Questionnaire–Disability Index

JOHN D. CARTER, ABDUL B. LODHI, SONIA R. NAGDA, LOUIS RICCA, COLLEEN WARD, ERICA TRAINA, ZACHARY J. THOMPSON, YANGXIN HUANG, JOANNE VALERIANO, and FRANK B. VASEY

ABSTRACT. *Objective.* To test rheumatologists' accuracy in determining functional disability of their patients with rheumatoid arthritis (RA).

Methods. We used the Health Assessment Questionnaire–Disability Index (HAQ-DI) as our guide at assessing functional disability in patients with RA. Included were male and female patients, 18 to 65 years of age, diagnosed with RA. Demographic data collected included the patients' age, disease duration, rheumatoid factor (RF) status, presence of rheumatoid nodules, absence or presence of erosive disease, and class and stage of their disease. The primary endpoint was the mean difference in the patients' HAQ-DI scores versus that of the physicians' (mHAQ-Diff). Secondary endpoints were the mean difference in pain assessment scale (mPAS-Diff) score; and assessing to see if the physicians' HAQ-DI was altered by the patients' age, disease duration, RF status, rheumatoid nodules, absence or presence of erosive disease, and class or stage of the patient's RA.

Results. A total of 223 patients (139 female, 84 male) were evaluated. The mHAQ-Diff score was statistically significant at -0.3 ($p = 0.03$) with the rheumatologists more often overestimating the degree of functional disability in their RA patients. The mPAS-Diff score was 0.16 , but this was not significant ($p = 0.53$). There was no significant difference between the scores based on sex, presence or absence of RF, erosions, or rheumatoid nodules. However, the rheumatologists' estimated HAQ and PAS did seem to be more accurate in patients with lower class and stage of their RA.

Conclusion. Our results indicate that there is a clear difference between patients' and rheumatologists' assessment of patients' functional disability in RA, with the rheumatologists significantly overestimating the degree of this disability. Although the rheumatologists' accuracy at determining the amount of their patients' functional disability was poor overall, they were somewhat more accurate in patients with RA having less severe disease. (First Release April 1 2007; *J Rheumatol* 2007;34:958–63)

Key Indexing Terms:

RHEUMATOID ARTHRITIS

FUNCTIONAL ABILITY

HEALTH ASSESSMENT QUESTIONNAIRE-DISABILITY INDEX

From the Department of Internal Medicine, Division of Rheumatology, and Department of Biostatistics, University of South Florida; and James A. Haley Veterans' Hospital, Tampa, Florida, USA.

Dr. Carter and Dr. Lodhi contributed equally to this work.

J.D. Carter, MD, Assistant Professor of Medicine; E. Traina, Medical Student; J. Valeriano, MD, Associate Professor of Medicine; F.B. Vasey, MD, Professor of Medicine, Division of Rheumatology; Z.J. Thompson, PhD Student; Y. Huang, PhD, Department of Biostatistics, University of South Florida; L. Ricca, MD, Rheumatologist; C. Ward, DO, Rheumatologist, James A. Haley Veterans' Hospital, Tampa; S. Nagda, Medical Student, University of Florida, Gainesville, Florida; A. Lodhi, MD, Rheumatologist, Mississippi, USA.

Address reprint requests to Dr. J.D. Carter, 12901 Bruce B. Downs Blvd., MDC 81, Tampa, FL 33612, USA. E-mail: jocarter@hsc.usf.edu

Accepted for publication January 12, 2007.

Loss of functional capacity in rheumatoid arthritis (RA) is very common. RA affects about 2.1 million Americans, or 1% of the general population¹. Assessment of functional disability is very important in management of patients with RA. This is not only true when assessing the effectiveness of therapeutic interventions, but it is also important that patients' functional disability is adequately quantified with regard to lifestyle modifications and determining their employment capabilities. Work disability occurs in 85% of patients with RA under the age of 65 who had been working fulltime at disease onset². Determination of RA patients' amount of functional disability often hinges upon their rheumatologist's

assessment. The accuracy of rheumatologists' assessment of functional disability in patients with RA, however, has never been formally evaluated.

