The Patient-Rated Wrist Evaluation (PRWE): cross-cultural adaptation into German and evaluation of its psychometric properties

M. John[1], F. Angst[2,3], F. Awiszus[1], G. Pap[4], J.C. MacDermid[5], B.R. Simmen[2]

[1]Department of Orthopaedics, Otto-von-Guericke University, Magdeburg, Germany; [2]Schulthess Klinik, Zurich, Switzerland; [3]Rehaclinic Zurzach, Bad Zurzach, Switzerland; [4]Department of Orthopaedics and Traumatology, Park Hospital, Leipzig, Germany; [5]School of Rehabilitation Sciences McMaster University, Ontario, Canada.

Abstract

Objective

To cross-culturally adapt the Patient-Rated Wrist Evaluation form (PRWE) into German (PRWE-G) and to evaluate its reliability and validity.

Methods

A cross-cultural adaptation of the PRWE was carried out, according to established guidelines. 103 patients, who had undergone resection interposition arthroplasty (RIAP) for carpometacarpal osteoarthritis approximately 6.2 years earlier, completed a questionnaire booklet containing the PRWE-G, the Short Form 36 (SF-36), the Disabilities of Arm, Shoulder, and Hand (DASH); they also underwent clinical assessment with the Hand Function Index (HFI, Keitel) and Custom Score including grip and pinch strength tests. The results were used to assess the criterion and construct validity of the PRWE-G. To measure the re-test reliability, 51 patients completed a second PRWE-G within 2 weeks.

Results

The test-retest reliability of the PRWE-G was acceptable for the pain and function sub-scales and for the global score, with intraclass correlation coefficients of 0.78-0.87. The PRWE-G showed a high internal consistency (Cronbach's alphas of 0.92-0.97 for the scales and the total score). The typical error of measurement for the global score was 8.1 points, giving a minimal detectable change ($MDC_{95\%}$) of approximately 22.5 points. The PRWE-G scores correlated well with those of the DASH ($r=0.82$, $p<0.001$) but less well with those of the physical component summary of the SF-36 ($r=0.53$, $p<0.001$) and not at all with the mental component summary scores of the SF-36 ($r=0.04$, $p>0.05$). The PRWE-G scores correlated moderately with certain clinical findings of the HFI, Custom Score, and grip/pinch strength tests ($r=0.30$-0.59, $p<0.001$).

Conclusion

The PRWE-G represents a valid and reliable instrument to evaluate self-rated outcome in German-speaking patients with hand and wrist pathology.

Key words

PRWE, wrist, trapeziometacarpal, carpometacarpal arthritis, outcome, validity, reliability, self-assessment.

Michael John, MD
Felix Angst, MD, MPH
Friedemann Awiszus, MD, MSc
Géza Pap, MD
Joy C MacDermid, PhD
Beat R. Simmen, MD

This study was supported by the Georg and Bertha Schwyzer-Winiker Foundation (Zürich, Switzerland) and the Schulthess Klinik Research Funds.

Please address correspondence and reprint requests to:
Michael John, MD,
Orthopädische Universitätsklinik,
Otto-von-Guericke-Universität Magdeburg,
Leipziger Str. 44,
39120 Magdeburg,
Germany.
E-mail: michajohn100@web.de
or: michael.john@med.ovgu.de

Received on February 11, 2008;
accepted in revised form on May 30, 2008.

Conflict of interest:
Dr. J. MacDermid is supported by a New Investigator Award, Canadian Institutes of Health Research. The other co-authors have declared no competing interests.

Introduction

Osteoarthritis (OA) is the most prevalent form of arthritis and a major cause of pain and disability in people aged ≥65 years (1, 2). The prevalence is age-dependent and increases from 10% (in men) and 20% (in women) between the ages of 45 and 65 years to more than 50% in women aged 85 years and over (3). Approximately 30% of adults over 30 years of age present radiological evidence of hand OA (4). In a representative urban population aged 55-64 years, Kellgren and Lawrence found radiological signs of OA in the first carpometacarpal joint in 33.5% of women and 19.1% of men (5). Bade and Koebke reported nearly 50% of arthritic lesions in the saddle joint of the thumb in 100 anatomic specimens (6).

Clinically, in 50-90% of the cases, carpometacarpal arthritis is accompanied by scapho-trapezo-trapezoidal-osteoarthritis (6-8). Joint disease of the 1st ray leads to a restriction in function of the whole hand and wrist due to the special function of the thumb in opposing the fingers during pinching and grasping movements. Pain, functional disability and loss of strength are observed in 75-100% patients with carpometacarpal arthritis, and swelling, in approximately 57% (9).

The comprehensive concept of health, promoted by the World Health Organisation's International Classification of Functioning (ICF), has led to a reappraisal of the consequences of the disease on various important domains (10). In particular, more emphasis has been given to the individual's own rating of their limitations during the performance of everyday and social activities. The need for suitable patient-rated assessment methods has been further strengthened by the objectives and demands of evidence-based medicine.

The evaluation of outcome is of major importance in determining the quality and effectiveness of medical interventions (11), with a methodological approach to clinical practice being the quintessential element of evidence-based medicine. Many rheumatological and hand therapy societies and networks *e.g.,* IWI (International Wrist Investigators), OMERACT (Outcome Measures in Rheumatoid Arthritis Clinical Trials (11)) and AGREE (Appraisal of Guidelines for Research and Evaluation (12)) have endorsed this approach.

In addition to generic questionnaires used to assess quality of life (*e.g.*, the Short Form 36 (13)), other measurement instruments have been designed to evaluate functional capacity or disability at a joint, region, or symptom-specific level. The majority of the measurement instruments are in English and are intended for use in countries with the corresponding culture and language.

A number of patient-rated questionnaires are available for the hand/wrist. In addition to the more global Disability of the Arm, Shoulder, and Hand questionnaire (DASH) (14), the following have all been used in the subjective assessment of hand disorders: The Patient-Rated Wrist Evaluation (established in the anglo-american language region) (15, 16), the Patient-Rated Wrist/Hand Evaluation PRWHE (17), the Arthritis Impact Measurement Scales 1 & 2 (AIMS1/ AIMS2) (18), the Stanford Health Assessment Questionnaire (HAQ/ HAQ-DI, MHAQ) (19, 20), the Michigan Hand Outcomes Questionnaire (MHQ) (21), the Functional Index for Hand Osteoarthritis (FIHOA) (22), the Cochin Hand Functional Disability Scores (23) and the Australian/Canadian Osteoarthritis Hand Index (AUSCAN) (24).

