Validity and Responsiveness of the Michigan Hand Questionnaire in Patients With Rheumatoid Arthritis: A Multicenter, International Study

JENNIFER F. WALJEE,[1] KEVIN C. CHUNG,[1] H. MYRA KIM,[2] PATRICIA B. BURNS,[3] FRANK D. BURKE,[4] E. F. SHAW WILGIS,[5] AND DAVID A. FOX[1]

Objective. Millions of patients experience the disabling hand manifestations of rheumatoid arthritis (RA), yet few hand-specific instruments are validated in this population. Our objective was to assess the reliability, validity, and responsiveness of the Michigan Hand Questionnaire (MHQ) in patients with RA.

Methods. At enrollment and at 6 months, 128 RA patients with severe subluxation of the metacarpophalangeal joints completed the MHQ, a 37-item questionnaire with 6 domains: function, activities of daily living (ADL), pain, work, aesthetics, and satisfaction. Reliability was measured using Spearman's correlation coefficients between time periods. Internal consistency was measured using Cronbach's alpha. Construct validity was measured by correlating MHQ responses with the Arthritis Impact Measurement Scales 2 (AIMS2). Responsiveness was measured by calculating standardized response means (SRMs) between time periods.

Results. The MHQ demonstrated good test–retest reliability (r = 0.66, $P < 0.001$). Cronbach's alpha scores were high for ADL ($\alpha = 0.90$), function ($\alpha = 0.87$), aesthetics ($\alpha = 0.79$), and satisfaction ($\alpha = 0.89$), indicating redundancy. The MHQ correlated well with AIMS2 responses. Function (r = −0.63), ADL (r = −0.77), work (r = −0.64), pain (r = 0.59), and summary score (r = −0.74) were correlated with the physical domain. Affect was correlated with ADL (r = −0.47), work (r = −0.47), pain (r = 0.48), and summary score (r = −0.53). Responsiveness was excellent among arthroplasty patients in function (SRM 1.42), ADL (SRM 0.89), aesthetics (SRM 1.23), satisfaction (SRM 1.76), and summary score (SRM 1.61).

Conclusion. The MHQ is easily administered, reliable, and valid to measure rheumatoid hand function, and can be used to measure outcomes in rheumatic hand disease.

INTRODUCTION

Rheumatoid arthritis (RA) is a systemic inflammatory autoimmune disease that results in substantial disability and premature death for more than 1 million individuals in the US (1). Rheumatic hand disease is caused by progressive and irreversible inflammation of the synovial tissue, and joint destruction occurs early in the course of disease. Hand deformity and dysfunction is the most common manifestation of RA; 70% of patients with RA experience disfiguring and painful rheumatoid hand deformities. Up to 30% of patients have radiographic evidence of disease at the time of diagnosis, and more than 60% have radiographic joint changes within 2 years of diagnosis (2). Unlike other chronic diseases, such as osteoarthritis or hypertension, patients are typically diagnosed with RA in young adulthood, and this disease profoundly interferes with their future work productivity, their ability to perform activities of daily living (ADL), and their social interactions. On a societal level, the effect of these lost wages and expensive medical therapies consumes approximately $3.6 billion per year (3,4). Unfortunately, a standardized, hand-specific instrument to measure rheumatoid hand function remains elusive, and there are few accepted guidelines for defining hand disability among RA patients.

A variety of methods have been used to describe health

[1]Jennifer F. Waljee, MD, MPH, MS, Kevin C. Chung, MD, MS, David A. Fox, MD: University of Michigan Health System, Ann Arbor; [2]H. Myra Kim, ScD: University of Michigan, Ann Arbor; [3]Patricia B. Burns, MPH: University of Michigan Medical Center, Ann Arbor; [4]Frank D. Burke, MD: Pulvertaft Hand Centre, Derby, UK; [5]E. F. Shaw Wilgis, MD: Curtis National Hand Center, Baltimore, Maryland.

Address correspondence to Kevin C. Chung, MD, MS, Section of Plastic Surgery, University of Michigan Health System, 2130 Taubman Center, SPC 5340, 1500 East Medical Center Drive, Ann Arbor, MI 48109-5340. E-mail: kecchung@umich.edu.

outcomes related to the hand manifestations of RA. These have ranged from objective measures, such as painful joint counts or grip strength, to more subjective measures, such as patient satisfaction scores and quality of life measures. Although single, objective measures of function, such as range of motion, degree of finger extension lag, grip strength, and pinch power, are relatively simple to obtain, they often do not capture the full extent of patient disability. More complex functional tests can include a battery of tasks, such as the Jebsen-Taylor Test, the Grip Ability Test, or the Arthritis Hand Function Test (5–7). These may provide a better assessment of difficulty with ADL, but often do not account for other important end points such as pain, aesthetics, and patient satisfaction (8).

Patient perception of health status as measured by self-administered instruments has been shown to be a better predictor of functional status and disability compared with objective measures (9). Many instruments have been used to define patient-related outcomes in patients with RA, ranging from general quality of life to hand-specific surveys. General health assessment instruments, such as the visual analog scale, the Short Form 36 and its derivations, and the Health Assessment Questionnaire, may offer a global assessment of functioning, but they are not sensitive to detect the amount of patient disability related specifically to RA or to hand dysfunction (8,10–12). Other authors have used hypothetical scenarios to explore patient-related outcomes, using utility measures to estimate future quality of life. Although such models are useful in decision analyses, they are often difficult to implement in clinical practice and the concepts may be challenging for the patients to grasp (13,14). An ideal instrument should be hand specific, and include not only patient perception of function, but also measures of pain, satisfaction, and hand appearance.