It is important for physicians, especially rheumatologists, to be objective and accurate when assessing functional disability in RA. There is no standard designed for physicians in this regard and there is little information about the adequacy of a physician's evaluation in the literature. Physicians rely upon their history-taking and examination skills, but these may not be sufficient. Loss of functional capacity in RA is a result of the summation of loss of function in individual joints. Overall joint function depends upon numerous factors including severity of disease activity, joint damage, muscle strength, and psychosocial factors³⁻⁸. It is important to study this issue to determine the success of physicians' (specifically rheumatologists') ability to detect and quantify functional disability in patients with RA.

Patient questionnaires have become increasingly utilized in rheumatology since the publication of the Stanford Health Assessment Questionnaire-Disability Index (HAQ-DI) in 1980⁹. These functional disability indices, such as the HAQ-DI and the Arthritis Impact Measurement Scale Health Status Questionnaire (AIMS)¹⁰, are based on patient self-reporting. The value of these patient questionnaires is now well established in clinical research. Patient questionnaires may be more effective than traditional methods like joint count, radiographs, and laboratory tests to predict and document severe consequences of RA including functional indices¹¹.

Our study was designed to test rheumatologists' accuracy in determining functional disability of their patients with RA. This was accomplished by comparing the rheumatologists' HAQ-DI scores with those of patients. We propose that a difference of 0.22 on the HAQ-DI represents a significant difference between the patients' HAQ-DI scores and those of the physicians. It has been established that a minimal significant improvement in a patient's HAQ-DI score on consecutive occasions is 0.22¹². It has also been suggested that any difference greater than 0.48 on the HAQ-DI represents a significant change in disability from the patient's perspective¹³. There are no data available for the physician's HAQ-DI score of patients' functional disability.

MATERIALS AND METHODS

Study design. We used the HAQ-DI as our guide to assessing functional disability in patients with RA. The HAQ-DI is a 20-question questionnaire. Each question is scored on a 4-point scale (0 to 3, 0 being the least severe and 3 the most severe). It also includes a pain analog scale (PAS). The patients' scores were considered the gold standard (since the HAQ-DI is validated for the patient) and the rheumatologists' scores were compared to the patients'. At the conclusion of the study, the absolute value difference of the patients' scores versus the rheumatologists' scores was summed and the mean difference was obtained (mHAQ-Diff).

The PAS is a validated 15 cm visual analog scale (VAS). We used this PAS to assess the level of pain in these patients with RA. Both rheumatologists and patients marked the level of pain on the 15 cm line (VAS). The distance from the left side of the line was measured and results were plotted on a scale of 0 to 3. Similar to the mHAQ-Diff, the absolute value differences of

the patients' PAS scores versus the rheumatologists' scores were summed and the mean difference was obtained (mPAS-Diff) at the study's conclusion.

The study consisted of one evaluation per patient. This was performed at the time of their regular scheduled visit with their rheumatologist. The purpose of this visit was not a formal work disability assessment; rather it was part of their routine clinical care. This evaluation was referred to as the index visit. At the index visit, both the patient and the rheumatologist completed an entire HAQ-DI including the PAS. The rheumatologist would answer each question of the HAQ-DI as an estimate from his or her history and examination from that index visit. The patients and the rheumatologists filled out the HAQ-DI (including the PAS) independently immediately after the index visit, and their respective scores were not shared.

Our study was performed at the University of South Florida (USF) Rheumatology Clinics and James A. Haley Veterans Affairs Medical Center in Tampa, Florida. The USF Institutional Review Board and James A. Haley Veterans Administration's Research and Design Committee approved the study.

Patient population. Male and female patients 18 to 65 years of age were diagnosed with RA according to the American College of Rheumatology criteria. Demographic data were collected by chart review (with the exception of the presence of rheumatoid nodules that were determined by examination). These included the patients' age, disease duration, rheumatoid factor (RF) status, presence of rheumatoid nodules, absence or presence of erosive disease, and class and stage of their disease.

Because our study was designed to assess functional disability in the setting of RA, patients with class III and IV congestive heart failure, cerebrovascular accidents with hemiparesis, chronic obstructive pulmonary disease on home oxygen, and other severe comorbidities that would adversely affect functional status, or who were wheelchair-bound for any reason other than RA, were excluded. The study was explained to all patients and all patients signed the informed consent. No compensation was provided to the patients.

The sample size was calculated using 80% power to detect a 0.22-point absolute difference in the patients' HAQ-DI versus the physicians' HAQ-DI. The level of statistical significance was 5%. This yielded a sample size of 221 patients.

