



Published in final edited form as:

Plast Reconstr Surg. 2011 July ; 128(1): 208–220. doi:10.1097/PRS.0b013e318218fc51.

Development of a Brief, 12-item Version of the Michigan Hand Questionnaire

Jennifer F. Waljee, MD, MPH, MS¹, H. Myra Kim, ScD², Patricia B. Burns, MPH³, and Kevin C. Chung, MD, MS⁴

¹ House Officer, Section of Plastic Surgery, Department of Surgery, The University of Michigan Health System

² Associate Research Scientist, Center for Statistical Consultation and Research, University of Michigan

³ Research Associate, Section of Plastic Surgery, Department of Surgery, The University of Michigan Medical Center

⁴ Professor of Surgery, Section of Plastic Surgery, Department of Surgery, The University of Michigan Health System

Abstract

Background—The Michigan Hand Questionnaire (MHQ) is one of the most widely utilized, hand-specific surveys that measures health status relevant to patients with acute and chronic hand disorders. However, item redundancy exists in the original version, and an abbreviated survey could minimize responder burden and offer broader applicability.

Methods—Patients (n=422) with 4 specific hand conditions completed the MHQ at two time periods: rheumatoid arthritis (RA) (n=162), thumb carpometacarpal osteoarthritis (CMC) (n=31), carpal tunnel syndrome (CTS) (n=97), and distal radius fracture (DRF) (n=132). Correlation analysis identified 2 items from each of 6 domains (function, activities of daily living, work, pain, aesthetics, and satisfaction). A brief-MHQ score was calculated as the sum of the responses to the 12 items. Psychometric analysis was performed to describe the reliability, validity, and responsiveness of the brief MHQ.

Results—The brief-MHQ includes 12 items that were highly correlated with the summary MHQ score ($r = 0.99$, $p < 0.001$). The brief MHQ scores were highly correlated between the two time periods ($r = 0.78$, $p < 0.001$), and by disease type. Responsiveness of the brief MHQ was high for all diseases, and similar to that of the original MHQ.

Conclusions—The 12-item brief MHQ is an efficient and versatile outcomes instrument specific to hand disability that retains the psychometric properties of the original MHQ. The brief MHQ is an important tool to measure patient outcomes and the quality of care in hand surgery.

Keywords

Brief-MHQ; Outcomes; Hand Surgery; Item Reduction

Corresponding Author: Kevin C. Chung, MD, MS, Section of Plastic Surgery, University of Michigan Health System, 2130 Taubman Center, SPC 5340, 1500 E. Medical Center Drive, Ann Arbor, MI, 48109-5340, kecchung@umich.edu, Phone 734-936-5885, Fax 734-763-5354.

Nothing to Declare.

In the United States, health care quality is variable, difficult to measure, and not well correlated with health care expenditure. In response, health services research has focused on identifying ways to measure health care quality efficiently and accurately.(1) Efforts to measure health care quality range widely, from examining 30-day mortality or hospital readmission rates, to implementing re-certification standards for physicians in practice. However, hand disability presents a unique challenge for health services researchers in that traditional indicators, such as mortality or perioperative processes of care, are not sensitive enough to distinguish variations in patient outcomes. (2) However, hand function has a profound effect on a patient's quality of life. Therefore, measures of patient satisfaction and disability are essential quality indicators for clinicians, researchers, and health care policy-makers interested in the quality of care of upper extremity disability.

Over 25 years ago, the Socioeconomic Committee of the American Society for Surgery of the Hand advocated the development of a comprehensive instrument through rigorous methodology to measure patient outcomes following hand surgery.(3) In 1998, the Michigan Hand Outcomes (MHQ) was introduced as a hand-specific outcomes instrument that can effectively measure aspects of health status that are relevant to patients with acute and chronic hand disorders. (4) Today, the MHQ is a widely utilized, self-administered instruments to measure upper extremity disability and patient outcomes across a broad range of hand-related diseases. The MHQ has been translated into 8 different languages and used worldwide for a range of acute and chronic hand conditions including, trauma, nerve compression syndromes, arthritis, and inflammatory conditions. (5–19) The MHQ comprehensively gathers information on hand function, the ability to complete daily and occupational activities, patient satisfaction, pain, and aesthetic hand appearance. It is the only questionnaire currently in use that adjusts for hand dominance, and can distinguish the differences in disability between both hands. (20)

The MHQ was developed based on strict psychometric testing and has been extensively validated in field studies in the US and around the world. In its current form, the MHQ includes 37 items, and takes approximately 15 minutes to complete. Although it has been widely used for a variety of hand conditions, previous reliability studies have demonstrated redundancy in the items included in the original survey, which limits its application for larger studies.(4, 21) Studies of reliability measures, such as Cronbach's alpha values, have demonstrated that higher values of Cronbach's alpha greater than 0.9 indicate item redundancy and multiple dimensions in a given scale rather than the true reliability of a scale, and should be avoided.(22, 23) A shortened version of the MHQ is a more attractive research instrument for population studies because it is more time-efficient, reduces responder burden, and can therefore minimize missing data. Previous research demonstrates that longer surveys yield lower response rates and poorer data quality or incomplete data. (24, 25) While there is no defined optimal number of survey items, clearly longer surveys increase responder burden and the cost of survey production, distribution and coding. (24, 26–28)

It is challenging to shorten a survey while retaining its reliability, validity and sensitivity. Shortened versions of the SF-36 and DASH have been developed and tested.(24, 29, 30) Using the experience from these shortened questionnaires, we created an abbreviated version of the MHQ based on rigorous methodology in order to preserve the essential properties of the original instrument. To develop a brief version of the MHQ, we used data gathered prospectively from patients with four distinct hand conditions: distal radius fractures, carpal tunnel syndrome, rheumatoid arthritis, and thumb carpometacarpal (CMC) arthritis. We hypothesize that the shortened version of the MHQ will be as valid and reliable as the original version, and yet more practical for studying health care quality in multicenter or nationwide studies.