We have prospectively evaluated patients with RA from 3 centers in the US and the UK to determine patient outcomes following surgical intervention (silicone arthroplasty) for metacarpophalangeal (MCP) joint deformity using the Michigan Hand Questionnaire (MHQ). We have achieved excellent long-term followup in this sample, and we have a unique opportunity to describe the reliability, validity, and responsiveness of the MHQ in patients with RA. The MHQ is a self-administered questionnaire that contains 37 items and requires approximately 15 minutes to complete, and has been used successfully in the RA population in prior work (15–25). We hypothesize that the MHQ will accurately and reliably define rheumatoid hand performance, and effectively measure clinical change in hand function over time.

SUBJECTS AND METHODS

The study sample consisted of patients diagnosed with RA by their rheumatologists and referred to one of the following 3 institutions: The University of Michigan (Ann Arbor, Michigan), the Curtis National Hand Center (Baltimore, Maryland), and the Pulvertaft Hand Centre (Derby, UK). The study sample is part of a larger prospective study supported by the National Institutes of Health regarding the use of silicone MCP joint arthroplasty (SMPA) for joint deformities due to RA, which has been described in detail elsewhere (26,27). Patients were included in the study if they were diagnosed with RA, had severe MCP joint deformity, and were deemed appropriate candidates for surgical reconstruction. Additionally, subjects were eligible if they were age ≥18 years and able to complete the study questionnaire in English. Patients were excluded from the study if their comorbid conditions prohibited surgery, they experienced additional hand conditions (swan-neck and boutonniere deformities, extensor tendon ruptures) that would require intervention beyond SMPA, or they had previously undergone MCP joint replacement. After enrollment, participants either elected to receive an SMPA or remained as controls. Data were collected from the subjects at the time of enrollment and at a 6-month followup time. The study protocol was approved by the Institutional Review Boards at the University of Michigan, the Curtis National Hand Center, and the Pulvertaft Hand Centre.

All of the subjects completed the MHQ, which has been previously validated for use in a wide range of patient samples (15–20). The MHQ yields an overall summary score of hand function, as well as scores for 6 specific scales: hand function, ability to complete ADL, pain, work performance, aesthetics, and patient satisfaction. Scores for each domain range from 0–100, and higher scores indicate better performance for all domains with the exception of pain.

The subjects also completed the Arthritis Impact Measurement Scales 2 (AIMS2) questionnaire, a 45-item, self-administered outcomes tool designed to assess health status in patients with inflammatory arthritis and osteoarthritis (28). The AIMS2 is designed to provide a global, self-reported assessment of patient health status and yields information in 4 domains, including physical functioning, affect, symptom, and social interaction. Scores range from 1–10, with lower scores reflecting better health.

All of the subjects underwent the following assessments to provide objective and reproducible measures of hand function at baseline and subsequent followup: grip strength, lateral pinch, 2-point pinch, and 3-point pinch, all measured in kilograms. Subjects also completed the Jebsen-Taylor Test, which is a 7-part, standardized test designed to assess a subject's ability to complete everyday hand-related tasks (29). The writing portion was not included in this analysis due to difficulty of interpretation relating to hand dominance. Time to complete each task was measured in seconds.

Outcomes. *Reliability.* Reliability was defined as the ability of an instrument to consistently or precisely measure a concept of interest (30). In this study, we measured 2 aspects of reliability of the MHQ: the test–retest reliability of the MHQ and the internal consistency of the 6 scales within the MHQ. Test–retest reliability implies that the survey yields similar results from consecutive administration to a subject. To determine the test–retest reliability of the MHQ, we compared responses for each domain of the MHQ at baseline and at the 6-month followup interval. We used responses regarding the symptomatic hand of the

control patients who did not undergo SMPA and the non-operated hand of the SMPA patients. The degree of correlation between the consecutive responses was assessed using Spearman's correlation coefficients. Additionally, we used paired *t*-tests to determine the average difference in score for each domain between these time periods to determine if these means were significantly different. A mean difference of 0 indicates perfect test–retest reliability.

We determined the internal consistency, or homogeneity, of the items included in each scale of the MHQ by calculating Cronbach's alpha for each of the 6 scales in the MHQ. Cronbach's alpha is a measure of the homogeneity of items within a scale, and is based on the number of items included and their degree of correlation according to the following formula:

$$\alpha = \frac{N \cdot \bar{c}}{(\bar{v} + (N - 1) \cdot \bar{c})}$$

where N = the number of items in the scale, $\bar{v}$ = the average variance between subjects in the sample, and $\bar{c}$ = the average covariance between items among the subjects in the sample. Cronbach's alpha values range from 0–1, with higher values indicating greater internal consistency. In general, Cronbach's alpha values between 0.6 and 0.8 are considered acceptable (31). Values greater than 0.8 indicate that there may be redundancy of items in the scale. Cronbach's alpha values that are less than 0.6 indicate that items in the scale are not adequately related to one another to measure a concept (31). The MHQ queried patients regarding their performance in each domain separately for the right and left hand, and we stratified our analysis by hand dominance.

Validity. Validity was defined as the ability of an instrument to accurately measure a concept of interest. For example, patients who score poorly on the MHQ, indicating worse function, would be expected to also have poor performance in other aspects of hand functioning, such as strength and dexterity with specific tasks. Three important types of validity exist: content validity, criterion validity, and construct validity.

Content validity, or face validity, describes the extent to which an instrument appears logical or capable of measuring an outcome of interest to experts within a particular field. The MHQ was developed with strict attention to psychometric principles. It has been validated in a variety of acute and chronic hand conditions and translated into several languages, and therefore is considered appropriate to measure outcomes among patients with RA (16,17,19,20,32).

Criterion validity describes the extent to which an instrument compares with the accepted reference standard. For patients with RA, there is no established standard to measure health outcomes related to hand dysfunction. Therefore, construct validity is often used to establish the validity of outcomes questionnaires.