Study endpoints. The primary endpoint was the mean difference in the patients' HAQ-DI scores versus that of the rheumatologists (mHAQ-Diff). The differences were computed on individual patient-rheumatologist dyads and then the mean difference was calculated. This would determine how effective rheumatologists are at determining an RA patient's physical disability compared to the patient's own assessment. Secondary endpoints were the mean difference in the PAS scores (also calculated from individual patient-rheumatologist dyads), and assessing to see if the rheumatologists' HAQ-DI was influenced by the patients' age, sex, disease duration, RF status, rheumatoid nodules, absence or presence of erosive disease, and class or stage of the RA.

Statistical analysis. Based on our sample size determination of 221 study subjects, we collected data on 223 patients with RA. Eight rheumatologists participated in the study. The data were analyzed using the Pearson rank correlation test to evaluate the statistical significance of correlation between the patients' and the rheumatologists' scores. Correlation is a measure of linear association between 2 variables. Squaring the coefficient gives the percentage of the variation in the 2 variables that is in common. A paired t test and a Bland-Altman plot were used to determine the accuracy of the rheumatologists' scores compared to the study subjects'. The Bland Altman plot¹⁴ is a statistical method to compare 2 measurements where a paired difference can be plotted against one of the components to examine any systematic difference. A paired t test was performed to determine the effect of disease severity (stage and class of RA) or other clinical factors (age, sex, RF status, nodules, erosions) on the accuracy of the rheumatologists' scores. Finally, analysis of variance (ANOVA) was used to detect a difference in the means of the paired differences in HAQ-DI scores between strata defined by the subject's stage of RA. Point estimates and 95% confidence intervals (CI) of the mean difference in each of the 3 strata were calculated.

RESULTS

We evaluated 223 patients (139 female and 84 male; see Table 1 for patient demographics). The attending rheumatologists performed the vast majority of the physicians' HAQ-DI (207/223), with the remaining 16 performed by rheumatology fellows. Figure 1 displays the patients' HAQ-DI scores versus those of the rheumatologists. The Pearson correlation coefficient was 0.59 (95% CI 0.50–0.67). The plot tells us that only 35% of the variation in one variable can be predicted from the other (weak to moderate correlation).

In order to determine the accuracy of the rheumatologists'

Table 1. Patient demographic data.

	Patients (n = 223)
Average age, yrs	52.4 (SD 9.5)
Average disease duration, yrs	8.8 (SD 7.5)
Female, n (%)	139 (62)
Male, n (%)	84 (37)
Rheumatoid factor +, n (%)	126 (56)
Radiographic erosions, n (%)	87 (39)
Rheumatoid nodules, n (%)	44 (20)
Disease stage, n (%)	
I	55 (25)
II	117 (52)
III	48 (22)
IV	3 (1)
Disease class, n (%)	
I	48 (22)
II	135 (60)
III	33 (15)
IV	7 (3)

assessments a paired t test was conducted using the patient scores and the rheumatologist scores on the HAQ-DI (Figure 1). The test statistic was $t = 1.874$ on 222 degrees of freedom, giving a p value = 0.03. The mean of the differences was -0.3 (95% CI -0.37 to -0.22). This shows that the patients' versus the physicians' HAQ-DI were statistically significantly different.

Figure 2 is a version of a Bland-Altman plot, which plots the paired differences versus the patient HAQ scores. The differences were obtained by subtracting the rheumatologist's score from the patient's score. This plot allows us to more clearly see a systematic difference in the patient and rheumatologist scores. Scores overestimated by the rheumatologists were more common, as evidenced by the majority of the paired differences being less than 0 with a mean of -0.3 .

Similar to the HAQ-DI scores, PAS scores were also compared using a paired sample t test: the absolute mean difference between patient scores and rheumatologist scores was 0.16 ($p = 0.53$). This value was not statistically significant because of a relatively equal distribution of over- and under-estimation of the patient's pain by the rheumatologist.