Methods

Study Sample

We gathered data prospectively from 422 patients with four specific hand conditions: rheumatoid arthritis, thumb CMC osteoarthritis, carpal tunnel syndrome, and distal radius fracture. With the exception of patients with distal radius fractures, patients for the elective procedures were evaluated preoperatively and at a designated time period postoperatively.

In this dataset, 162 patients with rheumatoid arthritis (RA) were included in a larger, prospective study supported by the National Institutes of Health regarding the use of silicone metacarpophalangeal arthroplasty (SMPA) for joint deformities due to RA. The details of this study and the data collection have been described elsewhere.(31, 32) RA patients completed the MHQ at the time of enrollment and at 6 months follow-up. Additionally, 97 patients treated at the University of Michigan for carpal tunnel syndrome (CTS) completed the MHQ as part of a study designed to evaluate the complications associated with minimal-incision carpal tunnel release. Patients were included in the study based on a diagnosis of carpal tunnel syndrome from clinical presentation (history of hand dyesthesia along the distribution of the median nerve and a positive Phalen's flexion test finding and/or a positive Tinel's sign) and electrodiagnostic study confirmation of CTS. Patient completed the MHQ prior to carpal tunnel release, and at 6 months postoperatively.(18, 33) Data were also collected from 132 distal radius fracture (DRF) patients who underwent operative fixation using the volar locking plating system (VLPS). As it was not possible to survey patients prior to their injury, responses taken at 3 months following surgery were used as baseline, and responses taken at 6 months following surgery were used as follow-up. (34) Finally, 31 patients with thumb carpometacarpal (CMC) arthritis completed the MHQ as part of a larger study to determine patient outcomes following trapeziectomy and modified abductor pollicis longus suspension arthroplasty. (12) Patients completed the MHQ at their preoperative clinic visit and again at 3 months postoperatively. All study protocols were approved by the institutional review boards at the University of Michigan.

Additionally, the DRF, RA, and CMC arthritis patients were assessed at baseline and at each clinic visit with objective measures of functioning, specifically grip strength in kilograms, key or lateral pinch strength in kilograms, and by Jebsen-Taylor Test score. (35)The Jebsen-Taylor test is a battery of tasks for which the subject's speed of completion is calculated, and consists of the following components: (1) writing a short sentence; (2) turning over 3-by 5-in. cards; (3) picking up small objects and placing them in a container; (4) stacking checkers; (5) simulated eating; (6) moving large, empty cans; and (7) moving large, weighted cans. The time required in seconds to complete each task was recorded for the subjects' dominant and nondominant hands. A subset of patients completed the Arthritis Impact Measurement Scales 2 questionnaire, a 45- item, self-administered outcomes tool designed to assess health status in patients with inflammatory arthritis and osteoarthritis.(36) 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.

The MHQ

All subjects completed the original, 37-item version of the MHQ (see Appendix 1). Each item on the MHQ allows the subject to respond to the question along a Likert like scale ranging from 1 to 5. Responses are then summed to yield a domain score for each of the 6 domains, and respondents must respond to 50% or more of the items within the scale in order to consider their responses as sufficient to generate a domain score. Each domain scores are transformed to range from 0 to 100, and higher scores indicate better performance

for all scales, with the exception of the pain scale. The MHQ also yields a summary score, which is calculated by averaging the scores for each domain, after reversing the pain score.

Item Reduction

Item reduction was performed using a concept-retention technique in order to identify those items of the original MHQ that would be retained in the brief version. (29) Briefly, this technique allows the selection of items based on their clinical relevance rather than on statistical testing alone. This technique is ideal in that it allows investigators to retain those items that were clinically relevant, regardless of their statistical value. For example, with the pain scale, both “Describe the pain in your hand(s)/wrist(s)” and “How often did the pain in your hand(s)/wrists(s) make you unhappy” both had similar correlation coefficients at 0.18. However, all authors reviewed these items to choose the most clinically relevant item, which was included in the final instrument. Twelve items were selected, with two items chosen from each scale. This allows for each scale from the original survey to be represented in the brief version. The two items from each scale that were most correlated with the original MHQ score were retained. Items were selected from each domain by the authors with the goal to retain the all concepts of the original framework of the MHQ (Figure 1). We then confirmed those items that were retained by determining the correlation of each item of the MHQ with the summary MHQ score using correlation analysis. Because the original MHQ requires subjects to respond to items separately for the right and left hand, except for items in work domain, responses of the two hands were averaged for each item prior to item reduction for all patients except those with distal radius fractures. The design of the brief-MHQ will not distinguish between laterality of hand symptoms to make the brief-MHQ as economical as possible in content and response time.

MHQ and brief MHQ scoring algorithm

We generated a summary score from the items in the brief-MHQ by averaging the responses of the final 12 items. The minimum possible raw summary score for the brief-MHQ score is 1 and the maximum score is 5. The brief-MHQ score was normalized to a scale of 0 to 100 that is similar to the original MHQ. Summary scores were not calculated if any of the items along the brief MHQ were missing. The brevity of the brief-MHQ requires complete response to all 12 items. Therefore, brief MHQ scores were not calculated if any data were missing in the included items. For the full MHQ, scores were only calculated if at least 3 out of the 6 scales were complete. For the function, ADL for individual hands, pain and aesthetic scales, respondents must have no more than 2 missing values to calculate scores for these scales. The satisfaction and ADL using both hands scales required 3 or fewer items missing. The final survey and scoring algorithm is included in Appendix 2.

Psychometric testing of the Brief-MHQ

Reliability of the Brief-MHQ—Reliability is the ability of a survey instrument to precisely measure a concept. (37) We measured the reliability of the brief- MHQ among RA patients who did not undergo SMPA because their responses would be expected to remain similar over a relatively short period. The brief-MHQ summary score and the original MHQ summary score were compared between baseline and at 6 month follow-up. The degree of correlation between the measurements of the two time periods was obtained using Spearman’s correlation coefficients. Paired t-tests were used to determine the average difference in the summary score between these two time periods. A mean difference of 0 indicates perfect test-retest reliability.