The initiation of a quality management programme in our orthopaedic hospital, to evaluate surgical outcomes in the upper extremities unit, revealed the need for a German-version of a standardised questionnaire for the self-assessment of hand/wrist function. The DASH questionnaire already existed in German (25), but as this evaluates the symptoms and functional capacity of the whole upper extremity as one independent functional unit (14), it was considered too unspecific for our needs. The Patient Rated Wrist/Hand Evaluation PRWHE, formed by a simple modification of the PRWE with extension of the region-specific focus to the hand, was not available at the start of our quality assessment programme;

evidence of its responsiveness was subsequently published in 2004 (17). The PRWE represented the questionnaire of our choice for cross-cultural adaptation into German. It was not considered prudent to develop an entirely new questionnaire, since the consistent use of already established outcome instruments is essential to enable comparison/meta-analyses of results from studies carried out in different countries.

The aims of the present study were: 1) to carry out a cross-cultural adaptation into German of the English language version of the PRWE and 2) to examine the psychometric properties of the German PRWE within the framework of a long-term outcome study in patients who had undergone RIAP for carpometacarpal osteoarthritis.

Materials and methods

The Patient-Rated Wrist Evaluation form

The Patient-Rated Wrist Evaluation (PRWE) (15) is a patient self-rated, joint-specific questionnaire that enquires about symptoms of the wrist (presence, intensity and frequency of pain) and functional limitations in relation to activities of daily living (ADL). The response scale is a numeric rating scale from 0-10 points (0=best; 10=worst rating), comparable to a visual analogue scale in its metric characteristics. Approximately 3 minutes are required to answer all 15 questions. Five questions on pain (including one on pain frequency) and 10 questions on function allow the assessment of wrist status as a global score and/or separately as pain and function sub-scores. The 5 pain scale questions (0-10) yield a score between 0 and 50 points (0=no pain; 50=worst pain) and the 10 function scale questions, between 0 and 100 points (0=no difficulties whatsoever in carrying out activities; 100=carrying out activities is not possible). The function score is divided by 2, to allow calculation of a global score from 0-100 points, using the sum scores from the pain and function sub-scales. The reliability and validity of the English version of the PRWE was reported by MacDermid *et al.* (16), although no information was given regarding the total number of "missing" answers allowed. In the present study, PRWE global scores and subscale scores (pain, function) were calculated as long as 66.7% of the individual questions for each scale had been completed (*i.e.*, a minimum of 2/3), as recommended for other similar questionnaires (26). Hence, answering 3 pain questions and 7 function questions (maximum 5 unanswered questions) was the minimum requirement for calculating the global score.

Translation and cross-cultural adaptation of the PRWE

Translation and cross-cultural adaptation was performed according to the guidelines of the American Association of Orthopaedic Surgeons (AAOS) (27, 28). The following 6 steps were carried out and documented.

1. Forward translation from English into German by two independent native German speakers with different job profiles; a clinician familiar with hand scores T1 (informed translator) and a non-clinician T2 (uninformed translator).
2. Comparison of the two German versions with one another and with the English original. Synthesis to form one common translation (T-12) after reaching consensus in relation to any discrepancies.
3. Back-translation of T-12 into English by two bilingual English native speakers with German as a second language, to produce BT1 and BT2. A third bilingual person compared BT1 and BT2 with each other and with the original English questionnaire and highlighted any conceptual errors or inconsistencies in the translations.
4. Formation of a "pre-final" German version of the PRWE based on discussion within an expert committee of all of the forward and back-translations (T1, T2, BT1, BT2) and the reports made in carrying out/comparing the translations. During each step of the procedure all questions, problems and discrepancies were documented.
5. Testing of the pre-final version of the German PRWE (PRWE-G) in a heterogeneous group of 30 people, registering any questions or problems that arose. All the findings were evaluated by the expert committee before the final German version of the PREE was produced and subject to further psychometric testing.
6. Submission of the translated questionnaire and the documentation of the stages 1 to 5 to the developer(s) to ensure that the process was correctly carried out and a reasonable translation was achieved.

Psychometric properties of the German version of the PRWE (PRWE-G)

Subjects

Between 1996 and 1998, RIAP was performed in our hospital in 144 patients with symptomatic carpometacarpal arthritis. Between June and November 2003, these patients were invited to undergo a comprehensive follow-up cross-sectional catamnesis of their hand functional outcome. The average follow-up time at this point was 6.2 years (SD±0.83; Range: 4.6-7.8 years). Patients completed a questionnaire booklet containing the PRWE-G, the Short Form 36 (SF-36) (29), and the Disabilities of Arm, Shoulder, and Hand (DASH) (14, 25). The findings from a clinical assessment of the carpometacarpal joint and the hand were used to complete the clinical HFI of the Keitel Function Test (KFT) (30) and the Custom Form, a form specially designed to document the function of the finger joints and wrist (31). Joint range of motion was assessed with a goniometer, grip strength with a hand-grip dynamometer (JAMAR®, Sammons Preston, Bolingbrook, Illinois, USA), and key-pinch with a strength-measuring device.

Test-retest reliability

To examine the test-retest reliability of the PRWE-G, 65 patients were sent a second questionnaire by post, 3-4 days after completion of the first one.

Statistical analysis

The Kolmogorov-Smirnov-Test was used to examine the normality of the distribution of the data. Within-subjects effects (time and item number of

the PRWE-G) were examined using the repeated-measures analysis of variance (ANOVA).

Test-retest reliability was examined using the intraclass correlation coefficient (ICC), using the "alpha" and one-way random model. The calculation of the ICC was carried out for the global score, the two sub-scale scores and the individual items of the PRWE-G. The extent to which the repeated measurements led to varying results was determined by calculating the standard error of measurement (SEM) or "typical error" (TE) as an index of the absolute reliability (32). The reliability of the global scores and the sub-scores was also depicted graphically in the form of Bland and Altman curves, showing the limits of agreement (95% confidence interval) (33). Paired t-tests were used to examine the significance of the differences in mean values on each of the test days.

Spearman Rank Correlation coefficients were used to describe the interrelationships between the scores for the individual items of the pain and function scales, and between these individual scores and the sub-scales and the global score. In addition, the internal consistency of the questionnaire, a measure of the strength of the association of the items within the scales (pain, function) and the global index, was determined by Cronbach's alpha.

Correlation coefficients describing the strength of the relationship between the different questionnaire scores (SF-36 and DASH) were used to examine the criterion-related validity of the PRWE-G. The construct validity of the PRWE-G was further investigated by examining the relationship between the PRWE-G scores and those derived from the clinical joint assessment from the HFI/KFT and the Custom Form. Following the input of the results into a database, the various questionnaire scores were transformed to a scale from 0-10 points (0=worst, 10=best assessment). This facilitated comparison of the results of the different questionnaires, and comparison of the latter with those of other studies (31).

The statistical analyses were carried out using SPSS 15.0 for Windows® (SPSS Inc., Chicago, IL, USA). Statistical significance was accepted at the 5% level, but the precise *p*-values for each analysis are also given in the corresponding tables.