We analyzed the data in regard to patient age, sex, presence or absence of RF, nodules, erosions, and disease stage and functional class of RA by using paired sample t tests. There was no significant difference between the scores based on age, sex, presence or absence of RF, erosions, or rheumatoid nodules. However, the rheumatologists' HAQ-DI and PAS did seem to be influenced by the patients' stage or functional class. For example, independent 2-sample t tests showed that the rheumatologists significantly overestimated the patients'

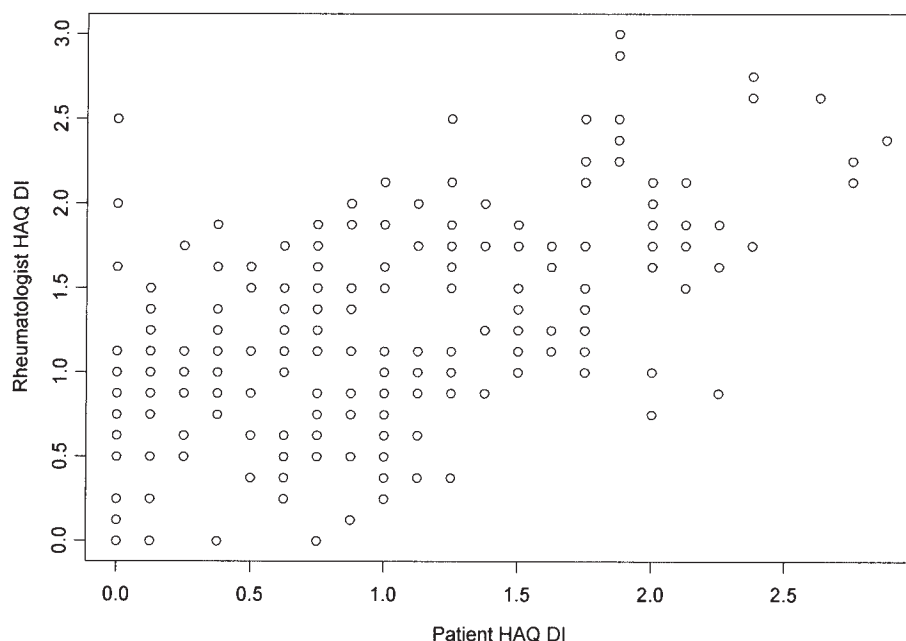


Figure 1. Scatterplot shows patient HAQ-DI scores plotted against the physician-estimated HAQ-DI scores for 223 subjects. Pearson correlation coefficient = 0.59 (95% CI 0.50, 0.67). Graph depicts weak to moderate correlation between patient and physician HAQ-DI.

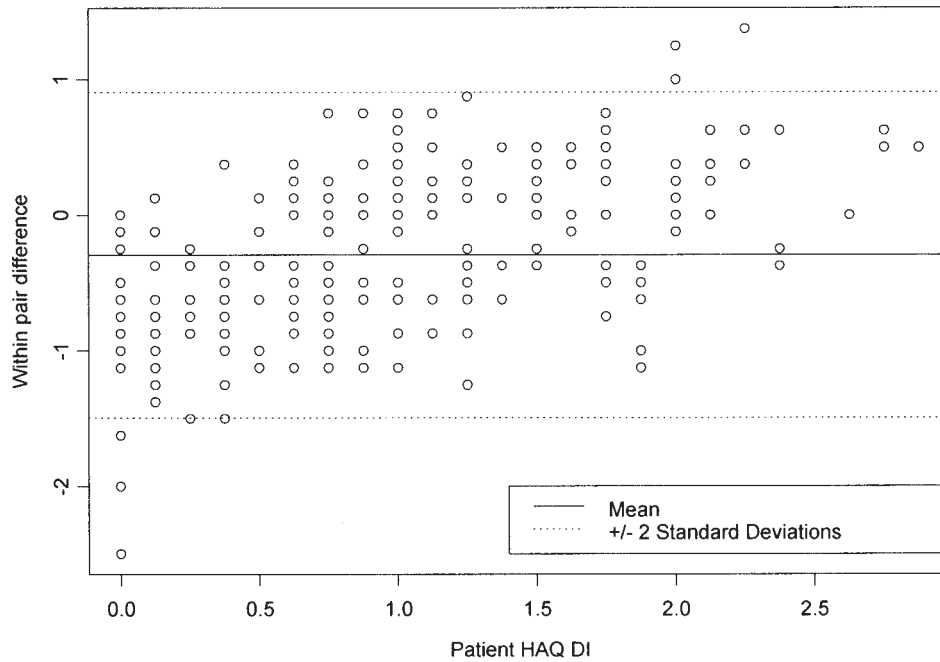


Figure 2. Bland-Altman plot of the paired difference obtained by subtracting rheumatologist's score from patient's score versus patient HAQ-DI scores. Solid line is the mean of the paired differences (0.3) and the 2 broken lines are the values \pm SD from the mean (0.9, -1.5).