As an additional measure of reliability, we calculated the values of the intraclass correlation coefficient (ICC), which is an additional measure of agreement of estimates between paired

data. The ideal values for the ICC should range between 1 (perfect correlation) and $-1/(k-1)$ (low correlation) where k is the number of subjects. (38)

Validity of the Brief-MHQ—Validity is defined as the ability of a survey instrument to accurately measure a concept of interest. Important types of validity to consider in survey development include content validity, criterion validity, and construct validity. Content validity refers to the extent to which the instrument appears capable of measuring the desired outcome. The original MHQ was developed with strict attention to psychometric principles, and has been validated in a variety of acute and chronic hand conditions, and translated into several languages. Therefore, we would expect that the brief-MHQ would also have similar content validity.(12, 14, 15, 39, 40) Criterion validity describes the extent to which a survey instrument compares with the accepted reference standard. Currently, there is no established standard by which to measure health outcomes related to hand dysfunction across a wide range of acute and chronic hand conditions. Therefore, we rely on construct validity to ensure that the brief-MHQ accurately measures hand disability and function. Construct validity describes the extent to which survey responses or scores correspond to expected values. If the brief MHQ is valid, we hypothesize that the brief MHQ scores and the original MHQ scores should be similar within disease groups. To analyze this, we used multiple linear regression to compare the mean brief-MHQ scores, adjusted for clinical and demographic characteristics, by disease type. Furthermore, we compared the brief and full MHQ scores with objective measures of hand function (grip and pinch strength, and Jebsen-Taylor test score) and subjective measures of hand function (AIMS2 scores) as an additional measure of validity.

Responsiveness of the Brief-MHQ—Responsiveness is the ability of a survey instrument to detect changes in an outcome of interest over time.(41) To determine the responsiveness of the brief-MHQ, we compared the summary brief-MHQ score at baseline and at follow-up after a surgical intervention for each of the disease groups using paired t -tests. In order to compare the change in brief-MHQ scores over time in a standardized fashion, we calculated the standardized response mean (SRM), calculated as the change in means at follow-up time from baseline divided by the standard deviation of the change-scores, for each disease types. We expect that a sensitive survey instrument should have a correspondingly high SRM.(42) Using Cohen's effect size definition, we can interpret an SRM of 0.2 as a small effect size, 0.5 as a medium effect size, and 0.8 as a large effect size. (43) As an additional point of comparison, the responsiveness of the brief MHQ was compared with that of the original MHQ, grip strength, key pinch strength, and Jebsen-Taylor testfor those patients who completed these measures.

Statistical Analysis

Descriptive statistics were used to describe the characteristics of the study sample. Linear regression was used to identify the survey items to be retained in the final version of the brief-MHQ, and to describe the correlation of these items to the summary score of the original MHQ. Reliability testing was performed using Spearman's correlation coefficients and paired t -tests as described above. Validity testing was performed by using linear regression to determine the brief-MHQ summary score for each disease type, adjusted for age, gender, and education level. Responsiveness of the brief-MHQ was determined using paired t -tests and calculating the SRM as described above. Multivariate linear regression was used to generate SRM measures controlling for disease type. Statistical significance was set at an alpha level of 0.05. All analyses were conducted using Stata 10.1. (Statacorps, College Station, Texas).

Results

The demographic characteristics of the study sample are detailed in Table 1. This study sample consisted of 422 patients, of whom 132 patients had suffered distal radius fractures, 97 patients had CTS, 162 patients had RA, and 31 patients with CMC arthritis. The average age of the patient sample was approximately 55 years, and 69.7% were women. Approximately 33% had a college education or higher.

The results of the item reduction of the MHQ are described in Table 2. In all of the domains, except satisfaction, the two survey items for each domain that were most strongly associated with the summary MHQ score were retained in the final version of the brief-MHQ for a total of 12 items. Within the function domain, the following two items “*Overall, how well did your hand work?*” ($R^2=0.41$) and “*How was the sensation in your hand?*” ($R^2=0.21$) were most correlated with the summary MHQ, and were retained in the brief-MHQ survey. Within the ADL domain, two items “*How difficult was it for you to hold a frying pan?*” ($R^2=0.16$) and “*How difficult was it for you to button a shirt/blouse?*” ($R^2=0.17$) were most strongly correlated with the summary MHQ score, and were retained in the brief-MHQ. The items within the work domain that were most strongly correlated with the summary MHQ score included “*How often were you unable to do your work in the last week because of your hands/wrists?*” ($R^2=0.25$) and “*How often did you take longer to do tasks in your work because of problems with your hands/wrists?*” ($R^2=0.34$). The items within the pain domain that were selected for the brief-MHQ included “*Describe the pain in your hands/wrists*” ($R^2=0.18$) and “*How often did the pain in your hands/wrists interfere with your daily activities?*” ($R^2=0.39$). Within the aesthetic domain, two items “*I am satisfied with the look of my hands*” ($R^2=0.33$) and “*The appearance of my hand interferes with my normal daily activities*” ($R^2=0.32$) were most strongly correlated with the summary MHQ score, and were retained in the brief-MHQ. Finally, the two items related to satisfaction that were retained included “*Satisfaction with the motion of your fingers*” ($R^2=0.14$) and “*Satisfaction with the motion of your wrist*” ($R^2=0.19$). These items were selected for inclusion in the brief-MHQ over others with higher correlation coefficients because these items contained concepts that were not represented in other portions of the survey. After these 12 items were selected for inclusion in the final brief-MHQ survey, regression analysis revealed that these survey items explained 97.8% of the variance of the original summary MHQ scores.

Table 3 details the reliability of the brief-MHQ. The reliability of the brief MHQ was examined by measuring the test-retest correlation over a 6 month period among a subset of 68 rheumatoid arthritis patients who did not undergo surgical intervention and had both baseline and 6 months measurements available. The correlation and ICC values between the brief-MHQ scores between each time period was high ($r=0.78$, $r_T=0.91$), and the mean difference between the two scores was not statistically significant (0.22, $p=0.87$). Similar findings were noted for the original MHQ scores. This indicates excellent test-retest reliability of the brief MHQ in this subset of patients.