Results

Response rates

Complete data (questionnaire and clinical results) were obtained for 103 (71.5%) patients with 112 treated carpometacarpal joints (9 both sides). Forty-one patients could not completely be assessed for the following reasons: 4 (2.8%) patients had died; 11 (7.6%) patients had changed address and could not be traced; and 19 (13.2%) patients declined participation (5 patients lived too far away from the clinic (>1000 km) to attend, 5 patients had multiple morbidity, and 9 patients simply did not wish to participate). The data of 7 (4.9%) patients were highly incomplete and had to be excluded.The average age of the 85 female (93 joints) and 18 male (19 joints) patients was 67.7 years (SD=9.8 years).

Test-retest reliability was examined on 51 PRWE joint assessments (39 women, 12 men). The second questionnaire was returned on average 14 days after the first one was completed. 9 patients did not return their questionnaire, and 5 patients did not fill out the PRWE questionnaire sufficiently completely.

Analysis of completeness of the questionnaires (from 93 patients; 112 joints), in relation to the individual items of the PRWE, showed the following distribution: items 1, 2, 3, 8, 12 and 13 were answered by more than 95%, the remaining 9 questions by 94.6% to 88.4% of the patients.

Cross-cultural adaptation process

Overall, the translation and back-translation of the questionnaire proceeded without any major difficulties. Only few questions required discussion or modification by the committee as a result of notable deviations from the original. Consensus was always found within the committee regarding any questionable wording.

In the English version, the introduction to the questionnaire instructs the patient to leave the response blank if he has never performed the activity being enquired about. However, the expert committee considered that this could lead to confusion in the later interpretation of the data (inability to determine whether the question was indeed not applicable, or had simply been overlooked/missed), and hence suggested that a further response option "not applicable" be included for items 6 to 15.

The back-translation revealed considerable discrepancies with the original English version for 4 items. For item 8 "Fasten buttons on my shirt" it was decided to dispense with the notion of the buttoning being done on a shirt that was being worn, since this ability is not only dependent on hand function but also elbow and shoulder function. Hence, this was simply translated as the equivalent of "Buttoning a shirt or blouse". The function item 9: "Use my affected hand to push up from a chair" caused problems during translation into German, on account of the two possible ways of translating the word chair ("Stuhl" or "Sessel"). In accordance with the decision made for dealing with this in the adaptation of the PREE (34) we chose the option "Stuhl". In the English version of the PRWE, item 10 in the 'function' section was formulated with imperial units of weight ("Carry a 10-pound object in my affected hand"); this was converted to the nearest whole kilogram in forming the German version (=5 kg). The second part of the function score: "usual activities" presented an unexpected problem. The 2 German forward translations were "Normale Tätigkeit" and "Alltägliche Tätigkeit", with the latter being chosen in the consensus version. This was back-translated into English as "Daily Activity" and "Everyday Activities". The expert committee considered that this term could present problems, in terms of whether the activity was literally carried out "every day". As such, the term "gewöhnliche Tätigkeiten" was chosen for the final German version, as a less ambiguous alternative.

In order to avoid the difficulties already experienced in adapting the PREE (34), in relation to the performance of household duties (item 13 in the PRWE and item 18 in the PREE), the words

"kleine Reparaturen" were added to the list of examples for this item, in order to address "gender-specific" duties.

The testing of the pre-final version of the German PRWE (PRWE-G) did not present any relevant problems necessitating further changes to the questionnaire. Two people suggested that the function question "use bathroom tissue with my affected hand" may be problematic. One patient interpreted the question as only meaning the removal/rolling off of paper from the toilet roll, whilst another mentioned that as a left-hander with an affected right hand, this would still not present any difficulty (this served to confirm the utility of the "not applicable" response option for this and other functions done only with the dominant hand).

The cross-cultural adaptation process and the final version were submitted to, discussed with and finally approved by J. C. MacDermid, who developed the original English PRWE in 1996 (16). The final version of the PRWE-G is shown in the Appendix.

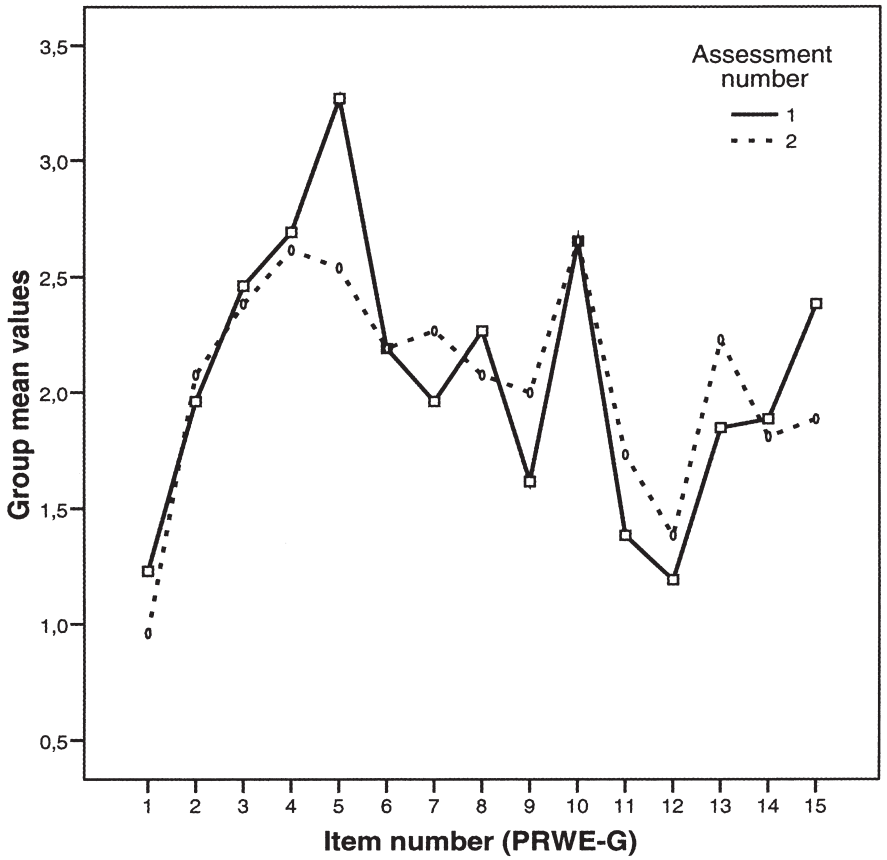

Fig. 1. Mean values for each item (1-15) for the first and the second (re-test) PRWE-G assessment.

Psychometric properties of the PRWE-G

Distribution of the scores

None of the PRWE-G scores or individual item scores were normally distributed (Kolmogorov-Smirnov-Test, $p<0.005$). 22/112 (19.6%) joints showed a perfect outcome (rating=0) on the pain sub-scale, 27/112 (24.1%) on the function scale and 18/112 (16.1%) on the global PRWE-G scale. These proportions (approx 16-24%) represent the ceiling effects for the various scales. The scores were heavily skewed to the left, with a score of between 0 and 2 (on the 0-10 scale) being achieved by 57.4% of the patients for pain, 61% for function and 53.9% for the global score. These PRWE-G results reflect minimal problems in relation to the wrist or carpometacarpal joint at the long-term follow-up in this patient group.

