Assessment of construct validity of the Finnish versions of the Disabilities of Arm, Shoulder and Hand Instrument and the Michigan Hand Outcomes Questionnaire

Assessment of construct validity of the Finnish versions of the Disabilities of Arm, Shoulder and Hand Instrument and the Michigan Hand Outcomes Questionnaire

Mikko M. Uimonen, BM¹; Sina M. Hulkkonen, MD², Jorma Ryhänen, MD, PhD²; Ville T. Ponkilainen, BM¹; Arja H. Häkkinen, PhD³,⁴; Jaro Karppinen, MD, PhD⁵,⁶; Jussi P. Repo, MD, PhD¹

1. Department of Surgery, Central Finland Central Hospital, Jyväskylä, Finland.

2. Department of Hand Surgery, Helsinki University Hospital and University of Helsinki, Finland.

3. Health Sciences, Faculty of Sport and Health Sciences, University of Jyväskylä, Jyväskylä, Finland.

4. Department of Physical Medicine, Central Finland Health Care District, Jyväskylä, Finland.

5. Medical Research Center Oulu, University of Oulu and Oulu University Hospital, Oulu, Finland

6. Finnish Institute of Occupational Health, Aapistie 1, 90220 Oulu, Oulu, Finland

Corresponding author:

Mikko M. Uimonen

Department of Surgery, Central Finland Central Hospital, Jyväskylä, Finland

Email: uimonen.mikko.m@student.uta.fi
Abstract

Background: There is a lack of information on the measurement properties of patient-reported upper extremity instruments and their association to health-related quality of life. The existing upper extremity specific measures need further validation. The aim of this study was to measure and compare the measurement properties and construct validity of the Disabilities of Arm, Shoulder, and Hand (DASH) Instrument and the Michigan Hand Questionnaire (MHQ) using a heterogeneous sample of patients with hand and wrist problems.

Methods: Two hundred-fifty consecutive patients visiting a general orthopedic outpatient clinic due to hand/wrist problems were invited to participate in the study. Altogether 230 patients agreed to participate and finally, a total of 193 (77%) participants provided sufficient patient-reported outcome data and were included in the analysis. Participants with various hand or wrist complaints completed the DASH, the MHQ, the EQ-5D-3L health-related quality of life (HRQoL) and Pain on a visual analogue scale instruments. Grip and key pinch forces were measured. Scale targeting, relatedness of demographics and construct validity of the DASH and the MHQ were assessed.

Results: Both the DASH and the MHQ had good targeting, but the DASH had wider coverage. The convergence between the DASH and the MHQ was high. The DASH was more closely related to HRQoL than the MHQ in terms of EQ-5D scores.

Conclusions: The DASH instrument appeared to measure hand function and disability from a perspective of HRQoL superior to the MHQ among patients with heterogeneous hand and wrist complaints.
Introduction

Increased interest in the outcomes of medical treatment has accelerated the development and use of patient-reported outcome (PRO) instruments as a part of clinical outcome assessment\(^1\text{-}^3\). Extensive research and validation of such instruments have shown them to be useful in evaluating and comparing treatment outcomes \(^4\). However, as the applicability of different PRO instruments may vary in diverse study samples \(^2\). In order to more accurately select a proper instrument for the population under examination \(^5\), it is beneficial to understand the measurement properties of different PRO instruments in head-to-head comparison. A systematic review of Van de Ven-Stevens et al. \(^6\) reviewed the clinimetric properties of 23 instruments for assessing hand function after hand injury, including five PRO instruments. There were vital shortages in the reported properties of all the surveyed PRO instruments. This finding refers to insufficient understanding of the key features of these hand-specific measures.

The Disabilities of Arm, Shoulder and Hand (DASH) instrument \(^7\) and the Michigan Hand Outcomes Questionnaire (MHQ) \(^8\) are widely adopted PRO instruments for evaluating the performance and disability of upper limbs or hands \(^5\). The clinimetric properties of the DASH have been investigated using classical test theory \(^9\) and the Rasch measurement theory \(^10\text{-}^12\), and several reports have assessed the validity of the MHQ among hand patients \(^8,^{13}\text{-}^{16}\). A study by Dias et al. compared three upper extremity-specific PRO instruments, the Patient evaluation measure (PEM), the DASH, and the MHQ \(^17\). It found the DASH and the MHQ to be valid and reliable for a sample of patients with various wrist or finger complaints although there were shortages in construct validity of all PRO instruments measured by correlation testing between the instruments scores and hand symptom severity \(^17\). Nonetheless, thus far there has been a lack of high-quality comparison of these two PRO instruments with a perspective of health-related quality of life (HRQoL) outcomes.