Figure 2 shows the adjusted means of the brief-MHQ summary score and the original MHQ summary score, stratified by disease type and adjusted for age, gender, and education. We hypothesize that the relative hand health status by the different disease types shown using the original MHQ will also be shown using the brief-MHQ scores. We observed that patients with DRF had the highest summary MHQ score (77.8 ± 1.60), and patients with RA had the lowest summary MHQ scores (51.7 ± 1.38). Similarly, patients with DRF had the highest brief MHQ score (77.8 ± 1.42) and patient with RA had the lowest brief MHQ scores (47.6 ± 1.34). Of note, DRF patients were surveyed postoperatively, and have significantly better hand outcomes than each of the other three conditions.

Table 4 details the responsiveness of the brief-MHQ to clinical change among patients who underwent surgical intervention by disease type. Overall, the responsiveness of the brief MHQ was high for all disease types, even in the DRF patients whose measurements reflect the improvements between 3 and 6 months post VLPS, and similar to that of the original MHQ. Responsiveness was highest among RA patients who have undergone SMPA, and was similar for the brief-MHQ (SRM=1.28) and the original MHQ (SRM=1.36).

Table 5 compares the changes in scores between the two time periods and responsiveness to clinical change for the brief MHQ, the original MHQ, grip and pinch strength, and Jebsen-Taylor test score. All analyses were adjusted for disease type, and CTS patients were admitted from this portion of the analysis as they did not complete these parameters. Responsiveness to clinical change was highest for the brief MHQ (SRM=4.15), followed by the original MHQ (SRM=3.30), Jebsen-Taylor test score (SRM=1.26), and pinch strength (SRM=1.60). Responsiveness was lowest for grip strength (SRM=0.36).

Figure 3 describes the correlation between the brief MHQ and objective measures of hand function, specifically grip and pinch strength, and Jebsen-Taylor test score. The brief MHQ was moderately correlated with each of these measures, after adjusting for disease type, age, gender, and education level. The correlation values were highest for grip strength ($r=0.38$), followed by pinch strength ($r=0.35$), and Jebsen-Taylor test ($r=0.35$). Similar trends were identified in the correlation of the full MHQ with these parameters.

Figure 4 describes the correlation between the brief MHQ and patient-reported hand function among a subset of the patient sample. Rheumatoid arthritis patients completed the AIMS2 survey, and responses to each domain were compared with brief MHQ score, after controlling for age, gender, and education level. The brief MHQ score was most highly correlated with function ($r=0.68$) and least correlated with social interaction ($r=0.15$). Brief MHQ score was moderately correlated with affect ($r=0.36$) and disease symptoms ($r=0.53$). Similar trends were identified in the correlation of the full MHQ with each of these domains.

Discussion

Health services research relies on the ability to accurately and reliably measure patient outcomes to describe variations in health care quality. To describe patient outcomes related to upper extremity disability and surgery, surgeons need on patient-reported data regarding hand function, pain, and satisfaction that cannot be captured by objective testing. (44) The MHQ is popular because it comprehensively measures multiple aspects of hand disability and adjusts for hand dominance and injury. It includes measures of satisfaction and aesthetic appearances, which are important for patients with chronic and deforming hand conditions such as osteoarthritis and rheumatoid arthritis. (29) (45) (46–48) In this study, we have created a shortened, 12-item version of the original MHQ by eliminating certain items and selecting those items with the greatest correlation with the original MHQ. The brief MHQ was reliable on repeated testing, well correlated with the original version, and can detect clinical change among patients who undergo surgery. It was correlated with objective measures of hand function, including Jebsen-Taylor test score and pinch strength, and subjective measures of hand function, including domains of the AIMS2 instrument.

Previous research has demonstrated that longer surveys are associated with lower response rates.(49) For example, Jepson et. al. conducted a study of physicians to determine the response rate with varying of lengths of surveys. In this study, response rates declined sharply with increasing length. (50) However, reducing the number of survey items can also lead to reduction in the comprehensiveness and precision of the data that can be collected. (51) The decisions regarding these tradeoffs must be weighed against the benefit of a shorter

survey, and the ultimate purpose of the instrument. Often, scales that were previously present in the original instrument are unable to be maintained in the shorter instrument. Furthermore, homogeneity of items in measuring a distinct concept is traded for heterogeneity of items, in order to capture data. For example, when shortening the SF-36 down to abbreviated versions such as the SF-12 and SF8, representative items from scales were able to be maintained.(24, 25) However, the selected items were heterogeneous and unique, but with a reliable variance and able to predict the outcome of interest. Therefore, the concepts measured by shorter surveys are assessed with less precision and greater variance. Regardless, these differences in precision are not likely to be detrimental when used in large studies with greater than 500 subjects, as confidence intervals of estimates are primarily related to sample size. Therefore, it is important to take into account the purpose of the instrument and the scale of the study when developing and implementing shorter survey instruments in order to account for these tradeoffs.(25, 52)

Scientific investigation relies on the ability to accurately and efficiently measure phenomena in a given population. An ideal measure should measure concepts consistently with as few items as possible. Achieving this ideal is challenging in areas where concepts such as pain and satisfaction are ambiguous and dynamic. To create a new version of the MHQ, we chose to use the concept-retention technique to identify two items most highly correlated with the original MHQ score to include in the final version. This approach has been used successfully for item reduction of other instruments, such as the Disabilities of the Arm, Shoulder, and Hand (DASH) instrument. (29) This methodology has been criticized previously in that it is more subjective compared with other psychometric approaches, such as the Rasch or equidiscriminative techniques. Input from clinicians and patients in survey development can potentially lead to more heterogeneous scales in survey development. However, clinician input is often correlated with greater validity, and is not limited by specific mathematical or statistical parameters, such as item difficulty, which may exclude important items that are clinically relevant. (53, 54)

While the original MHQ remains a robust tool for comprehensive analysis, the brief MHQ will be an important adjunct to the original instrument for collecting data on a large scale. The responsiveness of the brief MHQ in this analysis indicates that it will detect clinical change with similar sensitivity as the full MHQ. For example, for studies with smaller sample size, the comprehensive questions in the 6 scales of the original MQH can minimize “noise” inherent in blunter survey instruments. Conversely, in larger studies, the brief-MHQ is ideal to reduce responder burden and maximize response rates. This shortened instrument will allow the domains of the MHQ to be applied more efficiently and with greater versatility across a broad range of hand diseases. The brief-MHQ will likely excel as a screening tool, or to capture a cross-sectional snapshot of patient outcomes in larger-scale audits of practice outcomes. Surgeons could use the brief-MHQ to document their practice outcomes and follow these longitudinally as a way to compare care against normative values and identify areas for attention or improvement. The brief-MHQ will offer greater efficiency than the original MHQ for this purpose by reducing responder burden and eliminating the need for complicated scoring algorithms.