functional disability (HAQ-DI) when comparing patients with stage III versus stage I ($p = 0.04$). Regarding the PAS, the rheumatologists' estimates were more accurate in patients with stage I versus stage II ($p = 0.007$) and class I versus class II ($p < 0.001$). The rheumatologists were most accurate at determining the level of functional disability and pain in patients with stage I or class I RA. There were too few patients with stage IV or class IV RA to draw any conclusions.

Paired differences in HAQ-DI scores between stages of RA were compared using one-way ANOVA with stage as the factor (Table 2). These data indicate that there is a significant difference ($p < 0.0001$) in the means between at least 2 of the 3 groups. Multiple comparisons showed that the differences in

Table 2. Analysis of variance: the number (7.67) under sum of squares and in the stage row is known as the treatment sum of squares (SSTR). This is a measure of the extent of the difference between the estimated factor level means (stages). 71.57 is the error sum of squares (SSE). This is a measure of the random variation of the observations around the respective estimated factor level means. Dividing both the SSTR and the SSE by their respective degrees of freedom they become treatment mean square (MSTR) and error mean square MSE. Then the test statistic F^* is calculated by the formula $F^* = MSTR/MSE = 11.57$. F^* has an F distribution and the probability of obtaining a result as large or larger than 11.57 is $p < 0.0001$. Thus, there is a significant difference between the means.

	Df	Sum of Squares	Mean Sq	F* Value	Pr(F)
Stage	2	7.67	3.83	11.57	< 0.0001
Residuals	216	71.57	0.33		

stage 1 subjects were smaller than those in stage 2 or stage 3. Paired t tests were conducted after stratifying the subjects based on the stage of RA. Confidence intervals were also calculated for the mean of the paired differences in each stage (1–3). The paired t test for those in stage 1 resulted in a t-statistic = 0.50 and p value = 0.62. Thus, the mean of the differences is not considered statistically different from zero. However, the t tests for the differences in subjects in stages 2 and 3 were statistically significant from zero. The t-statistics were -7.34 for stage 2 and -4.57 for stage 3 ($p < 0.0001$ for both). These results and the confidence intervals (Table 3) also show that the rheumatologists were more accurate in predicting the HAQ-DI scores in subjects with stage 1 RA.

We also analyzed individual rheumatologists' accuracy at determining the HAQ-DI (Figure 3). Each rheumatologist's mHAQ-Diff was rather similar; however, this study was not powered to determine whether each mHAQ-Diff was significant for each rheumatologist. Three attending rheumatologists (physicians A, B, and C in Figure 3) analyzed 194/223 (87%) patients. All 3 of these rheumatologists are board certified in rheumatology with the following years of experience (A: 30

Table 3. Point estimates (95% CI) for the mHAQ-Diff.

Stage of RA	mHAQ-Diff (95%CI)
1	0.03 (-0.11 to 0.18)
2	-0.40 (-0.51 to -0.29)
3	-0.39 (-0.57 to -0.22)