Comparison of the association of the DASH and the MHQ to HRQoL provides valuable information for researchers and clinicians dealing with hand and wrist problems. The results could
potentially guide us to choose the right instrument for assessing the function or disability of
patients.

The aim of this study was to measure and compare the scale targeting and construct
validity of the Finnish versions of the DASH and the MHQ and their association to HRQoL using a
heterogeneous sample of patients with hand and wrist problems in order to better understand the
clinimetrics of these two widely used PRO instruments.

Materials and Methods

In 2017, two hundred-fifty consecutive patients with hand and wrist problems treated at the general
orthopedic outpatient clinic in Länsi-Pohja Central Hospital in Kemi, Finland, were invited face-to-
face to participate in the study. The inclusion criteria were age of 18 years or above, full
understanding of spoken and written Finnish, a lack of cognitive disabilities, and the ability to give
signed informed consent to participating in the study. Overall, 230 invited patients were willing to
participate in the study. Of these 230 patients, 193 had completed all the PRO instruments
sufficiently and were included in the analysis, giving us an effective response rate of 77%.

The participants were clinically examined and cognitively debriefed by a surgeon
during their appointment at the hospital. Both hands were examined. The participants were asked to
complete the Finnish versions of the DASH instrument and the MHQ. Both PRO instruments are
available in Finnish \(^{12,18}\). If both hands were affected, we selected the MHQ scores of the hand that
the participant considered worse for further analysis whereas the DASH does not distinguish
between hands and thus, we analyzed the DASH scores regardless of which hand was affected. The
participants also filled in the EQ-5D-3L generic health-related quality of life (HRQoL)
questionnaire. In addition, the participants were asked to complete a visual analogue scale (VAS)
pain instrument (“Place a line perpendicular to the line from 0 to 100 at the point that represents
their pain intensity with 0 representing ‘No pain at all’ and 100 representing ‘The worst imaginable
They returned the completed questionnaires during the final appointment at the hospital. We collected the clinical and demographic details. Grip and key pinch forces were measured using a Baseline hydraulic hand dynamometer together with a pinch gauge (White Plains, NY, USA). The Ethics Committee of the Northern Ostrobothnia Hospital District approved the study protocol. All the participants gave their written informed consent.

**Disabilities of the arm, shoulder and hand (DASH)**

The DASH is an upper limb-specific PRO instrument. It has been validated among patients with various hand and upper limb complaints. The instrument was developed to evaluate patients’ disability and performance. The DASH has been widely adopted in use and is available in several languages. It consists of two sections, one covers physical activities (23 items) and the other covers symptoms (7 items). All items are on a Likert scale from 1 to 5, with higher score indicating higher disability over the preceding week. At least 27 items must be completed to enable calculation of the total score. The total score is scaled from 0 to 100 by dividing it by the number of items responded to and then subtracting one. Finally, the result is multiplied by 25. Cronbach’s alpha indicates high internal consistency at 0.97. Test-retest reliability is estimated as high, as intraclass correlation coefficient (ICC) has been reported as 0.96 (95% CI: 0.93-0.98) with a three- to five-day interval between the assessments. A difference of 10 points is considered the minimum significant change.

**Michigan Hand Outcomes Questionnaire (MHQ)**

The MHQ is a 37-item questionnaire, which was developed to evaluate the health state of patients with hand disorders. The questionnaire includes six domains: Overall hand function, Activities of daily living (ADL), Pain, Work performance, Aesthetics, and patient’s Satisfaction with hand function. The raw scores for each domain are calculated as a sum of the items in the category. The raw score is converted to 0 to 100 scale, in which a higher score indicates better hand function.
The Pain domain is reversed, as a higher score indicates less pain. The developers of the questionnaire estimated Cronbach’s alphas to be from 0.86 to 0.97 for all dimensions indicating high internal consistency. Test-retest reliability has proven to be high, as the ICC for all dimensions ranges from 0.81 to 0.97. The minimum significant change between the dimensions has been presented as 17.