This study has several notable limitations. First, the brief-MHQ was developed in a group of patients with four specific hand diseases, and its applicability for other conditions requires testing in large field studies. However, the diseases included in this prospective dataset represent both acute and chronic disorders, and include patients with degenerative, inflammatory, and nerve compression syndromes. Additionally, we performed validity testing by comparing our abbreviated instrument against the original instrument. However, there is no existing “gold standard” survey instrument for hand disability for comparison. Objective measures, such as grip and pinch strength, may not correlate with other important

aspects of hand dysfunction, such as aesthetic deformity.(20, 55–57) Furthermore, the shortened version of this instrument does not account for laterality, and does not retain the detail of the 6 domains of the original version. Finally, our analysis was conducted retrospectively on a subset of patients in our dataset. This work captures our experience in the initial development of the brief MHQ, and details our efforts to perform item reduction of the MHQ and define psychometric properties of the resulting instrument. Rigorous survey development and item reduction demands a sequential process including systematic review of existing knowledge, qualitative and quantitative preliminary testing with eventual expansion to larger multicenter studies for survey refinement. The brief MHQ will be further refined through a larger, prospective study to describe the logistics of surgery administration and applicability to other conditions.

The brief, 12-item version of the MHQ has the potential for many applications for surgeons in practice, researchers, and policy-makers interested in improving the delivery of care to patients with hand disability. This shortened version has similar psychometric properties as the original MHQ, and demonstrates excellent reliability and validity across a variety of acute and chronic hand conditions. The brief-MHQ will be an essential tool for measuring the quality of care and influencing practice patterns to optimize the practice of hand surgery.

Acknowledgments

Supported in part by a grant from the National Institute of Arthritis and Musculoskeletal and Skin Diseases (R01 AR047328) and a Midcareer Investigator Award in Patient-Oriented Research (K24 AR053120) (to Dr. Kevin C. Chung).

References

1. DeLisa JA. Maintenance of certification: continuing assessment of physician quality with respect to their commitment to quality patient care, lifelong learning, ongoing self-assessment, and improvement. *Am J Phys Med Rehabil.* 2009 Sep; 88(9):775–9. [PubMed: 19661777]
2. Chung KC, Shauver MJ. Measuring quality in health care and its implications for pay-for-performance initiatives. *Hand Clin.* 2009 Feb; 25(1):71–81. vii. [PubMed: 19232918]
3. Chase RA. Costs, risks, and benefits of hand surgery. *J Hand Surg Am.* 1983 Sep; 8(5 Pt 2):644–8. [PubMed: 6415154]
4. Chung KC, Pillsbury MS, Walters MR, Hayward RA. Reliability and validity testing of the Michigan Hand Outcomes Questionnaire. *J Hand Surg Am.* 1998 Jul; 23(4):575–87. [PubMed: 9708370]
5. Weber RA, DeSalvo DJ, Rude MJ. Five-year follow-up of carpal tunnel release in patients over age 65. *J Hand Surg Am.* Feb; 35(2):207–11. [PubMed: 20060236]
6. Dziedzic K, Thomas E, Hill S, Wilkie R, Peat G, Croft PR. The impact of musculoskeletal hand problems in older adults: findings from the North Staffordshire Osteoarthritis Project (NorStOP). *Rheumatology (Oxford).* 2007 Jun; 46(6):963–7. [PubMed: 17329350]
7. Pike J, Mulpuri K, Metzger M, Ng G, Wells N, Goetz T. Blinded, prospective, randomized clinical trial comparing volar, dorsal, and custom thermoplastic splinting in treatment of acute mallet finger. *J Hand Surg Am.* Apr; 35(4):580–8. [PubMed: 20353859]
8. Deshmukh SC, Kumar D, Mathur K, Thomas B. Complex fracture-dislocation of the proximal interphalangeal joint of the hand. Results of a modified pins and rubbers traction system. *J Bone Joint Surg Br.* 2004 Apr; 86(3):406–12. [PubMed: 15125130]
9. Johnston P, Larson D, Clark IM, Chojnowski AJ. Metalloproteinase gene expression correlates with clinical outcome in Dupuytren's disease. *J Hand Surg Am.* 2008 Sep; 33(7):1160–7. [PubMed: 18762113]
10. Sambandam SN, Priyanka P, Gul A, Ilango B. Critical analysis of outcome measures used in the assessment of carpal tunnel syndrome. *Int Orthop.* 2008 Aug; 32(4):497–504. [PubMed: 17370071]