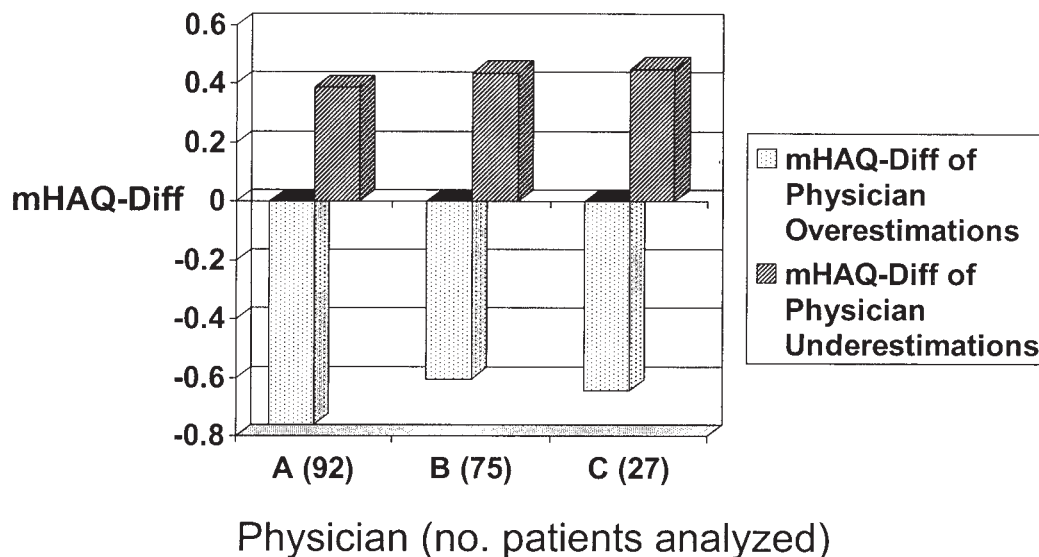


Figure 3. HAQ-DI prediction error (mHAQ-Diff = patient mHAQ-Diff – physician mHAQ-Diff). Three attending rheumatologists (physicians A, B, and C) analyzed 194/223 (87%) patients.

yrs; B: 5 yrs; C: 15 yrs). These rheumatologists consistently overestimated the functional disability (HAQ-DI) of their patients. Overall, the rheumatologists overestimated the HAQ-DI in 154 of the 223 index visits (69% of the time).

DISCUSSION

There are no data available regarding a physician's accuracy at estimating RA patients' functional disability by using the HAQ-DI. However, a previous smaller study of 46 subjects suggested that functional disability of patients with RA could be assessed with sufficient precision by performing a measure of disease activity [Ritchie Articular Index (RAI)]¹⁵. A more recent study suggested that the number of swollen and tender joints is less important than pain and joint mobility in determining a patient's HAQ-DI⁶. Our study suggests that physicians (specifically rheumatologists) are not accurate at determining the level of functional disability of their patients with RA, as represented by a significant difference between their HAQ-DI scores and the patients'. This conclusion hinges on the premise that patients are indeed the gold standard for determining the amount of their own functional disability. Because the HAQ-DI is validated for the patient, we felt it was appropriate to consider their scores as the standard, thus justifying our conclusion that rheumatologists are inaccurate at determining the level of disability in their patients with RA.

Although the primary endpoint of this study evaluated the rheumatologists' accuracy at determining the HAQ-DI, it was not our purpose to suggest that a rheumatologist's HAQ-DI might be a suitable replacement for the patient's. The patient's HAQ-DI is a validated useful measure of functional disability. Rather, physicians (specifically rheumatologists) are often asked to comment on their patients' degree of functional disability by completing formal assessments. These determina-

tions may relate to work disability or their functional status in everyday life. It is our experience that most physicians (including rheumatologists) do not employ formal validated measures, such as a RAI or complete joint mobility scores, when determining the functional disability of their patients. Rather, their patients are seen and examined as part of their routine care and these assessments of employment or daily lifestyle disability are determined based on impressions from their routine care. In our study the rheumatologists performed their routine visit with these patients with RA and then performed a HAQ-DI on these same patients immediately thereafter. This routine assessment of their patients is not adequate at determining their functional disability.

Another interesting finding in our study was that rheumatologists almost universally overestimated the functional disability of their patients with RA. The reasons for this are unclear, but may include empathy or an unrealistically low expectation of the degree of activities of daily living of which patients with RA are capable. These factors may or may not apply when patients are formally evaluated for employment ability.

While the patients' age, sex, RF status, rheumatoid nodules, or presence of erosions had no influence on the rheumatologists' estimated HAQ-DI or PAS scores, the patients' stage and class of their RA did seem to affect their scores. Overall, the rheumatologists' HAQ-DI scores were significantly different than the patients', but this was particularly true for patients with advanced class and stage of RA. Patients' gross visible deformities tend to correlate with their stage/class of RA. Thus, in spite of a rheumatologist's in-depth knowledge of the pathophysiology of RA, visible deformities seem to be the most important factor influencing the rheumatologists' overestimation of their patients' functional

disability.

Why was there a significant difference between the rheumatologists' and the patients' assessment of their disability? Did the rheumatologists fail? There are many factors involved in the evaluation of functional disability of a patient with RA, such as disease activity, the patient's perception of pain, disease duration, and joint deformities. However, the rheumatologists significantly and consistently overestimated the patients' level of functional disability. It can be argued that there can be no better authority on the patient's amount of functional disability than the patient. Alternatively, there may also be motivation for the patient to overestimate their functional disability, such as during employment ability examinations. There are very few, if any, reasons for a patient to underestimate their amount of functional disability. Therefore, the rheumatologists' perception of the patients' functional disability appears to be a true overestimation.

A potential limitation of our study relates to the extrapolation of functional capacity as it relates to employment capability in patients with RA. The HAQ-DI scores in our study were obtained at the patient's regularly scheduled visit for their RA. There were no specific maneuvers or measures other than those performed at a routine visit. The functional capacity of patients with RA as it relates to their employment capability is sometimes evaluated in a more formal setting in which their physician measures specific criteria that correlate with specific job functions. However, persons with RA tend to select jobs that they can perform with their RA limitations. Specific employment evaluations (such as the Work Limitations Questionnaire) of people with RA incorporate these limitations into the assessment, making the HAQ-DI a more accurate assessment of overall functional limitations¹⁶.

Our study is important because, as rheumatologists, we have to assess the functional disability of patients with RA on a continual basis. We often have to assess the functional disability of patients with RA as it relates to their employment capabilities; this is frequently done as part of their routine care. The results indicate that there is a clear difference between patients' and rheumatologists' assessment of their functional disability in RA.

REFERENCES

1. Spector TD. Rheumatoid arthritis. *Rheum Dis Clin North Am* 1990;16:513-37.
2. Pincus T, Callahan LF, Sale WG, Brooks AL, Payne LE, Vaughn WK. Severe functional declines, work disability, and increased mortality in seventy-five rheumatoid arthritis patients studied over nine years. *Arthritis Rheum* 1984;27:864-72.
3. Escalante A, del Rincon I. How much disability in rheumatoid arthritis is explained by rheumatoid arthritis? *Arthritis Rheum* 1999;42:1712-21.
4. Welsing PM, van Gestel AM, Swinkels HL, Kiemeny LA, van Riel PL. The relationship between disease activity, joint destruction, and functional capacity over the course of rheumatoid arthritis. *Arthritis Rheum* 2001;44:2009-17.
5. Makisara GL, Makisara P. Prognosis of functional capacity and work capacity in rheumatoid arthritis. *Clin Rheumatol* 1982; 1:117-25.
6. Hakkinen A, Kautiainen H, Hannonen P, Ylinen J, Arkela-Kautiainen M, Sokka T. 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.
7. Hakkinen A, Kautiainen H, Hannonen P, Ylinen J, Makinen H, Sokka T. Muscle strength, pain, and disease activity explain individual subdimensions of the Health Assessment Questionnaire Disability Index, especially in women with rheumatoid arthritis. *Ann Rheum Dis* 2006;65:30-4.
8. Sokka T, Kankainen A, Hannonen P. Scores for functional disability in patients with rheumatoid arthritis are correlated at higher levels with pain scores than with radiographic scores. *Arthritis Rheum* 2000;43:386-9.
9. Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcomes in arthritis. *Arthritis Rheum* 1980;23:137-45.
10. Meenan RF, Gertman PM, Mason JH. Measuring health status in arthritis: the Arthritis Measurement Scales. *Arthritis Rheum* 1980;23:146-52.
11. Wolfe F, Cathey MA. The assessment and prediction of functional disability in rheumatoid arthritis. *J Rheumatol* 1991;18:1298-306.
12. Wells GA, Tugwell P, Kraag GR, Baker PR, Groh J, Redelmeier DA. Minimum important difference between patients with rheumatoid arthritis: the patient's perspective. *J Rheumatol* 1993;20:557-60.
13. Greenwood MC, Doyle DV, Ensor M. Does the Stanford Health Assessment Questionnaire have potential as a monitoring tool for subjects with rheumatoid arthritis? *Ann Rheum Dis* 2001;60:344-8.
14. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res* 1999;8:135-60.
15. Singer F, Kolarz G, Mayrhofer F, Scherak O, Thumb N. The use of questionnaires in the evaluation of the functional capacity in rheumatoid arthritis. *Clin Rheumatol* 1982;1:251-61.
16. Walker N, Michaud K, Wolfe F. Work limitations among working persons with rheumatoid arthritis: results, reliability, and validity of the work limitations questionnaire in 836 patients. *J Rheumatol*