EuroQol instrument (EQ-5D-3L)

The EQ-5D-3L is a generic HRQoL instrument. It consists of five dimensions: Mobility, Self-Care, Usual Activities, Pain/Discomfort, and Anxiety/Depression. Each dimension has three response categories: no problems, some problems or severe problems. The instrument also includes an item on overall health state, on a 0–100 visual analogue scale (EQ-VAS). The result is presented as a five-digit number, a health state, and contains responses to all dimensions. This can be further converted into a single index, varying from -0.011 to 1 in the Finnish version of EQ-5D-3L; a lower score indicating poorer health-related quality of life.

Statistical Methods

The scores of all the instruments were converted to scale from 0 to 100. To obtain a parallel effect direction, the scores of the EQ-5D-3L index, EQ-VAS and the MHQ were inverted by subtracting the score from 100. After conversion, lower scores indicated better outcomes in all scales and higher scores indicated worse outcomes. Clinical, demographic and questionnaire data are presented as means with standard deviations (SD) or 95% confidence intervals (95% CI) or counts with percentages. Scale targeting was assessed by evaluating score distributions, and floor and ceiling effects. Floor or ceiling effects were considered confirmed if 15% of the participants scored the minimum or maximum points. Relatedness of the DASH and the MHQ with demographic features was examined by calculating Pearson correlation coefficients between the instruments’ scores and the ages of the participants, and by comparing the mean scores of male and female participants with
independent samples t-test. To produce comparable and generalizable results, the instruments should ideally be unrelated of demographic factors.

Construct validity refers to the extent to which the PRO instrument measures what it is supposed to. We evaluated construct validity of the DASH and the MHQ by assessing the convergence between these two hand specific PRO instruments scores, association with objective hand outcome measures, grip and key pinch forces, and relationship with HRQoL.

The relationship between grip and key pinch forces and the DASH and the MHQ was examined by calculating Pearson correlation coefficients. To evaluate convergence of the DASH and the MHQ, we calculated the Pearson correlation coefficients between the instruments. The correlations were represented according to the Rule of Thumb for interpreting the size of the correlation coefficient \(^2\): 0.00–0.30 negligible, 0.30–0.50 low, 0.50–0.70 moderate, 0.70–0.90 high and 0.90–1.00 very high correlation. To further examine the underlying constructs measured by the DASH and the MHQ from a perspective of HRQoL, we conducted Principal Component Analysis (PCA) to reduce the variables to the main factors to enable us to examine their influence on the DASH and MHQ scores. All EQ-5D scores and VAS Pain were included in the PCA. The continuous variables (the EQ-5D index, EQ-VAS, and VAS Pain) were transformed into a logarithmic scale \(^3\). The Principal Component (PC) was required to have an eigenvalue either equal to or higher than 1 to be selected for further examination, in accordance with the Kaiser criteria \(^4\).

We evaluated the distributions of the DASH and MHQ scores against the selected PCs. Locally estimated scatterplot smoothing (LOESS) curves with 95% CIs were formed to illustrate the correlation of the PCs with the DASH and MHQ scores. Similar distributions of the DASH and the MHQ scores around the PCs indicate high convergence of the instruments from the perspective of HRQoL related PCs. We used HRQoL instrument EQ-5D and its subscales, EQ-VAS and Pain VAS as a
criterion measures when assessing relatedness of the hand outcome instruments of HRQoL. Pearson correlation coefficients between the criterion measures and the DASH and the MHQ were calculated. In addition, we used age- and gender-standardized linear regression analyses to identify the strength of the influence of the DASH and MHQ scores on the perceived HRQoL of the participants. Regression coefficients (beta, β) indicated how strongly the DASH and MHQ scores influenced the criterion variables’, EQ-5D scores and VAS Pain scores. The β was measured in units of SD. The Cohen reference values for regression coefficient β were <0.1, <0.3 and <0.5 for small, moderate and strong influence, respectively. In addition, we examined the association between HRQoL and the instruments’ scores by dividing the participants into subgroups according to the EQ-5D index. The 25%, 50% and 75% percentiles were used as cut-off values between subgroups. Residual analysis of linear regression of the DASH and the MHQ scores against the EQ-5D index was conducted to assess the strength of linear relationship within the subgroups.