Inter-item correlations and internal consistency

Cronbach's alpha values of 0.92-0.97 indicated a high internal consistency for the PRWE-G (pain items: 0.92; function items: 0.96; total score: 0.97).

Test-retest reliability

For the within-subjects analysis of variance of the PRWE-G scores (assessment time point, item number, time point x item number (interaction)), only the item number was significant ($p<0.001$), *i.e.*, the mean scores for the items differed significantly from each other, but there was no significant difference between the test and re-test mean values. A plot of the mean values for each item at the two time points (test and re-test) showed largely similar curves (Fig. 1). If the curves had been perfectly superimposed upon one another, this would have indicated identical results on each occasion.

The curves indicate a graded and differentiated mean rating for each item, even for items within the same sub-section. The items hence appear to show little overlap in terms of their content. Conspicuous differences between the curves could be seen in relation to the mean values generated on the two test occasions for items 5, 9, 13 and 15. The repeated assessment yielded better ratings for item 5 and 15 (*i.e.*, less pain and difficulties) and worse ratings for item 9 and 13 (*i.e.*, more difficulties). At re-test, the mean score for the pain sub-scale was significantly lower (less pain; $p<0.05$), whereas the mean function score and the global score each showed no significant differences compared with their values in the first assessment.

The ICCs (indicating test-retest reliability) ranged from 0.69 to 0.87 for 14 of the individual items, the separate pain and function sub-scales, and the total score; the ICC was 0.42 for item 1 (Table I).

The error of measurement of the PRWE-G, given by the TE, was relatively high (Table II). The pain scale, in particular, was associated with relatively large measurement error. Figure 2 (Bland-Altman-Plots) shows the limits of agreement for the repeated scores of the PRWE-G; the best results (narrowest 95% confidence interval; mean: -1.45±CI: 11.48*1.96) were for the global score and the function sub-score (mean: 1.13±CI: 11.05*1.96).

Table I. Intraclass correlation coefficients (ICC) for the test-retest reliability of the PRWE-G with 95% confidence intervals (n=51).

Pain scale			Function scale					
Item	ICC	95% C.I.	Item	ICC	95% C.I.	Item	ICC	95% C.I.
1	0.42	0.18-0.62	6	0.78	0.64-0.87	11	0.86	0.76-0.92
2	0.74	0.58-0.84	7	0.69	0.50-0.81	12	0.85	0.74-0.91
3	0.79	0.66-0.87	8	0.79	0.67-0.88	13	0.69	0.52-0.81
4	0.75	0.59-0.85	9	0.77	0.62-0.86	14	0.84	0.72-0.91
5	0.72	0.56-0.83	10	0.83	0.72-0.90	15	0.73	0.56-0.84
Pain sub-scale score:			Function sub-scale score:					
	0.78	0.64-0.87		0.87				0.78-0.92
Global score:				0.86				0.76-0.92

Table II. Precision of measurement for the PRWE-G. Values relate to the results for the comparison of test and re-test values.

	Mean difference	Standard deviation	Typical error	95% C.I. #
Pain	-3.99	14.91	10.54	-33.92-25.94
Function	1.13	11.05	7.81	-21.06-23.32
Total	-1.45	11.48	8.12	-24.49-21.59

C.I. #confidence interval (Mean (Δ) $\pm$ $t_{50,\ 0,025}$ (2,00744768) x SD).

Correlations between the PRWE-G scores and other questionnaire scores (Table III)

The scores for the PRWE-G subscales (pain, function) and the global score each showed moderate to high correlations with the corresponding DASH scores (r=0.70-0.82, p<0.001).

The scores of the PRWE showed moderate correlations (0.32-0.54, p<0.001) with those of various physical scales of the SF-36. None or only low correlations were found for the PRWE-G scores and the mental scores of the SF-36 (Table III).

The PRWE scores (sub-scales and whole score) correlated moderately with hand-grip strength, pinch-grip strength, and all the Custom Scores apart from Custom Deformity and ROM (r=0.39-0.59, p<0.001); the correlations with the HFI were low but nonetheless significant (r=0.30-0.35, p<0.005).

Discussion

The cross-cultural adaptation of the PRWE into the PRWE-G provided an outcome instrument for the hand/wrist, which is to our knowledge the only one available in the German language.

The translation and adaptation of the PRWE were carried out in relation to the language and cultural conditions that prevail in Switzerland and were tested on a German-Swiss group of patients. Although there are considerable differences in "spoken" German within the German-speaking countries/regions in Europe, there are no grammatical or semantic differences in the use of the written language. Thus, with the only prerequisite for the use of our PRWE-G being an adequate understanding of written German, we believe that the PRWE-G can be used in other German-speaking European countries.

The successful pre-testing phase of the translated version confirmed the face/content validity of the PRWE-G. The closeness of the English and German languages meant that the translation could be carried out with mostly literal equivalence. The test-retest reliability of the PRWE-G was moderate to high showing almost those ICCs as reported by MacDermid *et al*. In contrast to that study, we had a stable but less diverse patient-group, *i.e.,* 2 had more variability (16), and we chose a 2-week time interval to minimise the possible memory effect and to provide a more realistic view of the degree of score change that may occur for non-specific reasons (random error) (35). In agreement with our earlier study on the PREE (34), the ICCs for the individual pain questions and the pain sub-scales of the PRWE tended to be lower than those of the function questions/scales. It has been proposed that the ICC should be >0.70 for reliable group comparisons and 0.90-0.95 for reliable comparisons of individual measurements over time (36). This suggests that all the PRWE-G subscales and most of the individual items can be safely recommended for future prospective studies of group change, but that changes in individual measurements recorded over time are to be interpreted with caution.

The "typical error" (TE), expressed as a percentage of the mean value for the score (pain 13.9%, function 9.7%, total 10.4%), was comparable to that reported for other joint-specific musculoskeletal outcome measures (37, 38). The TE is an expression of the precision of a measurement instrument and hence, also, its sensitivity to change. The TE can be used to indicate the "minimum detectable change" ($MDC_{95\%}$) for the instrument, *i.e.*, the degree of change required in an individual's score, in order to establish it (with a given level of confidence) as being a "real change" over and above measurement error (39). At the 95% confidence level, this is defined as $1.96*\sqrt{2}*TE$. With a TE of 8.12 for the PRWE-G (total score), this would suggest that changes of the order of 22.5 points (out of 100) would be required to state with confidence that real change in a patient has occurred; this again calls for caution in the interpretation of individual PRWE-G score-changes over time.

The internal consistency of the PRWE-G was high (Chronbach's alpha=0.97). Values of >0.70 are typically strived for (36) in order to demonstrate that the individual items are sufficiently homogeneous in relation to the instrument's content and intended construct. Coefficients greater than 0.95 must be viewed critically as they may indicate redundancy in the instrument's content and hence threaten its content validity (40).

The PRWE-G scores in the patient

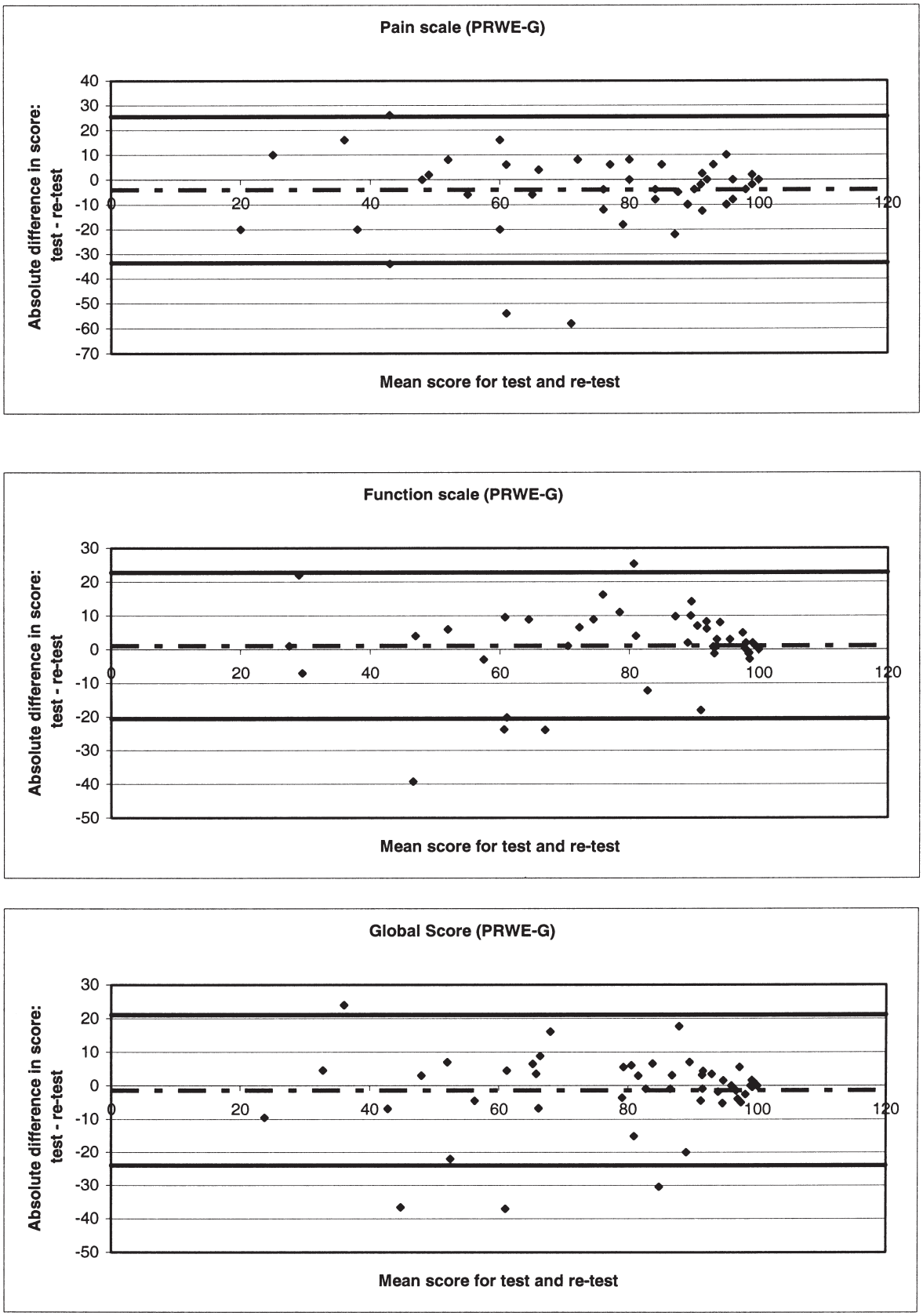

Fig. 2. Bland-Altman Plots for the pain sub-scale, function sub-scale and the global score of the PRWE-G (n=51 patients). In each case, the plots show the mean value for the test and re-test scores (x axis) against the absolute difference between the test and re-test scores (y axis). The dotted horizontal line indicates the mean difference (systematic bias) between the two trials, and the solid horizontal lines indicate the 95% confidence interval (SD x 1.96) *i.e.*, the limits of agreement.

Table III. Strength of association between the scores from the various questionnaires used to assess validity of the PRWE-G. (Spearman's rank – correlation).

PRWE-G	*Pain* r	Function r	Total r
SF-36 Physical functioning	**0.41*****	**0.46*****	**0.46*****
SF-36 Role physical	**0.32*****	**0.42*****	**0.39*****
SF-36 Bodily pain	**0.52*****	**0.51*****	**0.54*****
SF-36 General Health	**0.33*****	**0.34*****	**0.36*****
SF-36 Vitality	**0.36*****	**0.28****	**0.32****
SF-36 Social functioning	**0.26****	**0.28****	**0.28****
SF-36 Role emotional	**0.24****	**0.18****	**0.19***
SF-36 Mental health	**0.26****	**0.21****	**0.21***
SF-36 PCS	**0.47*****	**0.54*****	**0.53*****
SF-36 MCS	0.12	0.03	0.04
DASH Symptoms	**0.70*****	**0.72*****	**0.75*****
DASH Function	**0.68*****	**0.80*****	**0.80*****
DASH	**0.72*****	**0.81*****	**0.82*****
KFT/ HFI	**0.30****	**0.31****	**0.35*****
Custom ROM	0.16	**0.20***	**0.25***
Custom Deformity	**0.17***	**0.18***	**0.21***
Custom Grip Strength	**0.39*****	**0.51*****	**0.49*****
Custom Pinch Strength	**0.46*****	**0.48*****	**0.50*****
Custom Whole Score	**0.46****	**0.59*****	**0.57*****

Significant correlations in bold: $p<0.001^{***}$, $p<0.01^{**}$, $p<0.05^{*}$

group examined were not normally distributed and showed rather high ceiling effects. The left-skewed distributions observed in the current study are likely the result of the long-term nature of the follow-up and the typically good results associated with RIAP. Other authors have also reported over 80% satisfied and/or pain free patients at long-term follow-up (9, 31). The low variability, caused by the good PRWE-G score results given, would tend to lead to lower ICCs (Table I).

Correlation coefficients showing the strength of correlation between the PRWE-G and those of the comparision questionnaires and various clinical parameters were already given by Angst *et al*. for the total score results only (31). We presented the missing data for the sub-scale results of the given instruments. The criterion and construct validity of the PRWE-G was good and almost comparable to that of the original PRWE (15, 16). Comparable to the reported results by MacDermid *et al*. (16) for a patient group being treated for scaphoid-pseudarthrosis we found sizeable correlations of the PRWE-G subscales and global score with the physical dimensions of the SF-36 (16). Similar to the validation results of the German DASH (25) only low correlations existed between the PRWE-G scales and the mental dimensions of the SF-36. This stands in contrast to the findings of MacDermid *et al*. (16).