Construct validity describes the extent to which the scales in the survey instrument behave as expected. For example, patients who report high pain scores would be expected to endorse difficulty with functioning, ADL, and

Table 1. Characteristics of the study sample (n = 128 patients)*	
	Value
Age, mean ± SD years	60.9 ± 9.16
Ethnicity, white	124 (96.9)
Sex	
Female	95 (74.2)
Male	33 (26.8)
Education, less than high school	34 (25.6)
Income, less than $60,000 annual	87 (68.0)
Right hand dominance	116 (90.6)
Underwent SMPA	51 (39.8)

* Values are the number (percentage) unless otherwise indicated. SMPA = silicone metacarpophalangeal joint arthroplasty.

work performance. We established a priori hypothetical relationships between the scales of the MHQ and used Spearman's correlation coefficients to test construct validity against each scale of the MHQ. Additionally, we compared responses to the MHQ among SMPA and control patients with their responses to the AIMS2 questionnaire, an existing, validated measure of health status in RA patients, in order to establish the construct validity.

Responsiveness. Responsiveness was defined as the ability of an instrument to detect important changes in an outcome of interest over time (33). Because the greatest change after surgery occurs within 6 months after SMPA surgery, we used paired *t*-tests to compare mean scores at baseline and at the 6-month followup for each scale and for the summary score. In order to compare the change in scores over time using a standardized method, we calculated the standardized response mean (SRM) for each scale of the MHQ. Ideally, a more sensitive instrument should have a higher SRM (34). Using Cohen's effect size definition, we assumed that an SRM of 0.2 corresponds to a small effect size, 0.5 corresponds to a medium effect size, and 0.8 corresponds to a large effect size (35). The responsiveness of the MHQ was determined separately for SMPA patients and for control patients. Finally, we compared the responsiveness of the MHQ with the responsiveness of each domain of the AIMS2 survey as an additional validation measure.

Statistical analysis. Descriptive statistics were generated to describe the sociodemographic characteristics of the study sample. Statistical significance was set at an alpha level of 0.05. All of the analyses were conducted using Stata, version 10.1. (StataCorp).

RESULTS

The characteristics of the study sample are detailed in Table 1. Of the 160 patients enrolled, 128 patients completed followup at 6 months, with a loss to followup rate of 20%. The majority of the patients were white (97%), women (74%), and right-hand dominant (91%), with a mean age of 60.9 years. Approximately 26% had less than a high school education, and 70% earned an annual income of less than $60,000 per year.

Table 2. Test–retest correlation comparing baseline and 6-month followup scores for the 6 domains of the MHQ (n = 128 patients)*

	Control patients: symptomatic hand (n = 77)		SMPA patients: nonoperated hand (n = 51)	
	Spearman's correlation coefficient	P	Spearman's correlation coefficient	P
Overall function	0.63	0.001	0.71	< 0.001
Activities of daily living	0.63	0.001	0.79	< 0.001
Work performance	0.75	0.001	0.67	< 0.001
Aesthetics	0.70	0.001	0.50	< 0.001
Pain	0.69	0.001	0.63	< 0.001
Patient satisfaction	0.70	0.001	0.76	< 0.001
Summary MHQ score	0.75	0.001	0.71	< 0.001

* MHQ = Michigan Hand Questionnaire; SMPA = silicone metacarpophalangeal joint arthroplasty.

The test–retest reliability of the MHQ was measured by calculating the correlation between responses to the MHQ at baseline and at 6 months of followup (Table 2). Responses regarding the affected hand of the control patients and the nonoperated hand of the case patients were used for this portion of the analysis, and the results are stratified to reflect the differences between these 2 groups. Overall, correlations between responses for each time period were high, indicating good reliability of the MHQ. Among the control patients, responses for work performance were most strongly correlated (r = 0.75, $P < 0.001$), as well as aesthetic appearance (r = 0.70, $P < 0.001$), patient satisfaction (r = 0.70, $P < 0.001$), and the summary MHQ score (r = 0.75, $P < 0.001$). Responses regarding function (r = 0.63, $P < 0.001$) and ADL (r = 0.63, $P < 0.001$) were the least strongly correlated (r = 0.61). Among the case patients, responses were highly correlated for ADL (r = 0.79, $P < 0.001$) and patient satisfaction (r = 0.76, $P < 0.001$). Responses regarding aesthetics were the least well correlated (r = 0.50, $P < 0.001$). Figure 1 describes the mean differences between scores for each scale between time points of administration. Figure 1A describes the mean difference in responses for the affected hand of control patients. Overall, the differences in score between each administration were small, without statistically significant differences between the time periods. Figure 1B describes the mean difference in responses regarding the nonoperated hand of the patients who underwent SMPA. Interestingly, responses regarding the control hand were significantly different between each time period for all domains, with the exception of pain. These patients reported higher improved scores regarding function (baseline mean 51.3, followup mean 60.5; $P < 0.001$), ADL (baseline mean 44.2, followup mean 56.7; $P < 0.001$), work (baseline mean 41.9, followup mean 52.3; $P < 0.002$), aesthetics (baseline mean 50.5, followup mean 59.7; $P < 0.006$), and patient satisfaction (baseline mean 44.2, followup mean 53.4; $P < 0.001$).

Reliability of the MHQ was also assessed by determining the internal consistency of items within each scale of the MHQ, as measured by Cronbach's alpha, stratified by hand dominance (Table 3). As described above, the ideal value for Cronbach's alpha should lie between 0.6 and 0.8. Cronbach's alpha values that are less than 0.6 indicate heterogeneity among items in the scale, and values greater than 0.8 indicate redundancy of items within a scale. Across all of the domains, Cronbach's alpha values were similar by hand dominance. For the MHQ, Cronbach's alpha scores were within the ideal range for aesthetics ($\alpha = 0.75$ for right hand, right-hand dominant; $\alpha = 0.78$ for left hand, left-hand dominant). Cronbach's alpha was high for ADL ($\alpha = 0.91$ for right hand, right-hand dominant; $\alpha = 0.87$ for left hand, left-hand dominant), function ($\alpha = 0.86$ for right hand, right-hand dominant; $\alpha = 0.83$ for left hand, left-hand dominant), pain ($\alpha = 0.85$ for right hand, right-hand dominant; $\alpha = 0.81$ for left hand, left-hand dominant), satisfaction ($\alpha = 0.89$ for right hand, right-hand dominant; $\alpha = 0.90$ for left hand, left-hand dominant), and work performance ($\alpha = 0.91$ for right-hand dominant; $\alpha = 0.88$ for left-hand dominant). These high values may indicate that redundant items exist within the scales of the MHQ.

To test the construct validity of the MHQ, we tested the responses to each scale against the other scales in the MHQ to determine if each scale behaves in an expected manner using Spearman's correlation coefficients (Table 4). For example, we would expect a higher correlation between function and ability to complete ADL than between aes-
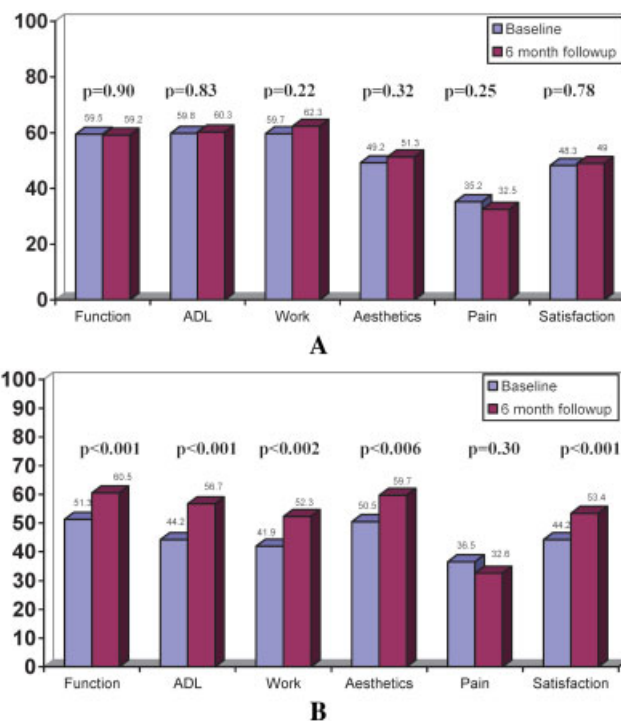
Figure 1. The mean difference in responses for each scale of the Michigan Hand Questionnaire measured at baseline and at 6 months of followup. **A**, Patients who did not undergo silicone metacarpophalangeal joint arthroplasty (symptomatic hand; n = 77). **B**, Patients who underwent silicone metacarpophalangeal joint arthroplasty (control hand; n = 51). ADL = activities of daily living.

Table 3. Internal consistency of the Michigan Hand Questionnaire in rheumatoid arthritis patients as measured by Cronbach's alpha (n = 128), stratified by hand dominance

	Cronbach's α
Function	
Right hand	
Dominant hand	0.86
Nondominant hand	0.77
Left hand	
Dominant hand	0.83
Nondominant hand	0.88
Activities of daily living	
Right hand	
Dominant hand	0.91
Nondominant hand	0.89
Left hand	
Dominant hand	0.87
Nondominant hand	0.95
Both hands	
Left-hand dominant	0.93
Right-hand dominant	0.85
Work performance	
Right-hand dominant	0.91
Left-hand dominant	0.88
Pain	
Right hand	
Dominant hand	0.85
Nondominant hand	0.89
Left hand	
Dominant hand	0.81
Nondominant hand	0.84
Aesthetics	
Right hand	
Dominant hand	0.75
Nondominant hand	0.83
Left hand	
Dominant hand	0.78
Nondominant hand	0.80
Patient satisfaction	
Right hand	
Dominant hand	0.89
Nondominant hand	0.87
Left hand	
Dominant hand	0.90
Nondominant hand	0.88

thetics and ability to complete ADL. For the majority of scales, responses to the MHQ were correlated with the other scales in the expected direction. For example, function was more correlated with ADL (r = 0.83), work performance (r = 0.65), and pain (r = −0.65) than with aesthetics (r = 0.43). As expected, aesthetics was least correlated with work performance (r = 0.38). Satisfaction was correlated most strongly with function (r = 0.81) and ADL (r = 0.73) than with pain (r = −0.69), aesthetics (r = 0.41), or work performance (r = 0.54).

To further assess construct validity, we compared responses to each scale of the MHQ with responses to the AIMS2 domains, as well as objective measures of hand functioning (Table 5). As expected, function (r = −0.63), ADL (r = −0.77), work performance (r = −0.64), pain (r = 0.59), and summary MHQ score (r = −0.74) were strongly correlated with the physical domain of the AIMS2 survey. The affect domain of the AIMS2 was most strongly correlated with the summary MHQ score (r = −0.53), and the symptom domain of the AIMS2 was most strongly correlated with pain (r = 0.70). The social domain was not well correlated with any of the MHQ scales. This suggests that the MHQ may not capture some elements of the effect of RA on social interaction and patient affect that are measured by the AIMS2.

The MHQ was also correlated well with objective measures of functioning, including grip and pinch strength and the Jebsen-Taylor score (Table 5). The function, ADL, and satisfaction domains were most highly correlated with these measures; aesthetics was the least correlated. For example, correlations were highest with Jebsen-Taylor scores (function r = −0.50, ADL r = −0.50, satisfaction r = −0.38; *P* < 0.001) and palmar pinch (function r = 0.45, ADL r = 0.47, satisfaction r = 0.32; *P* < 0.001). Summary MHQ score was most correlated with grip strength (r = 0.34, *P* < 0.12), palmar pinch (r = 0.37, *P* < 0.001), and Jebsen-Taylor score (r = −0.46, *P* < 0.001).