The statistical analyses were performed using R (version 1.1.453) and IBM SPSS 25.0 statistics software. The results of this study are interpreted and reported in adherence to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement and the Consensus-based Standards for the selection of health status Measurement Instruments (COSMIN) guidelines.

**Results**

Table 1 presents the sociodemographic and clinical data of the participants. We observed no floor or ceiling effects in either PRO instrument, although six (3.1%) participants scored the minimum score in the DASH. Figure 1 shows the distributions of DASH and MHQ scores. The distribution of the MHQ scores followed normal distribution, whereas the DASH score distribution was skewed towards lower disability. However, the DASH scores covered the scale more comprehensively than the MHQ scores. In addition, the EQ-5D score was distributed normally.
There was similar pattern of the DASH and the MHQ score distributions in diagnostic subgroups. The subgroups of Distal radius fracture (mean DASH score = 41, MHQ score = 51), Other fractures (DASH score = 36, MHQ score = 49) and Carpometacarpal 1 joint osteoarthritis (DASH score = 39, MHQ score = 44) obtained the highest scores from both the DASH and the MHQ indicating worst hand function, whereas the subgroups of Dupuytren’s disease (mean DASH score = 17, MHQ score = 32), Ulnar nerve entrapment (DASH score = 19, MHQ score = 32) and Ganglion cyst (DASH score = 20, MHQ score = 33) obtained the lowest scores indicating, in turn, best hand function.

The Pearson correlation coefficients between the instruments’ scores and age showed no significant correlations. On the other hand, female participants obtained higher scores from both instruments as the mean DASH scores were 35 for female vs. 25 for male (p = 0.003) and the mean MHQ scores were 42 vs. 38 (p = 0.018). Both instruments’ scores correlated negatively with grip and key pinch forces as the Pearson correlation coefficients of the DASH were -0.50 (p < 0.001) for grip pinch and -0.45 (p < 0.001) for key pinch and of the MHQ were -0.53 (p < 0.001) for grip pinch and -0.48 (p < 0.001) for key pinch forces.

Pearson correlation coefficient r between the DASH and MHQ scores was 0.75 (95% CI = 0.68 to 0.82; p < 0.001), which indicates high convergent validity of the instruments (Figure 2). We observed moderate correlations between the DASH score and the EQ-5D index (Figure 3A), Self-care, Usual activity, Pain/Discomfort, EQ-VAS and theVAS Pain scales (Table 2). We also observed low but still notable correlations between MHQ score and the EQ-5D index (Figure 3B), Self-care, Usual activity, Anxiety/Depression, EQ-VAS and VAS Pain, as well as between the DASH score and Anxiety/Depression. All the correlations of the DASH with the reference outcomes were higher than those of the MHQ.

Figure 4 presents the age- and sex-adjusted regression coefficient β of the DASH and MHQ scores on the HRQoL measures. According to the Cohen reference values, all scores except
those of Mobility and Pain/Discomfort reached beta values of over 0.3 against the MHQ score, indicating a moderate influence. The MHQ scores had the strongest influence on Usual activity ($\beta = 0.49$). The DASH scores had a strong influence on the EQ-5D index, Self-care, Usual activity, EQ-VAS, and Pain VAS with betas over 0.5 and at least a moderate influence on all scores except that of Mobility with betas over 0.3. The highest $\beta$ was for the EQ-5D index ($\beta = 0.61$). All of the associations were statistically significant (Figure 3). Overall, according to the regression coefficients, the DASH had a notably stronger association with the HRQoL measures than the MHQ. Thus, the HRQoL related construct validity of the DASH was considered higher than that of the MHQ.

Figure 5 presents the MHQ and DASH scores in each HRQoL group. The DASH scores show consistent improving trend when HRQoL improves from Very bad to Very good. In addition, the MHQ scores show improving trend when HRQoL improves from Very bad to Good but when HRQoL improves from Good to Very good, the trend seems to plateau. The residual analysis of linear regression of the DASH and the MHQ scores against the EQ-5D index within the HRQoL groups showed consistent strong relationship between the DASH and the EQ-5D index across the HRQoL groups. However, the regression coefficient beta values showed that the strength of the relationship decreases as the HRQoL improves as the beta values were 1.39 ($p < 0.001$), 1.00 ($p < 0.001$), 0.85 ($p < 0.001$) and 0.53 ($p < 0.001$) for Very bad, Bad, Good and Very good HRQoL, respectively. On the other hand, there was significant linear relationship between the MHQ and the EQ-5D index in only subgroup of Very bad HRQoL ($\beta = 0.47$, $p < 0.001$).