Moderate to high correlations were found between the DASH and PRWE-G for both their total scores and their respective sub-scales. The strength of the correlation between the scores on the two instruments indicates a large degree of overlap in relation to the questionnaire content. Discrepancies in the outcomes as measured with the two instruments may arise due to deficits in the neighbouring shoulder and elbow joints. MacDermid *et al*. reported that the PRWE, compared with the DASH, had a higher joint-related specificity (41), indicating that they are not delivering exactly the same information.

The clinical parameter strength showed a moderate correlation with the PRWE-G subscales/score which was the strongest correlation result among the other clinical joint assessments. Our findings are in accordance with MacDermid *et al*. (16, 41, 42), who also found slight to moderate correlations between patient-rated function and clinically determined parameters. Karnezis and Fragkiadakis (43) demonstrated that handgrip strength but not ROM of the wrist is a significant predictor of the PRWE score. Furthermore, PRWE results of patients with distal radius fracture also correlate strongly with the time lost from work (RTW=return to work) (47). However, patient satisfaction with treatment correlates more highly with subjective ratings of pain alleviation and improvement in ADL than with any "objective" clinical or radiological measurements (41, 42).

The PRWE is a standardized, self-assessment questionnaire designed for the evaluation of wrist impairment and disability. It allows a brief, reliable and valid measurement of wrist-related symptoms independent from the underlying disease/cause of the disorder (15, 16). The development process of the PRWE used surveys of 100 specialists in wrist treatment (IWI) and patient interviews (15, 16). This ensured not only the content validity of the PRWE but also a maximum of its clinical appropriateness and practicality. The PRWE is more responsive than the generic SF-36 and even more sensitive than the DASH in wrist-specific pathologies (41). The PRWE is used on an international basis (*i.e*.,15, 16, 31, 41, 43-48) and has already been cross-culturally adapted for other languages (49, 50). A search of the literature in PubMed using the key words "PRWE" or "patient-rated wrist evaluation" revealed more than 60 publications dealing with or using this instrument since its first publication in 1996. A short questionnaire with easily understandable questions and no requirement of special instruments makes this tool valuable for the clinical hand practise. The questionnaire is quick to complete (2-3 minutes) for the patient and easy to evaluate for the investigator.

The PRWE was originally developed and validated in patients with radial fracture and has mostly been used as a joint-specific self-assessment instrument to measure the effects of treatment after fracture of the radius (43, 45, 46). It has only rarely been used as a measurement instrument in non-traumatic disorders of the wrist, *e.g*., in osteoarthritis of the carpometacarpal joint (31, 47) and proximal row carpectomy (48). We were able to

demonstrate the reliability and validity of the PRWE-G in patients with a joint disorder in the area of the metacarpus. In her original studies, MacDermid and colleagues (15, 16) examined patients with pseudarthrosis of the scaphoid. The results of the present study, on patients previously treated with RIAP, did not differ in any notable way to those presented in the original PRWE (16, 41). When confronted with a questionnaire, patients will automatically consider its content in relation to the symptomatic or recently treated (operated) region of the hand/wrist as a "global entity". The development of the PRWHE, by simple adaptation of the PRWE to include the "wrist/hand" instead of just the wrist, illustrates the case. The PRWHE, with its identical wording to the PRWE, represents a region-specific measurement instrument for traumatised hands/wrists, simply by virtue of a modification of the anatomical region under focus. Its responsiveness is largely equivalent to that of the PRWE (16, 17, 41, 42). Hence, we believe that a German version of the PRWHE (PRWHE-G) can be created and used in clinical practice, through simple modification of the instructions from "wrist" to "wrist/hand", without further testing of its reliability and validity.

Since the English version of the PRWE has been shown to be a sensitive instrument, we would expect the same qualities from the PRWE-G but examination of the responsiveness of the PRWE-G will be the subject of future prospective studies.

In conclusion, the present study has served to provide a reliable and valid German version of the PRWE, and thereby extends the list of self-assessment questionnaires currently available for use with groups of German-speaking patients. This fulfils the increasing demand for international, standardised measurement instruments to assess health and quality of life and the efficacy of medical interventions for the purposes of outcomes research and quality management programmes.

Acknowledgments

We thank Hans-Kasper Schwyzer for providing the infrastructure to carry out the study, and Roberta Schefer for the management of the patients, the questionnaires, and the database. We are grateful to Anne Mannion for directing and co-ordinating the cross-cultural adaptation of the questionnaire, for her support and guidance during the study, and for her thorough content and language editing of the manuscript. We are also grateful to Astrid Junge for her help with the translation of the PRWE into German, and to Kathrin Knecht and Gordana Balaban for carrying out the pre-testing phase of the cross-cultural adaptation. We thank Charles McCammon for his help in preparing the manuscript.

References

1. BUCKWALTER JA, SALTZMAN C, BROWN T: The impact of osteoarthritis: implications for research. *Clin Orthop Relat Res* 2004; 427S: 6-15.
2. LAWRENCE RC, HELMICK CG, ARNETT FC *et al*.: Estimates of the prevalence of arthritis and selected musculoskeletal disorders in the United States. *Arthritis Rheum* 1998; 41: 778-99.
3. MARCH LM, BAGGA H: Epidemiology of osteoarthritis in Australia. *Med J Aust* 2004; 180: 6-10.
4. GARSTANG SV, STITIK TP: Osteoarthritis: epidemiology, risk factors, and pathophysiology. *Am J Phys Med Rehabil* 2006; 85: 2-11.
5. KELLGREN JH, LAWRENCE JS: Osteo-arthrosis and disk degeneration in an urban population. *Ann Rheum Dis* 1958; 17: 388-97.
6. BADE H, KOEBKE J: Arthrosis in the peritrapezial joints of the hand. *Surg Radiol Anat* 1996; 18: 209-14.
7. ARMSTRONG AL, HUNTER JB, DAVIS TR: The prevalence of degenerative arthritis of the base of the thumb in post-menopausal women. *J Hand Surg* (Br) 1994; 19: 340-41.
8. BHATIA A, PISOH T, TOUAM C *et al*.: Incidence and distribution of scaphotrapezotrapezoidal arthritis in 73 fresh cadaveric wrists. *Ann Chir Main Memb Super* 1996; 15: 220-5.
9. WITTEMANN M, DEMIR E, SAUERBIER M *et al*.: The Epping resection-suspension arthroplasty procedure. A standard procedure in the operative treatment of trapeziometacarpal osteoarthrosis?. *Handchir Mikrochir Plast Chir* 2002; 34: 49-58.
10. WORLD HEALTH ORGANISATION (WHO): ICF – International Classification of Functioning, Disability and Health. ed. *World Health Organisation Geneva*, Switzerland 2001: 10-20.
11. BOERS M, BROOKS P, SIMON LS *et al*.: An international initiative to improve outcome measurement in rheumatology. *Clin Exp Rheumatol* 2005; 23: 10-3.
12. MACDERMID JC, BROOKS D, SOLWAY S *et al*.: Reliability and validity of the AGREE instrument used by physical therapists in assessment of clinical practice guidelines. *BMC Health Serv Res* 2005; 5: 18.
13. WARE JE, SNOW KK, KOSINSKI M *et al*.: *SF-36 Health survey: Manual and interpretation guide*. 2nd ed. Lincoln (RI): QualityMetric Incorporated 2000.
14. HUDAK PL, AMADIO PC, BOMBARDIER C: Development of an upper extremity outcome measure: the DASH (disabilities of the arm, shoulder and hand) [corrected]. The Upper Extremity Collaborative Group (UECG). *Am J Ind Med* 1996; 29: 602-8.
15. MACDERMID JC: Development of a scale for patient rating of wrist pain and disability. *J Hand Ther* 1996; 9: 178-83.
16. MACDERMID JC, TURGEON T, RICHARDS RS *et al*.: Patient rating of wrist pain and disability: a reliable and valid measurement tool. *J Orthop Trauma* 1998; 12: 577-86.
17. MACDERMID JC, TOTTENHAM V: Responsiveness of the disability of the arm, shoulder, and hand (DASH) and patient-rated wrist/hand evaluation (PRWHE) in evaluating change after hand therapy. *J Hand Ther* 2004; 17: 18-23.
18. MEENAN RF, MASON JH, ANDERSON JJ *et al*.: The content and properties of a revised and expanded Arthritis Impact Measurement Scales Health Status Questionnaire. *Arthritis Rheum* 1992; 35: 1-10.
19. BRUCE B, FRIES JF: The Health Assessment Questionnaire (HAQ). *Clin Exp Rheumatol* 2005; 23: 14-18.
20. FRIES JF, SPITZ P, KRAINES RG *et al*.: Measurement of patient outcome in arthritis. *Arthritis Rheum* 1980; 23: 137-45.
21. CHUNG KC, PILLSBURY MS, WALTERS MR *et al*.: Reliability and validity testing of the Michigan Hand Outcomes Questionnaire. *J Hand Surg* (Am) 1998; 23: 575-87.
22. DREISER RL, MAHEU E, GUILLOU GB *et al*.: Validation of an algofunctional index for osteoarthritis of the hand. *Rev Rhum Engl Ed* 1995; 62: 43-53.
23. DURUOZ MT, POIRAUDEAU S, FERMANIAN J *et al*.: Development and validation of a rheumatoid hand functional disability scale that assesses functional handicap. *J Rheumatol* 1996; 23: 1167-72.
24. BELLAMY N, CAMPBELL J, HARAOUI B *et al*.: Dimensionality and clinical importance of pain and disability in hand osteoarthritis: Development of the Australian/Canadian (AUSCAN) Osteoarthritis Hand Index. *Osteoarthritis Cartilage* 2002; 10: 855-62.
25. OFFENBAECHER M, EWERT T, SANGHA O *et al*.: Validation of a German version of the disabilities of arm, shoulder, and hand questionnaire (DASH-G). *J Rheumatol* 2002; 29: 401-2.
26. ANGST F, PAP G, MANNION AF *et al*.: Comprehensive assessment of clinical outcome and quality of life after total shoulder arthroplasty: usefulness and validity of subjective outcome measures. *Arthritis Rheum* 2004; 51: 819-28.
27. GUILLEMIN F, BOMBARDIER C, BEATON D: Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol* 1993; 46: 1417-32.
28. BEATON DE, BOMBARDIER C, GUELLEMIN

F *et al*.: Guidelines for the process of cross-cultural adaption of self-report measures. *Spine* 2000; 25: 3186-91.
29. BULLINGER M, KIRCHBERGER I: SF-36 Fragebogen zum Gesundheitszustand. Handanweisung. (The SF-36 questionnaire to assess health status. A manual). Göttingen, Germany: Hogrefe, 1998.
30. KEITEL W, HOFFMANN H, WEBER G *et al*.: [Evaluation of the percentage of functional decrease of the joints using a motor function test in rheumatology]. *Dtsch Gesundheitsw* 1971; 26: 1901-3.
31. ANGST F, JOHN M, GOLDHAHN J *et al*.: Comprehensive assessment of clinical outcome and quality of life after resection interposition arthroplasty of the thumb saddle joint. *Arthritis Rheum* 2005; 53: 205-13.
32. HOPKINS WG: Measures of reliability in sports medicine and science. *Sports Med* 2000; 30: 1-15.
33. BLAND JM, ALTMAN DG: Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; 1: 307-10.
34. JOHN M, ANGST F, PAP G *et al*.: Cross-cultural adaptation, reliability and validity of the Patient Rated Elbow Evaluation (PREE) for German-speaking patients. *Clin Exp Rheumatol* 2007; 25: 195-205.
35. BEURSKENS AJ, DE VET HC, KOKE AJ *et al*.: Measuring the functional status of patients with low back pain. Assessment of the quality of four disease-specific questionnaires. *Spine* 1995; 20: 1017-28.
36. LOHR KN: Assessing health status and quality-of-life instruments: attributes and review criteria. *Qual Life Res* 2002; 11: 193-205.
37. ANGST F, GOLDHAHN J, PAP G *et al*.: Cross-cultural adaptation, reliability and validity of the German Shoulder Pain and Disability Index (SPADI). *Rheumatology* (Oxford) 2007; 46: 87-92.
38. MANNION AF, JUNGE A, FAIRBANK JC *et al*.: Development of a German version of the Oswestry Disability Index. Part 1: cross-cultural adaptation, reliability, and validity. *Eur Spine J* 2006; 15: 55-65.
39. BEATON DE: Understanding the relevance of measured change through studies of responsiveness. *Spine* 2000; 25: 3192-9.
40. MICHENER LA, LEGGIN BG: A review of self-report scales for the assessment of functional limitation and disability of the shoulder. *J Hand Ther* 2001; 14: 68-76.
41. MACDERMID JC, RICHARDS RS, DONNER A *et al*.: Responsiveness of the short form-36, disability of the arm, shoulder, and hand questionnaire, patient-rated wrist evaluation, and physical impairment measurements in evaluating recovery after a distal radius fracture. *J Hand Surg* (Am) 2000; 25: 330-40.
42. MACDERMID JC, ROTH JH, RAMPERSAUD YR *et al*.: Trapezial arthroplasty with silicone rubber implantation for advanced osteoarthritis of the trapeziometacarpal joint of the thumb. *Can J Surg* 2003; 46: 103-10.
43. KARNEZIS IA, FRAGKIADAKIS EG: Association between objective clinical variables and patient-rated disability of the wrist. *J Bone Joint Surg Br* 2002; 84: 967-70.
44. MACDERMID JC, ROTH JH, MCMURTRY R: Predictors of time lost from Work following a distal radius fracture. *J Occup Rehabil* 2007; 17: 47-62.
45. ADAMS BD, GROSLAND NM, MURPHY DM *et al*.: Impact of impaired wrist motion on hand and upper-extremity performance (1). *J Hand Surg* (Am) 2003; 28: 898-903.
46. KARNEZIS IA, PANAGIOTOPOULOS E, TYLLIANAKIS M *et al*.: Correlation between radiological parameters and patient-rated wrist dysfunction following fractures of the distal radius. *Injury* 2005; 36: 1435-39.
47. MACDERMID JC, WESSEL J, HUMPHREY R *et al*.: Validity of self-report measures of pain and disability for persons who have undergone arthroplasty for osteoarthritis of the carpometacarpal joint of the hand. *Osteoarthritis Cartilage* 2007; 15: 524-30.
48. DE SMET L, ROBIJNS F, DEGREEF I: Outcome of proximal row carpectomy. *Scand J Plast Reconstr Surg Hand Surg* 2006; 40: 302-6.
49. VOCHE P, DUBERT T, LAFFARGUE C *et al*.: Patient-rated wrist questionnaire: preliminary report on a proposed French version of a North American questionnaire designed to assess wrist pain and function. *Rev Chir Orthop Reparatrice Appar Mot* 2003; 89: 443-48.
50. WAH JW, WANG MK, PING CL: Construct validity of the Chinese version of the Patient-rated Wrist Evaluation Questionnaire (PRWE-Hong Kong Version). *J Hand Ther* 2006; 19: 18-26.

