The Journal of Rheumatology

VOLUME 21: NO. 7 JULY 1994

Editorials

Articles

Contents continued opposite inside back cover . . .

PRINTED IN CANADA — ISSN 0315-162X

Pediatric Rheumatology

Case Reports

Correspondence

Letters

Book Reviews

Evaluation of a German Version of the Physical Dimensions of the Health Assessment Questionnaire in Patients with Rheumatoid Arthritis

PIUS BRÜHLMANN, GEROLD STUCKI, and BEAT A. MICHEL

ABSTRACT. ***Objective.*** Our objective was to translate and adapt the disability section of the health assessment questionnaire (HAQ) into German (HAQ-G) to suit Swiss-German conditions and to test its metric properties, reliability, and validity.

Methods. We tested 62 consecutive patients with rheumatoid arthritis (RA) attending the outpatient Clinic of the Department of Rheumatology, University Hospital Zurich. All patients fulfilled the American Rheumatism Association 1987 revised criteria for RA. The translation was done by 2 translators aware of the objective of the questionnaire and some questionable items were discussed and resolved in a panel by 4 rheumatologists including one bilingual clinical researcher. Test-retest reliability was assessed with Pearson's correlation coefficient on the scores of 2 questionnaire mailed in a 10-day interval. The internal consistency was assessed with Cronbach's coefficient alpha. To assess the construct validity, we compared the HAQ scores to clinical, laboratory, and radiological variables of disease activity and outcome. To assess criterion validity, we compared physicians' assessment of functional class (observed disability) to the HAQ (referred disability). The content validity was assessed in a multivariate model explaining HAQ scores with a variety of other measurements of disease activity and outcome.

Results. The test-retest reliability was 0.94; the internal consistency was 0.92; the criterion validity was 0.76; and correlations with other disease variables ranged from 0.39 (Larsen radiological score) to 0.66 (grip strength).

Conclusion. The HAQ-G is a reliable and valid instrument for measuring functional disability in a German speaking population with RA. (*J Rheumatol 1994;***21:1245–9**)

Key Indexing Terms:
PHYSICAL FUNCTIONAL DISABILITY RHEUMATOID ARTHRITIS
GERMAN VERSION HEALTH ASSESSEMENT QUESTIONNAIRE

Assessment of rheumatoid arthritis (RA) includes measurements of process, prognosis, and outcome[1]. Outcome is the suffering and loss of health experienced by an individual as a result of the process of disease[2]. Death, physical disability, discomfort, treatment side effects, and economic impact have been identified as primary outcome dimensions. Although the importance of physical disability has been recognized since the earliest days of the speciality of rheumatology[3], reliable and valid measures were not developed until 1980[4]. One of the most widely used instruments to measure functional disability is the Health Assessment Questionnaire (HAQ)[2,5]. It measures function in dressing, arising, walking, hygiene, reaching, gripping, and in other activities. The questionnaire is self-administered and may take less than 5 min to complete[5]. Its reliability and validity has been demonstrated in different languages and contexts[5]. In a number of studies, the HAQ has been shown to detect changes in clinical status and it has been used as a primary endpoint in rheumatic disease trials[6]. The HAQ has predictive validity for work disability, resource utilization[7], peptic ulceration[8], and death[5].

We translated and adapted the disability section of the HAQ into German (HAQ-G) to suit Swiss-German conditions and tested its metric properties, reliability, and validity.

MATERIALS AND METHODS

Patients. We enrolled 62 consecutive patients with RA attending the rheumatology outpatient clinic, University Hospital, Zurich. All patients fulfilled the American Rheumatism Association (ACR) 1987 revised criteria[9] for the classification of RA.

Translation and cultural adaptation of the HAQ. The primary translation was done by 2 translators aware of the objective of the questionnaire. The emphasis was to attempt the best idiomatic rather than pure vocabulary equivalence[10]. A few questionable items were discussed and resolved by 4 rheumatologists including one bilingual clinical researcher with experience in the use of the instrument in clinical studies, 2 with a general internal medicine background, and one with additional training in physical medicine. The following modifications in the Swiss-German adaptation (Table

From the Department of Rheumatology and Physical Medicine, University Hospital, Zurich, Switzerland and the Robert B. Brigham Multipurpose Arthritis and Musculoskeletal Disease Centers, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA.

P. Brühlmann, MD, Assistant Professor of Rheumatology, University Hospital Zurich; G. Stucki, MD, Research Fellow in Rheumatology, Brigham and Women's Hospital; B.A. Michel, MD, Professor of Rheumatology, University Hospital Zurich.

Address reprint requests to Dr. Med. P. Brühlmann, Rheumaklinik und physikalisches Institut, Universitätsspital, Gloriastrasse 25, 8091 Zürich, Switzerland.