The principal component analysis generated two principal components that met the Kaiser criteria $^{31}$ of an eigenvalue over 1, and were thus included for further examination. The eigenvalues of the first principal component (PC1) and the second principal component (PC2) were 3.16 and 1.26, respectively. PC1 explained 39.6% of the total variance of reference outcome measures, whereas PC2 explained 15.7%. Overall, the selected PCs explained 55.3% of the total
variance. The loadings of the reference outcome measures on the PCs are presented in Figure 6. According to the loading plot, the loading vectors can be divided to two groups. The first group includes the EQ-5D index, EQ-VAS, Pain/Distress, and the VAS Pain scales, representing the measures of general HRQoL and pain. The second group includes Self-care, Usual activities, and Anxiety/Depression, which can be interpreted as representing the independency and performance of participants. The mobility dimension was loaded towards the first group, but the loading was merely weak. Figure 7A-D presents scatter plots and the LOESS curves with the 95% CIs of the DASH and the MHQ against PCs. The distributions of the DASH and MHQ scores were similar for both PCs.

Discussion

The main finding of this study was that both the DASH and the MHQ instruments’ scores strongly correlated with each other as well as with the generic HRQoL instrument scores. Both instruments had comparable measurement properties when the outcomes of various hand complaints were evaluated (Table 3). However, in terms of its strong relationship with HRQoL, the DASH instrument seemed to have more suitable measurement properties for the study sample than the MHQ when evaluating HRQoL related outcomes after hand complaints.

The differences in the distributions of the DASH and MHQ scores pointed to wider spread of the DASH scores than of the MHQ scores. Although, the distribution of the DASH scores was skewed, it still covered the participants well, whereas the MHQ scores focused on more narrow spectrum. On the other hand, we observed no floor or ceiling effects, indicating that both instruments measured the spectrum of the hand complaints in the present sample well, which in turn indicates good targeting for both instruments. Previous studies on the DASH and the MHQ have also found similar distributions ⁹, ³⁴⁻³⁶.

The evaluation of relatedness of the instruments with demographics showed similar limitations for both instruments as female participants scored higher than male indicating worse
outcomes for female than male. The finding might be due to male participants having more strength which in turn might help compensating the problems caused by hand complaints. On the other hand, age was not associated with the instruments’ scores. As there were similar pattern between the DASH and the MHQ concerning the relatedness with sex, the issue should be taken into account when these instruments are applied in further studies.

As expected, the correlation between the DASH and the MHQ scores was high, in line with prior studies 17, 37. Furthermore, the principal component analysis gave parallel results as the DASH and the MHQ scores were distributed similarly for both recognized principal components. Lastly, both instruments correlated significantly with grip and key pinch forces. The findings indicated that both instruments measured the same constructs, which, in turn, proposes good construct validity of each instruments.

Both instruments correlated strongly with the generic HRQoL scores (EQ-5D) although the DASH correlations were stronger than those of the MHQ. In addition, the regression analysis results indicated that the influence of the DASH score on HRQoL was stronger than that of the MHQ. We also observed this in the comparison of the instruments’ scores in the subgroups of HRQoL, as the DASH scores shifted consistently in compliance with HRQoL while the MHQ scores were unable to distinguish the differences when the HRQoL was good or very good. The residual analysis of linear regression in HRQoL subgroups showed strong relationship between the DASH score and the EQ-5D index across the subgroups while regarding the MHQ, the relationship was found only in Very bad-subgroup. The findings propose that the DASH score was superior to the MHQ in assessing the HRQoL outcomes in our sample. The highest correlation and regression coefficients of each hand instrument were in the EQ-5D’s Daily activity dimension and the EQ-5D index, suggesting that the management of everyday tasks plays a key role in the assessment of hand function. Prior studies have had parallel results, proposing that hand performance measures associate with general physical function 8, 38.
Although both instruments focus mostly on the same themes on hand outcomes, there are differences in the perspectives of the instruments. While the subscales of the DASH focus on the functions of daily living and symptoms of the hand, the MHQ has more comprehensive point of view as the subscales also cover the management of daily chores and work as well as aesthetics of the hand. However, despite the wider perspective of the MHQ on hand outcomes, the DASH score was more closely related to HRQoL measured by EQ-5D than the MHQ score. According to our findings, the DASH seems to capture the key aspects of hand outcomes related to HRQoL better than the MHQ in a sample of patients with various hand complaints. On the other hand, there is a fundamental difference in the scopes of the DASH and the MHQ. As the MHQ scores were analyzed for the worse hand, the DASH does not distinguish between the hands. Thus, the DASH scores might not be as sensitive to hand complaints if the unaffected hand compensate the issues with affected hand. Hence, if the DASH score shows impairment, it is more likely, that the hand issues affect daily living and HRQoL of the patient. This may explain the stronger association between the DASH and the EQ-5D index than between the MHQ and the EQ-5D index. In addition, the distributions of the DASH and the MHQ scores supported this idea as the MHQ scores showed more hand issues than the DASH. Thus, the MHQ might be more sensitive if only one hand is affected.