Appendix

German version of the PRWE (PRWE-G)-Score (Patient-Rated Wrist Evaluation)

Die unten aufgeführten Fragen sollten uns helfen, das Ausmass der Schwierigkeiten, die Sie wegen Ihres Handgelenks in der letzten Woche hatten, zu verstehen. Sie werden gebeten, auf einer Skala von 0-10 anzugeben, wie stark Ihre Handgelenksbeschwerden bei den folgenden Tätigkeiten in der letzten Woche durchschnittlich waren.
Bitte beantworten Sie ALLE Fragen. Wenn Sie eine (oder mehrere) der Tätigkeiten in der letzten Woche nicht ausgeführt haben, SCHÄTZEN Sie bitte das Ausmass der Schmerzen oder Schwierigkeiten ein, die Sie erwarten würden. Wenn Sie eine Tätigkeit niemals ausgeführt haben, kreuzen Sie bitte jeweils, „trifft nicht zu“ an.

SCHMERZEN
Bitte geben Sie die durchschnittliche Stärke der Schmerzen in Ihrem Handgelenk in der letzten Woche an, indem Sie die Zahl auf der Skala von 0 bis 10 ausfüllen, die Ihre Schmerzen am besten beschreibt. Null (0) bedeutet, dass Sie keinerlei Schmerzen hatten and Zehn (10) bedeutet, dass Sie die schlimmsten Schmerzen hatten, die Sie je erlebt haben oder, dass Sie die Tätigkeit aufgrund der Schmerzen nicht ausführen konnten.

	0 1 2 3 4 5 6 7 8 9 10
	keinerlei Schmerzen — schlimmste Schmerzen
Bitte geben Sie die Stärke Ihrer Schmerzen an:	
1. In Ruhe	0 1 2 3 4 5 6 7 8 9 10
2. Bei Tätigkeiten mit wiederholter Bewegung des Handgelenks	0 1 2 3 4 5 6 7 8 9 10
3. Beim Heben eines schweren Gegenstandes	0 1 2 3 4 5 6 7 8 9 10
4. Wenn sie am stärksten sind	0 1 2 3 4 5 6 7 8 9 10
	Nie — Immer
5. Wie häufig haben Sie Schmerzen?	0 1 2 3 4 5 6 7 8 9 10

FUNKTION

Funktionsfähigkeit - Bestimmte Tätigkeiten

Bitte geben Sie an, wie viele Schwierigkeiten Sie in der letzten Woche bei jeder der unten aufgeführten Tätigkeiten hatten. Bitte kreuzen Sie die Zahl von 0-10 an, die das Ausmass Ihrer Schwierigkeiten am besten beschreibt.
Null (0) bedeutet, dass Sie keinerlei Schwierigkeiten hatten and Zehn (10) bedeutet, dass die Schwierigkeiten so gross waren, dass Sie die Tätigkeit nicht ausführen konnten.
Wenn Sie eine Tätigkeit niemals ausgeführt haben, kreuzen Sie bitte den Punkt „trifft nicht zu“ an.

0	1	2	3	4	5	6	7	8	9	10	
keinerlei Schwierigkeiten										nicht möglich	trifft nicht zu
6. Mit der betroffenen Hand einen Türknauf drehen											
0	1	2	3	4	5	6	7	8	9	10	❑
7. Mit der betroffenen Hand Fleisch mit dem Messer schneiden											
0	1	2	3	4	5	6	7	8	9	10	❑
8. Ein Hemd oder eine Bluse zuknöpfen											
0	1	2	3	4	5	6	7	8	9	10	❑
9. Die betroffene Hand nutzen, um von einem Stuhl aufzustehen											
0	1	2	3	4	5	6	7	8	9	10	❑
10. Mit der betroffenen Hand einen 5 kg schweren Gegenstand tragen											
0	1	2	3	4	5	6	7	8	9	10	❑
11. Mit der betroffenen Hand das Toilettenpapier benutzen											
0	1	2	3	4	5	6	7	8	9	10	❑

Funktionsfähigkeit - Gewöhnliche Tätigkeiten

Bitte geben Sie an, wie viele Schwierigkeiten Sie in der letzten Woche bei Ihren gewöhnlichen Tätigkeiten in jedem der unten genannten Bereiche hatten. Bitte füllen Sie die Zahl von 0 bis 10 aus, die das Ausmass Ihrer Schwierigkeiten am besten beschreibt. Unter „gewöhnliche Tätigkeiten“ verstehen wir die Aktivitäten, die Sie ausführten, bevor die Probleme mit Ihrem Handgelenk begannen. Null (0) bedeutet, dass Sie keinerlei Schwierigkeiten hatten and Zehn (10) bedeutet, dass Sie so große Schwierigkeiten hatten, dass Sie keine dieser gewöhnlichen Tätigkeit ausführen konnten. Wenn Sie eine Tätigkeit niemals ausgeführt haben, kreuzen Sie bitte den Punkt „trifft nicht zu“ an.

0	1	2	3	4	5	6	7	8	9	10	
keinerlei Schwierigkeiten										nicht möglich	trifft nicht zu

12. Persönliche Körperpflege (Anziehen, Waschen)

0	1	2	3	4	5	6	7	8	9	10	❑

13. Hausarbeit (Putzen, Aufräumen, kleine Reparaturen)

0	1	2	3	4	5	6	7	8	9	10	❑

14. Arbeit (Beruf oder Alltagstätigkeiten)

0	1	2	3	4	5	6	7	8	9	10	❑

15. Freizeitaktivitäten

0	1	2	3	4	5	6	7	8	9	10	❑