Submitted August 30, 1993 revision accepted December 3, 1993.

Table 1. *German version of the HAQ physical disability dimension [all dimensions start with the question: are you able to ("Können Sie")]*

Ankleiden und Körperpflege

- sich selber ankleiden, Kleider zuknöpfen und Schuhe binden?
- Ihre Haare waschen?

Aufstehen

- von einem Stuhl ohne Armlehne aufstehen?
- ins Bett gehen und aufstehen?

Essen

- das Fleisch mit dem Messer schneiden?
- ein gefülltes Glas zum Munde führen?
- einen Milchkarton (Tetrapack) öffnen?

Gehen

- auf ebener Strasse gehen?
- Treppen steigen?

Körperpflege

- sich ganz waschen und abtrocknen?
- ein Vollbad nehmen?
- auf die Toilette gehen?

Heben

- einen 2 kg schweren Gegenstand (z.B. einen Sack Kartoffeln) über Kopfhöhe heben bzw. herunternehmen?
- sich bücken, um ein Kleidungsstück vom Fussboden aufzuheben?

Greifen und Öffnen

- eine Autotüre öffnen?
- ein Konfitürenglas öffnen, welches schon einmal offen war?
- einen Wasserhahn auf- und zudrehen?

Andere Tätigkeiten

- einkaufen gehen?
- in ein Auto ein- und aussteigen?
- Haushaltarbeiten (z.B. Staubsaugen) oder Gartenarbeiten verrichten?

1) were made: (a) "Climb up five steps" was changed to "Can you climb stairs"; (b) "Reach and get down a 5-pound object (such as a bag of sugar) from just above your head" was changed to "Can you lift and reach a 2 kilo object above your head"; (c) the heading "grip" was changed to "grip and open"; (d) aids and devices: "walker" and "built up or special utensils" were dropped since walkers are rarely used as the only device (patients who need a walker have usually also a wheelchair for greater distances; built up or special utensils were specified under "other").

Scoring. Scoring of questionnaires was done in accordance with the original HAQ using the following categories (points): without any difficulty (0), with some difficulty (1), with much difficulty (2), unable to do (3). The highest component score in each category determines the score for the category unless aids or assistance are required. Dependence on equipment or physical assistance increases a category score of < 3 by plus 1. Category scores are averaged to give the disability index, a value from 0 to 3.

Study design. The questionnaire was given to the patients at a regularly scheduled outpatient clinic visit. The same day the patient was evaluated clinically, laboratory tests were performed, and radiographs taken. The metric properties of the instrument and the validity were studied on this data. To assess the test-retest reliability the questionnaire was mailed twice with an

interval of 10 days, 3 months after the cross sectional evaluation. Patients were therefore already familiar with the questionnaire from Time Point 1 with the advantage of excluding a learning effect.

Reliability. The test-retest reliability was assessed with Pearson's correlation coefficient. To assess the internal reliability, an analysis of internal consistency using Cronbach's coefficient alpha was performed on the overall scale as well as on the 3 components addressing upper extremity function (eating, reach, grip) and the 3 components addressing lower extremity function (arising, walking, activity). The relationships between the 8 components were evaluated in a correlation matrix using Spearman's correlation coefficient. In a factor analysis, the loading of the 8 components and total interperson variability explained by the factors was studied.

Validity. To assess criterion validity, we correlated the HAQ (referred disability) to functional disability as recorded by the physician (observed disability) and to the ACR functional class 1991 as recently published by Hochberg, *et al*[11].

To test the construct validity, we hypothesized a strong relationship between functional disability and variables of disease activity (swollen and tender joint count; morning stiffness; physician's global disease activity rated on a numerical rating scale; erythrocyte sedimentation rate (ESR), C-reactive protein (CRP), anatomical damage (Larsen radiological score)[12], pain (11 point numerical rating scale, range 0–10), and global health. We hypothesized that grip strength (measured with a vigorimeter) would correlate most strongly with the grip domain and less with the other domains of the upper extremity and less with lower extremity domains. For walking time we hypothesized that the walk domain would correlate most strongly. We also hypothesized that the physical disability score discriminates between patients with a short versus a long disease duration. For all correlations we used Spearman's correlation coefficients because the variables were either nonparametric or not normally distributed.

If the HAQ has a high content validity in reflecting the impact of the disease, different attributes of the disease process should explain important amounts of its variability. Thus we studied the relationship of physical disability to different aspects of the disease process in a biological multivariate model. The variables for the final model were selected in a stepwise regression analysis of groups of variables (level to stay: $p = 0.2$). The following groups and variables were considered: demographic measures (age, sex), laboratory variables (ESR, CRP, thrombocytes, erythrocytes, lymphocytes, Singer-Plotz titer, Waaler-Rose titer, hemoglobin, iron, ferritin), articular indices (tender joint count, Ritchie count, swollen joint count), performance measurement (grip), symptoms (pain, morning stiffness), anatomical damage (Larsen score), and disease duration. The variables selected in the group analysis were evaluated in a stepwise selection process with a level to enter of 0.1 and to stay of 0.05.

RESULTS

Sixty-two patients completed the questionnaire at Time Point 1 and 50 patients completed the questionnaire at Time Points 2 and 3. All items were filled in, no items gave rise to misunderstanding. The baseline characteristics of the patients are shown in Table 2.

The test-retest reliability for the instrument was 0.94. The reliability for the categories ranged from 0.74 (grip) to 0.94 (hygiene). The internal consistency reached an alpha of 0.92, the correlation of the individual items with the total ranged from 0.53 (walking) to 0.82 (reach). A similar high internal consistency (alpha = 0.92) was found for the 5 domains representing upper extremity functioning. The correlation of the domains with the average upper extremity score ranged from 0.74 (grip) to 0.82 (dressing). The 3 lower extremity domains showed a Cronbach's alpha of 0.72 and the correlation of the items with the average lower extremity score ranged from 0.52 (rise) to 0.56 (activity). The correlation matrix between the domains yielded correlation coefficients of 0.37 (walking and eating) to 0.78 (hygiene and reach) (Table 3). The principal component analysis resulted in one factor explaining 64% of the total variance of the scale.

Table 2. *Baseline demographic and clinical characteristics and HAQ-G scores of the 62 patients with RA*

	Number of Patients (N = 62)	HAQ-G Score (SD)
Sex		
Male	23	1.1 (0.8)
Female	39	1.1 (0.8)
Age group		
≤ 49 (5.7)	15	1.1 (0.8)
49 to 62 (7.2)	18	1.4 (0.7)
62 to 71 (9.7)	14	0.8 (0.8)
≥ 71 (10.3)	15	1.0 (0.8)
Functional class ACR 1991		
I	9	0.1 (0.1)
II	19	0.7 (0.7)
III	32	1.5 (0.4)
IV	2	2.6 (0.3)
Rheumatoid factor (Singer Plotz)	(n = 61)	
Negative	31	0.9 (0.8)
Positive	30	1.3 (0.8)

The correlation of the HAQ with the Steinbrocker scale was 0.6 ($p < 0.01$) and with the ACR 1991 functional scale 0.76 ($p < 0.01$) (Table 2). The correlations between functional disability and disease activity measurements, anatomical damage, discomfort, and global health were all significant ($p < 0.05$) and ranged from 0.39 (Larsen score) to 0.66 (grip strength) (Table 4). Age and sex were not univariate correlates of the HAQ.

The correlation of grip strength with the domains representing upper extremity function ranged from −0.50 (gripping) to −0.63 (reaching) and the correlation with lower extremity domains from −0.29 (walking) to −0.48 (activity). The correlation of walking time with the lower extremity domains ranged from 0.35 (rise) to 0.47 (walking) and with the upper extremity domains from 0.22 (eating) to 0.49 (hygiene).

Patients with a disease duration of < 1.5 years had a much lower average disability score of 0.7 than patients with a disease duration of > 11 years showing an almost doubled HAQ score of 1.3 on average.

For the multivariate model, the following variables had univariately a p value of less than 0.2 in the group selection and were considered for selection: CRP, ESR, ferritin, iron, erythrocyte count, grip and Ritchie joint count, Larsen score, pain, and morning stiffness. In the stepwise forward selection process, grip (explaining 38% of the total variation of the HAQ score), morning stiffness (explaining an additional 17%), ferritin (7%), CRP (4%), and Ritchie joint count

Table 3. *Spearman correlation coefficients among the components of the HAQ-G*

	Eating	Reaching	Grip	Dressing	Hygiene	Arising	Walking	Activity
Eating	1							
Reaching	0.71	1						
Grip	0.73	0.65	1					
Dressing	0.74	0.75	0.62	1				
Hygiene	0.66	0.78	0.63	0.77	1			
Arising	0.63	0.58	0.57	0.62	0.70	1		
Walking	0.37	0.47	0.40	0.45	0.53	0.50	1	
Activity	0.64	0.69	0.66	0.59	0.64	0.50	0.50	1

Table 4. *Relationship of the HAQ-G score with biological disease variables*

Variable	Median (Quartiles)	Correlation with HAQ-G Score†
Swollen joint count	5 (2; 11)	0.31**
Tender joint count	6 (2; 15)	0.57**
Morning stiffness (min)	15 (0; 60)	0.5**
ESR (mm/h)	18 (12; 30)	0.26*
CRP (mg/dl)	1.8 (0; 4, 6)	0.46**
Activity physician (NRS‡ 0–10)	3 (2; 5)	0.57**
Larsen score (range 0–40)	10 (8; 17, 5)	0.39**
Grip strength (kp/cm^2)	0.6 (0.38; 0.9)	−0.66**
Walktime (s/50 m)	32 (27; 35)	0.49**
Global health (Likert 0–3)	2 (1; 2)	0.56**
Pain (NRS 0–10)	5 (3; 6)	0.53**
Disease duration (years)	5.2 (1.44; 11.6)	0.29*

** p value < 0.01, * p value < 0.05.
† Spearman's correlation coefficient. ‡ NRS numerical rating scale.

(2%) were selected. The total model r^2 was 68% ($p < 0.01$).

DISCUSSION

The adaptation of the HAQ for a Swiss-German context required no major cultural adaptation since the cultural similarity is high. Because German and English are linguistic relatives, most items could be translated with vocabulary equivalence. Few questions required a different phrasing to avoid misunderstanding and to guarantee idiomatic equivalence.

The metric properties of the HAQ-G were similar to those presented in the original report[2] and in other adaptations[5]. The internal consistency was high and indicates that the components of the scale measure the same construct. The items correlated all with each other, but there was no redundancy; this indicates that each domain addresses a somewhat different aspect of functional disability. The test-retest reliability was similar to reported evaluations which reported correlations between 0.87 and 0.96[5]. The high correlation with other measurements of functional disability such as the Steinbrocker functional scale and even more with the recently developed ACR functional class 1991 demonstrates that both instruments measure a similar construct. As expected, we found strong correlations with variables of disease activity as well as joint destruction. This is consistent with the concept that functional disability is influenced by process as well as destruction and may explain why grip strength, another combined measurement of both dimensions, was the strongest correlate.

The finding that 68% of the variation of the HAQ-G can be explained by a few biological variables reflects the strong relationship between disease process and functional disability. Since heterogeneous dimensions of the disease including laboratory variables and signs and symptoms explained relevant proportions of the variability of the HAQ-G, it can be concluded that this concept of functional disability comprehensively addresses the impact of RA. The responsiveness of the scale has not yet been studied in this validation. However, the sensitivity of the HAQ has been demonstrated in multiple studies using the original version as well as other cross cultural adaptations. Since we were able to demonstrate the similarity of the instrument with respect to metric properties, reliability, and validity, we expect a similar responsiveness for this German version of the HAQ.

We conclude that the German version of the HAQ retains the characteristics of the American original and is a reliable and valid instrument to measure functional disability in German speaking patients with RA.

ACKNOWLEDGMENT

The authors thank Matthew H. Liang, MD, MPH, for helpful discussions and Susanne Stucki, MEd for help in preparation of the manuscript.

REFERENCES

1. Kirwan JR: A theoretical framework for process, outcome and prognosis in rheumatoid arthritis. *J Rheumatol 1992;*19:333–6.
2. Fries JF, Spitz P, Kraines RG, Holman HR: Measurement of patient outcome in arthritis. *Arthritis Rheum 1980;*23:137–45.
3. Steinbrocker O, Traeger CH, Battman RC: Therapeutic criteria in rheumatoid arthritis. *JAMA 1949;*140:659–62.
4. Liang MH, Katz JN: Measurements of outcome in rheumatoid arthritis. *Baillière's Clin Rheumatol 1992;*6:23–37.
5. Ramey DR, Raynauld JP, Fries JF: The health assessment questionnaire 1992. Status and review. *Arthritis Care Res 1992;*5:119–29.
6. Wolfe F, Pincus T: Standard self-report questionnaires in routine clinical and research practice — an opportunity for patients and rheumatologists. *J Rheumatol 1991;*18:643–4.

7. McNevitt MC, Yelin EH, Henke CJ, Epstein WV: Risk factors for hospitalization and surgery for rheumatoid arthritis: implications for capitated medical payments. *Ann Intern Med 1986;*105:421-8.
8. Taha AS, Morran C, Sturrock RD, Russell RI: The health assessment questionnaire as a predictor of non-steroidal peptic ulceration. *Br J Rheumatol 1993;*32:135-8.
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.
10. Secherst L, Fay TL, Zaidi SMH: Problems of translation in cross-cultural research. *J Cross-Cultural Psychol 1972;*3:41-56.
11. Hochberg MC, Chang RW, Dwosh I, Lindsey S, Pincus T, Wolfe F: The American College of Rheumatology 1991 revised criteria for the classification of global functional status in rheumatoid arthritis. *Arthritis Rheum 1992;*35:498-502.
12. Larsen A, Dale K, Eek M: Radiographic evaluation of rheumatoid arthritis and related conditions by standard reference films. *Acta Radiol Diagn 1977;*18:481-91.