Our study had some limitations. First of all, we did not examine the responsiveness to change of the instruments under observation, although it is an essential measurement property in the validity evaluation of PRO instruments. Second, our study was conducted in only one hospital district. However, the strength of this study was its comprehensive, large sample of various hand and wrist complaints. In addition, as the statistical methods it used have not previously been used for evaluating and comparing the clinimetric properties of the DASH and the MHQ, our study provides further insight into the construct validity of these two outcome instruments.

**Conclusion**
The scores of the DASH and the MHQ were highly correlated. The DASH scores had a stronger relationship with the HRQoL outcomes. Thus, the DASH instrument appeared to be superior to the MHQ in evaluating the outcomes from a perspective of HRQoL among patients with heterogeneous hand and wrist complaints. On the other hand, the MHQ might be a more specific instrument when measuring performance of the affected hand.
**Figures and tables**

**Table 1.** Participants’ sociodemographic and clinical details.

<table>
<thead>
<tr>
<th></th>
<th>N=193</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age in years, mean (SD)</td>
<td>54 (15)</td>
</tr>
<tr>
<td>Women, n (%)</td>
<td>114 (59)</td>
</tr>
<tr>
<td>BMI, mean (SD)</td>
<td>28 (5.6)</td>
</tr>
<tr>
<td>Right hand affected, n (%)</td>
<td>109 (56)</td>
</tr>
<tr>
<td>Grip force of affected hand, kg, mean (SD)</td>
<td>29 (16)</td>
</tr>
<tr>
<td>Pinch force of affected hand, kg, mean (SD)</td>
<td>7.2 (2.9)</td>
</tr>
<tr>
<td>Number of diagnoses</td>
<td></td>
</tr>
<tr>
<td>Carpal tunnel syndrome</td>
<td>82</td>
</tr>
<tr>
<td>Trigger finger</td>
<td>25</td>
</tr>
<tr>
<td>Distal radius fracture</td>
<td>20</td>
</tr>
<tr>
<td>Other fracture of the hand/wrist</td>
<td>20</td>
</tr>
<tr>
<td>Ganglion cyst</td>
<td>17</td>
</tr>
<tr>
<td>Dupuytren’s disease</td>
<td>17</td>
</tr>
<tr>
<td>Carpometacarpal 1 joint osteoarthritis</td>
<td>16</td>
</tr>
<tr>
<td>Ulnar nerve entrapment</td>
<td>3</td>
</tr>
<tr>
<td>Other</td>
<td>10</td>
</tr>
<tr>
<td>DASH score, mean (SD)</td>
<td>32 (22)</td>
</tr>
<tr>
<td>MHQ score, mean (SD)</td>
<td>40 (13)</td>
</tr>
<tr>
<td>EQ-5D index, mean (SD)</td>
<td>34 (17)</td>
</tr>
<tr>
<td>EQ-VAS, mean (SD)</td>
<td>26 (17)</td>
</tr>
<tr>
<td>VAS Pain, mean (SD)</td>
<td>36 (25)</td>
</tr>
</tbody>
</table>
Table 2. Pearson correlation coefficients between DASH and MHQ, and EQ-5D scores and VAS Pain.