We calculated the responsiveness of the MHQ to detect clinical change in hand function over the 6-month study period. The summary MHQ score and scores for each scale of the MHQ at baseline and at the 6-month followup and the SRMs for each scale of the MHQ are shown in Table 6. As expected, the MHQ demonstrated strong responsiveness to clinical change in the group of patients who underwent SMPA. SRMs were high for function (SRM 1.42), ADL (SRM 0.89), aesthetic appearance (SRM 1.23), satisfaction (SRM 1.76), and the summary MHQ score (SRM

Table 4. The correlation between the 6 scales of the MHQ to test the construct validity of the MHQ (n = 128)*

	Function	ADL	Work performance	Aesthetics	Pain	Patient satisfaction
Function	–					
ADL	0.83	–				
Work performance	0.65	0.67	–			
Aesthetics	0.43	0.46	0.38	–		
Pain	−0.65	−0.63	−0.58	−0.50	–	
Patient satisfaction	0.81	0.73	0.54	0.41	−0.69	–
Summary MHQ score	0.89	0.88	0.77	0.67	−0.82	0.87

* MHQ = Michigan Hand Questionnaire; ADL = activities of daily living.

Table 5. The correlation of the 6 scales of the Michigan Hand Questionnaire with the AIMS2 domains and objective measures of hand functioning (n = 128)*

	AIMS2				Objective hand function testing				
	Physical	**Affect**	**Symptom**	**Social**	**Grip strength, kg**	**Key (lateral) pinch, kg**	**2-point (tip) pinch, kg**	**3-jaw (palmar) pinch, kg**	**Jebsen-Taylor, seconds**
Function	−0.63	−0.41	−0.48	−0.23	0.33	0.34	0.44	0.45	−0.50
ADL	−0.77	−0.47	−0.50	−0.28	0.35	0.32	0.40	0.47	−0.50
Work performance	−0.64	−0.47	−0.55	−0.33	0.25	0.22	0.23	0.26	−0.44
Aesthetics	−0.38	−0.36	−0.35	−0.20	0.16	0.03	0.12	0.18	−0.22
Pain	0.59	0.48	0.70	0.27	−0.23	−0.12	−0.12	−0.14	0.25
Patient satisfaction	−0.54	−0.42	−0.49	−0.24	0.35	0.23	0.31	0.32	−0.38
Summary score	−0.74	−0.53	−0.63	−0.32	0.34	0.26	0.33	0.37	−0.46

* AIMS2 = Arthritis Impact Measurement Scales 2; ADL = activities of daily living.

1.61). Responsiveness was lower for pain (SRM 0.63) and work performance (SRM 0.47). Responsiveness was higher for all domains of the MHQ when compared with that of the AIMS2 instrument among patients who have undergone SMPA. With respect to the control patients, changes for all of the measures over a 6-month time period were modest, and there were no statistically significant differences between mean scores for any measure. Work performance was the most responsive measure over time (SRM

0.14), although this is overall a very low effect. Because these patients did not undergo surgical reconstruction, we would not expect to see a large change in their hand performance over this brief period of time.

DISCUSSION

The MHQ is a hand-specific outcome measurement tool that has been extensively studied in a variety of acute and

Table 6. Responsiveness of the MHQ to clinical change over a 6-month followup period*

	Baseline, mean ± SD	6 months, mean ± SD	P†	SRM‡
SMPA patients (n = 51)				
MHQ				
Summary score	38.3 ± 18.4	62.7 ± 20.8	0.001	1.61
Function	37.6 ± 23.0	65.2 ± 20.3	0.001	1.42
ADL	36.6 ± 27.4	55.9 ± 29.4	0.001	0.89
Work performance	41.9 ± 23.0	52.3 ± 29.1	0.002	0.47
Pain	48.2 ± 26.3	33.4 ± 24.9	0.001	0.63
Aesthetics	34.3 ± 22.4	71.0 ± 23.5	0.001	1.23
Satisfaction	27.6 ± 20.2	65.6 ± 25.0	0.001	1.76
AIMS2				
Physical	3.6 ± 2.4	3.2 ± 2.3	0.002	0.46
Affect	3.8 ± 1.9	3.5 ± 2.0	0.012	0.36
Symptom	5.3 ± 2.8	5.0 ± 2.7	0.18	0.19
Social	4.1 ± 2.1	3.8 ± 2.0	0.11	0.23
Controls (n = 77)				
MHQ				
Summary score	56.8 ± 19.0	58.3 ± 20.2	0.40	0.10
Function	59.5 ± 18.8	59.2 ± 21.2	0.87	0.02
ADL	59.8 ± 23.5	60.3 ± 25.9	0.83	0.02
Work performance	59.7 ± 22.9	62.3 ± 27.6	0.22	0.14
Pain	35.2 ± 25.6	32.5 ± 26.0	0.25	0.13
Aesthetics	49.2 ± 24.8	51.3 ± 23.1	0.32	0.11
Satisfaction	48.3 ± 25.8	49.0 ± 26.2	0.77	0.03
AIMS2				
Physical	2.3 ± 1.8	2.4 ± 2.0	0.35	0.11
Affect	2.9 ± 1.8	2.7 ± 1.7	0.09	0.20
Symptom	4.2 ± 2.4	4.0 ± 2.6	0.40	0.10
Social	3.4 ± 1.3	3.7 ± 1.5	0.08	0.20

* MHQ = Michigan Hand Questionnaire; SRM = standardized response mean; SMPA = silicone metacarpophalangeal joint arthroplasty; ADL = activities of daily living; AIMS2 = Arthritis Impact Measurement Scales 2.
† By paired *t*-test comparing 6 months with baseline.
‡ Calculated using the following formula: (6-month followup mean − baseline mean)/SD of the change.

chronic hand conditions, including nerve compression, distal radius fractures, Dupuytren's disease, and osteoarthritis (15–20). The MHQ is ideal for use in the RA population because it specifically encompasses measures of aesthetics and pain control, which have been demonstrated to be important motivators for surgical therapy among patients with RA (36,37).

The MHQ demonstrated good test–retest reliability among the control patients, with minimal change in scores between survey administrations without statistically significant differences. Interestingly, patients who underwent SMPA reported significantly higher improved scores for the control, nonoperated hand across all of the MHQ domains with the exception of pain. The perioperative experience has a profound influence on a patient's perception of their disease and disability. Other authors have also reported similar phenomena among patients with other chronic health conditions such as paralysis and renal failure (38–40). These biases should be taken into consideration when using any health outcomes instrument, including the MHQ.

Overall, the 6 scales of the MHQ demonstrated excellent internal consistency, although item redundancy exists within the MHQ domains. This indicates that the MHQ may be well suited for item reduction in the future, which may improve response rates by shortening the time to complete the survey. The ADL, work performance, and pain scales of the MHQ were correlated in the expected direction with the AIMS2 instrument, particularly with respect to the physical domain. The function, ADL, satisfaction, and summary MHQ score were well correlated in the expected direction with objective measures of hand functioning, including grip and pinch strength and the Jebsen-Taylor Test score. Finally, the MHQ showed excellent responsiveness among patients who underwent SMPA arthroplasty, and it was able to detect change in hand performance over a 6-month time period. As expected, changes in hand performance as measured by the MHQ among patients who had not undergone surgery were modest over the 6-month followup period.

Other methods for assessing the extent of disability related to rheumatic hand disease include describing the extent of anatomic deformity using scoring algorithms based on clinical examination or radiographic evidence of joint destruction. For example, the Hand Index uses simple hand measurements of the span and lateral height of the open and closed hand to create a standardized measure of deformity (41). The Joint Alignment and Motion Scale score and the mechanical joint score can be used to define hand deformity and dysfunction at the bedside, but are subject to observer variation (42,43). Radiographic evidence of disease and disease progression has been defined by other instruments such as the Sharp score (44). These measures have the advantage of documenting the progression of disease over time, but do not adequately predict the clinical manifestations of disease such as pain, occupational disability, and the need for joint replacement (45). Although describing anatomic deformity may be helpful as part of a global assessment of disability, it is problematic when taken alone because many patients are able to retain excellent hand functioning despite deformity. Pain, joint

instability, and exercise tolerance are more predictive of physical functioning and general health perception among patients with RA than clinical or radiologic joint appearance (8,46).

To our knowledge, the MHQ is the first self-administered instrument validated in patients with RA that comprehensively gathers information on functional ability and the ability to complete daily and occupational activities, as well as patient satisfaction, pain, and aesthetic hand appearance. It is the only questionnaire validated in this population that can adjust for hand dominance and the differences in hand disability between both hands. The MHQ queries patients regarding their left and right hands separately in each performance domain. Furthermore, the MHQ includes a separate item regarding hand dominance (right-hand dominant, left-hand dominant, or both). Responses to this item can then be used to stratify patients by hand dominance for subsequent analyses, or to adjust for hand dominance in multivariate analyses. Several upper extremity–specific instruments have been used to measure hand function in patients with rheumatism. The Disabilities of the Arm, Shoulder, and Hand questionnaire has been used to study patients with rheumatism, but its validity and responsiveness have not been documented in this population (47). Furthermore, it does not make a distinction between right and left hand disability and focuses on the entire upper extremity, not specifically hand dysfunction. The Patient-Rated Wrist Evaluation Questionnaire has been used to study pain and function among RA patients, but focuses primarily on wrist, not hand, dysfunction, which frequently coexist in patients with rheumatism (48). The Cochin Rheumatoid Hand Disability scale has been developed to measure the effectiveness of surgery on rheumatoid hand functioning with respect to ADL, and has been shown to be valid in this population and sensitive to changes in disease state (49,50). It does not, however, include important aspects of the patient experience, including an assessment of pain, patient satisfaction, or aesthetics (50,51).

Our findings are consistent with previous, smaller studies regarding the use of the MHQ to describe hand disability among patients with RA. In comparison with the Australian/Canadian Osteoarthritis Hand Index and the Sequential Occupational Dexterity Assessment, the MHQ yields reproducible results and is uniquely suited to measure outcomes in patients with RA because it can discern disability in both hands (24). Although the MHQ was less responsive to clinical change among control patients because, expectedly, the control patients should not have changes in hand performance, it demonstrated excellent responsiveness among patients undergoing SMPA. These results are consistent with prior work using the MHQ to measure outcomes in Dutch patients with RA. In this study, the MHQ demonstrated excellent responsiveness to clinical change over time, particularly in the domains of patient satisfaction and hand aesthetics (22).

Our study has several notable limitations. First, criterion validity cannot be assessed because there is no previously accepted "gold standard" instrument for measuring hand function among patients with RA. Additionally, the progression of rheumatoid hand dysfunction may be too slow

to detect appreciable clinical change, and longer followup may be needed to understand the responsiveness to change of the MHQ in patients who have not undergone surgery. However, we were able to demonstrate excellent responsiveness among patients who underwent surgical intervention. Finally, our sample size prevented us from stratifying our results based on disease severity and effects of medical and occupational therapies, which may have influenced our results. Our current data are limited by a lack of information regarding disease characteristics and medical therapies, such as laboratory values (e.g., C-reactive protein level) or the use of corticosteroids. Future studies will include such variables and will allow us to evaluate the performance of the MHQ in the context of these factors.

Nonetheless, this study demonstrates that the MHQ is an essential instrument to understand the extent of disability of rheumatic hand disease. The MHQ offers clinicians a systematic approach to defining patient disability. Additionally, the MHQ can be incorporated into future studies regarding the effectiveness of RA therapies because it offers a comprehensive assessment of hand functioning and patient-centered outcomes. In conclusion, the MHQ is an easily administered, reliable, valid tool to measure rheumatoid hand function, and an essential instrument to systematically guide clinical decision making and assess the quality of care of rheumatic hand disease.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be submitted for publication. Dr. Chung had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Chung, Burke, Wilgis, Fox.
Acquisition of data. Chung, Burns, Burke, Wilgis.
Analysis and interpretation of data. Waljee, Chung, Kim, Burke, Wilgis.

REFERENCES

1. Helmick CG, Felson DT, Lawrence RC, Gabriel S, Hirsch R, Kwoh CK, et al, for the National Arthritis Data Work Group. Estimates of the prevalence of arthritis and other rheumatic conditions in the United States: part I. Arthritis Rheum 2008; 58:15–25.
2. O'Dell JR. Therapeutic strategies for rheumatoid arthritis. N Engl J Med 2004;350:2591–602.
3. Sherrer YS, Bloch DA, Mitchell DM, Young DY, Fries JF. The development of disability in rheumatoid arthritis. Arthritis Rheum 1986;29:494–500.
4. Ward MM, Javitz HS, Yelin EH. The direct cost of rheumatoid arthritis. Value Health 2000;3:243–52.
5. Dellhag B, Bjelle A. A grip ability test for use in rheumatology practice. J Rheumatol 1995;22:1559–65.
6. Jebsen RH, Taylor N, Trieschmann RB, Trotter MJ, Howard LA. An objective and standardized test of hand function. Arch Phys Med Rehabil 1969;50:311–9.
7. Backman C, Mackie H. Arthritis hand function test: inter-rater reliability among self-trained raters. Arthritis Care Res 1995; 8:10–5.
8. Pap G, Angst F, Herren D, Schwyzer HK, Simmen BR. Evaluation of wrist and hand handicap and postoperative outcome in rheumatoid arthritis. Hand Clin 2003;19:471–81.
9. Pincus T, Yazici Y, Bergman MJ. Patient questionnaires in

rheumatoid arthritis: advantages and limitations as a quantitative, standardized scientific medical history. Rheum Dis Clin North Am 2009;35:735–43.
10. Kalyoncu U, Dougados M, Daures JP, Gossec L. Reporting of patient-reported outcomes in recent trials in rheumatoid arthritis: a systematic literature review. Ann Rheum Dis 2009; 68:183–90.
11. Boers M, Felson DT. Clinical measures in rheumatoid arthritis: which are most useful in assessing patients? J Rheumatol 1994;21:1773–4.
12. Boers M, Tugwell P, Felson DT, van Riel PL, Kirwan JR, Edmonds JP, et al. World Health Organization and International League of Associations for Rheumatology core endpoints for symptom modifying antirheumatic drugs in rheumatoid arthritis clinical trials. J Rheumatol Suppl 1994;41: 86–9.
13. Mandl LA, Burke FD, Shaw Wilgis EF, Lyman S, Katz JN, Chung KC. Could preoperative preferences and expectations influence surgical decision making? Rheumatoid arthritis patients contemplating metacarpophalangeal joint arthroplasty. Plast Reconstr Surg 2008;121:175–80.
14. Cavaliere CM, Chung KC. Total wrist arthroplasty and total wrist arthrodesis in rheumatoid arthritis: a decision analysis from the hand surgeons' perspective. J Hand Surg Am 2008; 33:1744–55.
15. Chung KC, Squitieri L, Kim HM. Comparative outcomes study using the volar locking plating system for distal radius fractures in both young adults and adults older than 60 years. J Hand Surg Am 2008;33:809–19.
16. Chung KC, Ram AN, Shauver MJ. Outcomes of pyrolytic carbon arthroplasty for the proximal interphalangeal joint. Plast Reconstr Surg 2009;123:1521–32.
17. Sammer DM, Fuller DS, Kim HM, Chung KC. A comparative study of fragment-specific versus volar plate fixation of distal radius fractures. Plast Reconstr Surg 2008;122:1441–50.
18. Herweijer H, Dijkstra PU, Nicolai JP, van der Sluis CK. Postoperative hand therapy in Dupuytren's disease. Disabil Rehabil 2007;29:1736–41.
19. Klein RD, Kotsis SV, Chung KC. Open carpal tunnel release using a 1-centimeter incision: technique and outcomes for 104 patients. Plast Reconstr Surg 2003;111:1616–22.
20. Chang EY, Chung KC. Outcomes of trapeziectomy with a modified abductor pollicis longus suspension arthroplasty for the treatment of thumb carpometacarpal joint osteoarthritis. Plast Reconstr Surg 2008;122:505–15.
21. Adams J, Burridge J, Mullee M, Hammond A, Cooper C. The clinical effectiveness of static resting splints in early rheumatoid arthritis: a randomized controlled trial. Rheumatology (Oxford) 2008;47:1548–53.
22. Van der Giesen FJ, Nelissen RG, Arendzen JH, de Jong Z, Wolterbeek R, Vliet Vlieland TP. Responsiveness of the Michigan Hand Outcomes Questionnaire Dutch language version in patients with rheumatoid arthritis. Arch Phys Med Rehabil 2008;89:1121–6.
23. Mandl LA, Galvin DH, Bosch JP, George CC, Simmons BP, Axt TS, et al. Metacarpophalangeal arthroplasty in rheumatoid arthritis: what determines satisfaction with surgery? J Rheumatol 2002;29:2488–91.
24. Massy-Westropp N, Krishnan J, Ahern M. Comparing the AUSCAN Osteoarthritis Hand Index, Michigan Hand Outcomes Questionnaire, and Sequential Occupational Dexterity Assessment for patients with rheumatoid arthritis. J Rheumatol 2004;31:1996–2001.
25. Goldfarb CA, Stern PJ. Metacarpophalangeal joint arthroplasty in rheumatoid arthritis: a long-term assessment. J Bone Joint Surg Am 2003;85A:1869–78.
26. Chung KC, Kotsis SV, Kim HM. A prospective outcomes study of Swanson metacarpophalangeal joint arthroplasty for the rheumatoid hand. J Hand Surg Am 2004;29:646–53.
27. Chung KC, Burns PB, Wilgis EF, Burke FD, Regan M, Kim HM, et al. A multicenter clinical trial in rheumatoid arthritis comparing silicone metacarpophalangeal joint arthroplasty with medical treatment. J Hand Surg Am 2009;34:815–23.
28. Meenan RF, Mason JH, Anderson JJ, Guccione AA, Kazis LE.

AIMS2: the content and properties of a revised and expanded Arthritis Impact Measurement Scales health status questionnaire. Arthritis Rheum 1992;35:1–10.

29. Sharma S, Schumacher HR Jr, McLellan AT. Evaluation of the Jebsen hand function test for use in patients with rheumatoid arthritis [corrected]. Arthritis Care Res 1994;7:16–9.

30. Trochim W, Donnelly JP. The research methods knowledge base. 2nd ed. Cincinnati (Ohio): Atomic Dog; 2007.

31. Streiner DL, Norman GR. Health measurement scales: a practical guide to their development and use. New York: Oxford University Press; 1989.

32. Chung KC, Pillsbury MS, Walters MR, Hayward RA. Reliability and validity testing of the Michigan Hand Outcomes Questionnaire. J Hand Surg Am 1998;23:575–87.

33. Chung KC, Hamill JB, Walters MR, Hayward RA. The Michigan Hand Outcomes Questionnaire (MHQ): assessment of responsiveness to clinical change. Ann Plast Surg 1999;42:619–22.

34. Liang MH, Lew RA, Stucki G, Fortin PR, Daltroy L. Measuring clinically important changes with patient-oriented questionnaires. Med Care 2002;40:II45–51.

35. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale (NJ): Lawrence Erlbaum Associates; 1988.

36. Alderman AK, Arora AS, Kuhn L, Wei Y, Chung KC. An analysis of women's and men's surgical priorities and willingness to have rheumatoid hand surgery. J Hand Surg Am 2006;31:1447–53.

37. Chung KC, Kotsis SV, Kim HM, Burke FD, Wilgis EF. Reasons why rheumatoid arthritis patients seek surgical treatment for hand deformities. J Hand Surg Am 2006;31:289–94.

38. Smith D, Loewenstein G, Jepson C, Jankovich A, Feldman H, Ubel P. Mispredicting and misremembering: patients with renal failure overestimate improvements in quality of life after a kidney transplant. Health Psychol 2008;27:653–8.

39. Damschroder LJ, Zikmund-Fisher BJ, Ubel PA. The impact of considering adaptation in health state valuation. Soc Sci Med 2005;61:267–77.

40. Ubel PA, Loewenstein G, Schwarz N, Smith D. Misimagining the unimaginable: the disability paradox and health care decision making. Health Psychol 2005;24:S57–62.

41. Highton J, Markham V, Doyle TC, Davidson PL. Clinical characteristics of an anatomical hand index measured in patients with rheumatoid arthritis as a potential outcome measure. Rheumatology (Oxford) 2005;44:651–5.

42. Johnson AH, Hassell AB, Jones PW, Mattey DL, Saklatvala J, Dawes PT. The mechanical joint score: a new clinical index of joint damage in rheumatoid arthritis. Rheumatology (Oxford) 2002;41:189–95.

43. Spiegel TM, Spiegel JS, Paulus HE. The Joint Alignment and Motion Scale: a simple measure of joint deformity in patients with rheumatoid arthritis. J Rheumatol 1987;14:887–92.

44. Sharp JT. Scoring radiographic abnormalities in rheumatoid arthritis. J Rheumatol 1989;16:568–9.

45. Yazici Y, Sokka T, Pincus T. Radiographic measures to assess patients with rheumatoid arthritis: advantages and limitations. Rheum Dis Clin North Am 2009;35:723–9.

46. Eurenius E, Brodin N, Lindblad S, Opava CH. Predicting physical activity and general health perception among patients with rheumatoid arthritis. J Rheumatol 2007;34:10–5.

47. Beaton DE, Katz JN, Fossel AH, Wright JG, Tarasuk V, Bombardier C. Measuring the whole or the parts? Validity, reliability, and responsiveness of the Disabilities of the Arm, Shoulder and Hand outcome measure in different regions of the upper extremity. J Hand Ther 2001;14:128–46.

48. MacDermid JC, Turgeon T, Richards RS, Beadle M, Roth JH. Patient rating of wrist pain and disability: a reliable and valid measurement tool. J Orthop Trauma 1998;12:577–86.

49. Duruoz MT, Poiraudeau S, Fermanian J, Menkes CJ, Amor B, et al. Development and validation of a rheumatoid hand functional disability scale that assesses functional handicap. J Rheumatol 1996;23:1167–72.

50. Poiraudeau S, Lefevre-Colau MM, Fermanian J, Revel M. The ability of the Cochin rheumatoid arthritis hand functional scale to detect change during the course of disease. Arthritis Care Res 2000;13:296–303.

51. Lefevre-Colau MM, Poiraudeau S, Fermanian J, Etchepare F, Alnot JY, Le Viet D, et al. Responsiveness of the Cochin rheumatoid hand disability scale after surgery. Rheumatology (Oxford) 2001;40:843–50.