11. Theivendran K, Pollock J, Rajaratnam V. Proximal interphalangeal joint fractures of the hand: treatment with an external dynamic traction device. *Ann Plast Surg.* 2007 Jun; 58(6):625–9. [PubMed: 17522484]
12. 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 Aug; 122(2):505–15. [PubMed: 18626369]
13. Chung KC, Burke FD, Wilgis EF, Regan M, Kim HM, Fox DA. A prospective study comparing outcomes after reconstruction in rheumatoid arthritis patients with severe ulnar drift deformities. *Plast Reconstr Surg.* 2009 Jun; 123(6):1769–77. [PubMed: 19483578]
14. Herweijer H, Dijkstra PU, Nicolai JP, Van der Sluis CK. Postoperative hand therapy in Dupuytren's disease. *Disabil Rehabil.* 2007 Nov 30; 29(22):1736–41. [PubMed: 17852228]
15. Chung KC, Ram AN, Shauver MJ. Outcomes of pyrolytic carbon arthroplasty for the proximal interphalangeal joint. *Plast Reconstr Surg.* 2009 May; 123(5):1521–32. [PubMed: 19407625]
16. Chung KC, Kotsis SV, Kim HM. Predictors of functional outcomes after surgical treatment of distal radius fractures. *J Hand Surg Am.* 2007 Jan; 32(1):76–83. [PubMed: 17218179]
17. McMillan CR, Binhammer PA. Which Outcome Measure is the Best? Evaluating Responsiveness of the Disabilities of the Arm, Shoulder, and Hand Questionnaire, the Michigan Hand Questionnaire and the Patient-Specific Functional Scale Following Hand and Wrist Surgery. *Hand (N Y).* 2009 Sep; 4(3):311–8. [PubMed: 19259747]
18. 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 Apr 15; 111(5):1616–22. [PubMed: 12655206]
19. Chung KC, Watt AJ, Kotsis SV. A prospective outcomes study of four-corner wrist arthrodesis using a circular limited wrist fusion plate for stage II scapholunate advanced collapse wrist deformity. *Plast Reconstr Surg.* 2006 Aug; 118(2):433–42. [PubMed: 16874215]
20. 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 Aug; 19(3):471–81. [PubMed: 12945645]
21. Waljee JF, Chung KC, Kim HM, Burns PB, Burke FD, Wilgis EF, et al. The validity and responsiveness of the Michigan Hand Questionnaire (MHQ) in patients with rheumatoid arthritis: A Multicenter, International Study. *Arthritis Care Res.* 2010 in press.
22. Boyle GJ. Does item homogeneity indicate internal consistency or item redundancy in psychometric scales? *Personality and Individual Differences.* 1991; 12:291–4.
23. Peterson R. A meta-analysis of Cronbach's coefficient alpha. *Journal of Consumer Research.* 1994; 21:381–91.
24. Ware J Jr, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care.* 1996 Mar; 34(3):220–33. [PubMed: 8628042]
25. Ware JE Jr. Improvements in short-form measures of health status: introduction to a series. *J Clin Epidemiol.* 2008 Jan; 61(1):1–5. [PubMed: 18083456]
26. Dorman PJ, Slattery J, Farrell B, Dennis MS, Sandercock PA. A randomised comparison of the EuroQol and Short Form-36 after stroke. United Kingdom collaborators in the International Stroke Trial. *BMJ.* 1997 Aug 23.315(7106):461. [PubMed: 9284664]
27. Holland R, Smith RD, Harvey I, Swift L, Lenaghan E. Assessing quality of life in the elderly: a direct comparison of the EQ-5D and AQoL. *Health Econ.* 2004 Aug; 13(8):793–805. [PubMed: 15322991]
28. (NCQA) NCfQA. Annual Member Health Care Survey Manual. Washington, D.C: 1995.
29. Beaton DE, Wright JG, Katz JN. Development of the QuickDASH: comparison of three item-reduction approaches. *J Bone Joint Surg Am.* 2005 May; 87(5):1038–46. [PubMed: 15866967]
30. Gandek B, Ware JE, Aaronson NK, Apolone G, Bjorner JB, Brazier JE, et al. Cross-validation of item selection and scoring for the SF-12 Health Survey in nine countries: results from the IQOLA Project. International Quality of Life Assessment. *J Clin Epidemiol.* 1998 Nov; 51(11):1171–8. [PubMed: 9817135]

31. 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 May–Jun; 34(5):815–23. [PubMed: 19410984]
32. Chung KC, Kotsis SV, Kim HM. A prospective outcomes study of Swanson metacarpophalangeal joint arthroplasty for the rheumatoid hand. *J Hand Surg Am.* 2004 Jul; 29(4):646–53. [PubMed: 15249089]
33. Kotsis SV, Chung KC. Responsiveness of the Michigan Hand Outcomes Questionnaire and the Disabilities of the Arm, Shoulder and Hand questionnaire in carpal tunnel surgery. *J Hand Surg Am.* 2005 Jan; 30(1):81–6. [PubMed: 15680560]
34. 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 Jul–Aug; 33(6):809–19. [PubMed: 18656749]
35. Jepsen RH, Taylor N, Trieschmann RB, Trotter MJ, Howard LA. An objective and standardized test of hand function. *Arch Phys Med Rehabil.* 1969 Jun; 50(6):311–9. [PubMed: 5788487]
36. 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 Jan; 35(1):1–10. [PubMed: 1731806]
37. Trochim, W.; Donnelly, JP. *The Research Methods Knowledge Base.* 2. Cincinnati, Ohio: 2007.
38. Doros G, Lew R. Design based on intra-class correlation coefficients. *American Journal of Biostatistics.* 2010; 6(1):1–8.
39. Sammer DM, Shah HM, Shauver MJ, Chung KC. The effect of ulnar styloid fractures on patient-rated outcomes after volar locking plating of distal radius fractures. *J Hand Surg Am.* 2009 Nov; 34(9):1595–602. [PubMed: 19896004]
40. 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 Nov; 31(9):1447–53. [PubMed: 17095372]
41. Chung KC, Hamill JB, Walters MR, Hayward RA. The Michigan Hand Outcomes Questionnaire (MHQ): assessment of responsiveness to clinical change. *Ann Plast Surg.* 1999 Jun; 42(6):619–22. [PubMed: 10382797]
42. Liang MH, Lew RA, Stucki G, Fortin PR, Daltroy L. Measuring clinically important changes with patient-oriented questionnaires. *Med Care.* 2002 Apr; 40(4 Suppl):II45–51. [PubMed: 12064581]
43. Cohen, J., editor. *Statistical power analysis for the behavioral sciences.* 2. Hillsdale, N.J: Lawrence Erlbaum Associates; 1988.
44. Schuind FA, Mouraux D, Robert C, Brassinne E, Remy P, Salvia P, et al. Functional and outcome evaluation of the hand and wrist. *Hand Clin.* 2003 Aug; 19(3):361–9. [PubMed: 12945632]
45. 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 Apr–Jun; 14(2):128–46. [PubMed: 11382253]
46. 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 Nov; 35(4):735–43. vii. [PubMed: 19962618]
47. Boers M, Felson DT. Clinical measures in rheumatoid arthritis: which are most useful in assessing patients? *J Rheumatol.* 1994 Sep; 21(9):1773–4. [PubMed: 7799368]
48. 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 Feb; 68(2):183–90. [PubMed: 18375533]
49. Halpern SD, Asch DA. Improving response rates to mailed surveys: What to we learn from randomized controlled trials? *International Journal of Epidemiology.* 2003; 32:637–8. [PubMed: 12913043]
50. Jepsen C, Asch DA, Hershey JC, Ubel PA. In a mailed physician survey, questionnaire length had a threshold effect on response rate. *J Clin Epidemiol.* 2005 Jan; 58(1):103–5. [PubMed: 15649678]

51. Hawthorne G. Assessing utility where short measures are required: development of the short Assessment of Quality of Life-8 (AQoL-8) instrument. *Value Health*. 2009 Sep; 12(6):948–57. [PubMed: 19490558]
52. Gandak B, Ware JE, Aaronson NK, Apolone G, Bjorner JB, Brazier JE, et al. Cross-validation of item selection and scoring for the SF-12 Health Survey in nine countries: results from the IQOLA Project. *Journal of Clinical Epidemiology*. 1998; 51(11):1171–8. [PubMed: 9817135]
53. Marx RG, Bombardier C, Hogg-Johnson S, Wright JG. Clinimetric and psychometric strategies for development of a health measurement scale. *J Clin Epidemiol*. 1999 Feb; 52(2):105–11. [PubMed: 10201650]
54. Green, KE.; Frantom, CG. Survey development and validation with the Rasch Model. University of Missouri; University of Denver; 2002.
55. Eurenus E, Brodin N, Lindblad S, Opava CH. Predicting physical activity and general health perception among patients with rheumatoid arthritis. *J Rheumatol*. 2007 Jan; 34(1):10–5. [PubMed: 17216673]
56. Yazici Y, Sokka T, Pincus T. Radiographic measures to assess patients with rheumatoid arthritis: advantages and limitations. *Rheum Dis Clin North Am*. 2009 Nov; 35(4):723–9. vi. [PubMed: 19962616]
57. Sears ED, Chung KC. Validity and Responsiveness of the Jepsen-Taylor Hand Function Test. *J Hand Surg Am*. 2009 Nov 30.

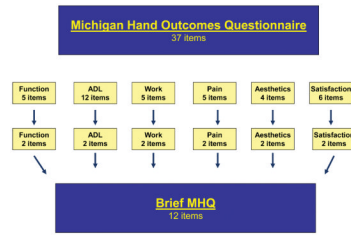


Figure 1.
Development of the Brief Michigan Hand Questionnaire

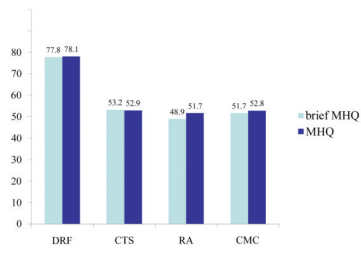


Figure 2. The comparison between the adjusted mean summary scores of the brief MHQ and the original MHQ, stratified by disease type. Adjusted for age, gender and education level.

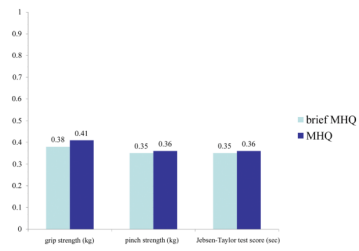


Figure 3. The correlation between the brief MHQ and objective measures of function, including grip and pinch strength, and Jebsen-Taylor test score, adjusted for disease type
Adjusted for age, gender and education level.

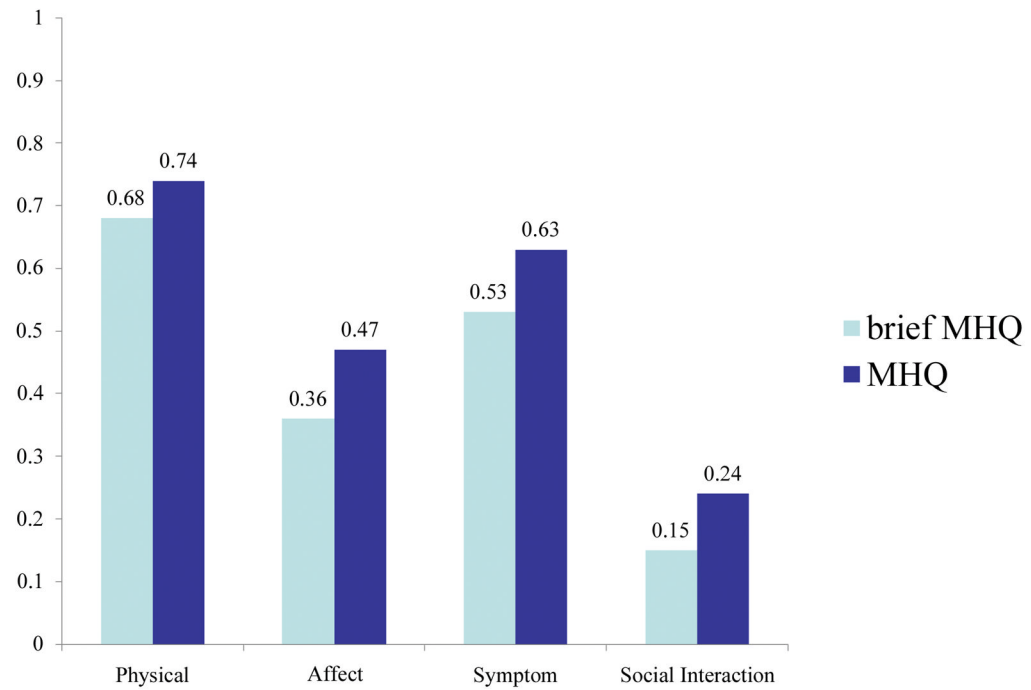


Figure 4. The correlation between the brief MHQ and patient self-assessment of function using the AIMS2 survey among RA patients
Adjusted for age, gender and education level.

Table 1

Descriptive statistics of the study population. (n=422)

Patient characteristics	n
Age (mean, SD)	Mean: 54.9 (range:18–90)
Gender (% Female)	291 (69.5%)
Education (% ≥ college)	134 (32.8%)
Disease type	
Distal radius fracture	132
Carpal tunnel syndrome	97
Rheumatoid arthritis	162
CMC arthritis	31

Table 2

The Correlation Between Individual Items of MHQ with the Summary MHQ Score (Items Retained for the Brief MHQ are Indicated in Bold Print).

<u>MHQ domain</u>	<u>Item (r, p-value)</u>
<u>Function</u>	<p>Overall, how well did you hand work (0.41, p<0.001) How was the sensation in your hand (0.21, p<0.001) How well did your fingers move (0.11, p<0.001)</p>
<u>ADLs</u>	<p>Hold a frying pan (0.16, p<0.001) Button a shirt/blouse (0.17, p<0.001) Hold a glass of water (0.06, p<0.16) Turn a key in a lock (0.06, p<0.25) Turn a door knob (0.02, p<0.63) Open a jar (0.11, p<0.001)</p>
<u>Work</u>	<p>How often were you unable to do your work in the last week because of your hands/wrists (0.25, p<0.001) How often did you take longer to do tasks in your work because of problems with your hands/wrists (0.34, p<0.001) How often do you have to take it easy at your work because of problems with your hands/wrists (0.009, p<0.85)</p>
<u>Pain</u>	<p>How often did the pain in your hands/wrists interfere with your daily activities (0.39, p<0.001) Describe the pain in your hands/wrists (0.18, p<0.001) How often do you have pain in your hands (0.10, p<0.012)</p>
<u>Aesthetics</u>	<p>I am satisfied with the look of my hands (0.33, p<0.001) The appearance of my hands interferes with my normal daily activities (0.32, p<0.001)</p>
<u>Satisfaction</u>	<p>Overall function of your hand (0.19, p<0.001) Motion of your fingers (0.14, p<0.003) Motion of your wrist (0.19, p<0.001)</p>
	<p>How was the strength in your hand (0.12, p<0.004) How well did your wrist move (0.16, p<0.001)</p> <p>Pick up a coin (0.14, p<0.002) Eat with a knife/fork (0.14, p<0.001) Carry a grocery bag (0.06, p<0.10) Wash dishes (0.09, p<0.05) Wash your hair (0.04, p<0.34) Tie shoelaces/knots (0.02, p<0.64)</p> <p>How often did you accomplish less in your work because of problems with your hands/wrists (0.07, p<0.18) How often did you have to shorten your work day because of problems with your hands/wrists (0.19, p<0.001)</p> <p>How often did the pain in your hands/wrists interfere with your sleep (0.11, p<0.006) How often did the pain in your hands/wrists make you unhappy (0.18, p<0.001)</p> <p>The appearance of my hands makes me depressed (0.21, p<0.001) The appearance of my hands makes me uncomfortable in public (0.09, p<0.10)</p> <p>Strength of your hand (0.09, p<0.012) Pain level of your hands (0.27, p<0.001) Sensation of your hands (0.13, p<0.001)</p>

Test-Retest Correlation Comparing Baseline and 6-month Follow Up Scores for the Brief MHQ score (n=68 RA Control Patients Who Did Not Undergo Surgical Intervention).

Table 3

	Brief MHQ score at baseline	Brief MHQ score at 6 months follow-up	Mean difference	Intra-class correlation	r
Brief MHQ	55.2± 16.1	55.0 ± 17.5	0.22	0.91	0.78
Full MHQ	59.8 ± 16.6	60.2 ± 18.4	0.38	0.90	0.76

Table 4
 Responsiveness of the Brief MHQ and Original MHQ Summary Score to Clinical Change, Stratified by Disease Type.*

	Brief MHQ			MHQ				
	Baseline	Follow-up	p value [♣]	SRM [†]	Baseline	Follow-up	p value [♣]	SRM [†]
Distal Radius Fractures (n=92)	76.0 ± 14.6	84.1 ± 14.1	0.001	0.91	79.3 ± 15.2	85.8 ± 13.7	0.001	0.73
Carpal Tunnel Syndrome (n=55)	53.9 ± 18.8	71.2 ± 20.4	0.001	1.01	55.0 ± 18.6	74.8 ± 21.0	0.001	1.00
Rheumatoid Arthritis^{††} (n= 55)	41.0 ± 18.5	57.0 ± 19.4	0.001	1.28	43.7 ± 18.4	59.7 ± 19.7	0.001	1.36
CMC Arthritis (n= 16)	51.8 ± 17.2	62.2 ± 14.0	0.006	0.80	54.6 ± 16.2	64.6 ± 15.7	0.008	0.71

* Cell values are mean ±SD.

♣ Paired t-test comparing follow up scores and baseline scores.

† Standardized Response Mean is calculated using the following formula: (follow-up mean – baseline mean)/standard deviation of the change

†† RA patients who have undergone SMPA arthroplasty and have both baseline and follow-up measurements available

Table 5

Comparison of responsiveness of the Brief MHQ score and objective function measures, controlling for disease type.*††

	Baseline value	Follow-up value	p value♣	SRM†
Brief MHQ Score	58.8 ± 14.4	70.5 ± 11.5	0.001	4.15
Original MHQ Score	59.7 ± 13.6	70.9 ± 10.2	0.001	3.30
Grip Strength (kg)	10.7 ± 3.3	11.6 ± 5.7	0.001	0.36
Pinch Strength (kg)	5.3 ± 1.7	14.1 ± 3.75	0.001	1.60
Jebsen-Taylor Test score (sec)	40.6 ± 8.57	44.2 ± 11.45	0.001	1.26

* Cell values are mean ±SD.

♣ Paired t-test comparing follow up scores and baseline scores.

† Standardized Response Mean is calculated using the following formula: (follow-up mean – baseline mean)/standard deviation of the change

†† Excluding patients with carpal tunnel syndrome, and RA patients who did not undergo SMFA arthroplasty. Patients who did not have both baseline and follow-up measurements were also excluded from analysis