<table>
<thead>
<tr>
<th>Variable</th>
<th>DASH, r (95% CI)</th>
<th>MHQ, r (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EQ-5D index</td>
<td>0.64 (0.53 to 0.73)</td>
<td>0.43 (0.28 to 0.56)</td>
</tr>
<tr>
<td>EQ-5D dimensions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mobility</td>
<td>0.23 (0.07 to 0.39)</td>
<td>0.21 (0.06 to 0.36)</td>
</tr>
<tr>
<td>Self-care</td>
<td>0.56 (0.40 to 0.69)</td>
<td>0.40 (0.23 to 0.53)</td>
</tr>
<tr>
<td>Usual activity</td>
<td>0.56 (0.42 to 0.67)</td>
<td>0.48 (0.34 to 0.60)</td>
</tr>
<tr>
<td>Pain/Discomfort</td>
<td>0.50 (0.38 to 0.61)</td>
<td>0.25 (0.07 to 0.41)</td>
</tr>
<tr>
<td>Anxiety/Depression</td>
<td>0.43 (0.25 to 0.60)</td>
<td>0.34 (0.17 to 0.47)</td>
</tr>
<tr>
<td>EQ-VAS</td>
<td>0.51 (0.36 to 0.64)</td>
<td>0.44 (0.28 to 0.57)</td>
</tr>
<tr>
<td>VAS Pain</td>
<td>0.55 (0.44 to 0.65)</td>
<td>0.34 (0.14 to 0.50)</td>
</tr>
</tbody>
</table>

* p < 0.05, ** p < 0.01, *** p < 0.001
Table 3. Clinimetric features, hypotheses and conclusions for the DASH and the MHQ.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Hypothesis</th>
<th>DASH</th>
<th>MHQ</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Scale targeting</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Floor effect</td>
<td>Min score $&lt;15%$</td>
<td>Confirmed</td>
<td>Confirmed</td>
</tr>
<tr>
<td>No Ceiling effect</td>
<td>Max score $&lt;15%$</td>
<td>Confirmed</td>
<td>Confirmed</td>
</tr>
<tr>
<td><strong>Relatedness with</strong></td>
<td>Non-significant or negligible associations with age and sex</td>
<td>Confirmed / Rejected</td>
<td>Confirmed / Rejected</td>
</tr>
<tr>
<td>demographics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Construct validity</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Convergence between the PRO instruments</td>
<td>Significant and at least low correlation with each other</td>
<td>Confirmed</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Similar distributions for all recognized PCs</td>
<td>Confirmed</td>
<td></td>
</tr>
<tr>
<td>Relatedness with objective hand function measures</td>
<td>Significant and at least low correlation with grip and pinch forces</td>
<td>Confirmed</td>
<td>Confirmed</td>
</tr>
<tr>
<td>Relatedness with HRQoL</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correlation</td>
<td>Significant and at least low correlation with EQ-5D index</td>
<td>Confirmed</td>
<td>Confirmed</td>
</tr>
<tr>
<td>Regression</td>
<td>Significant and at least low linear dependency with EQ-5D index</td>
<td>Confirmed</td>
<td>Confirmed</td>
</tr>
</tbody>
</table>
Figure 1. Distributions of DASH and MHQ scores.
Figure 2. Correlation between DASH and MHQ.
Figure 3A-B. Correlations of the DASH and the MHQ with the EQ-5D index.
**Figure 4.** Age- and sex-adjusted regression coefficient $\beta$ of predictors of DASH and MHQ in units of SD. Boxes represent mean scores and whiskers show 95% CIs. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. 
**Figure 5.** DASH and MHQ scores in HRQoL subgroups presented as 0 to 100 scores. Division in subgroups is made by 25%, 50%, and 75% percentiles of EQ-5D index.
Figure 6. Loading vectors of reference outcome measures on PCs. The dots represent the participants.
Figure 7A-D. Relationships of DASH and MHQ scores with PC1 (A-B) and PC2 (C-D). The LOESS curve shows the deterministic part of the variation in the data. The grey area around the curve describes the 95% CIs.
References


17. Dias J, Rajan R and Thompson J. Which questionnaire is best? The reliability, validity and ease of use of the Patient Evaluation Measure, the Disabilities of the Arm, Shoulder and Hand


34. Atroshi I, Gummesson C, Andersson B, et al. The disabilities of the arm, shoulder and hand (DASH) outcome questionnaire: Reliability and validity of the Swedish version evaluated in
176 patients. *Acta Orthopaedica Scandinavica* 2000; 71: 613-618. DOI:


