



JHT READ FOR CREDIT ARTICLE #594.

Scientific/Clinical Article

Polish version of the Patient-Rated Ulnar Nerve Evaluation in preoperative patients: Translation and psychometric testing



Mateusz Koziej MD^{a,b}, Marek Trybus MD, PhD^{b,*}, Anna Mydłowska MD^b, Katarzyna Piątek MD^a, Marta Banach MD, PhD^c, Mateusz Hołda MD^a

^aDepartment of Anatomy, Jagiellonian University Medical College, Kraków, Poland

^bSecond Department of General Surgery, Jagiellonian University Medical College, Kraków, Poland

^cDepartment of Neurology, Jagiellonian University Medical College, Kraków, Poland

ARTICLE INFO

Article history:

Received 20 December 2016

Received in revised form

19 August 2017

Accepted 24 August 2017

Available online 22 September 2017

Keywords:

Cubital tunnel syndrome

Michigan Hand Outcomes Questionnaire

Patient-Rated Ulnar Nerve Evaluation

PRUNE

Ulnar nerve entrapment

ABSTRACT

Study Design: Cross-sectional design.

Introduction: This study examined the translated English to Polish version of the Patient-Rated Ulnar Nerve Evaluation (PRUNE) for its internal consistency, test-retest reliability, and construct validity.

Methods: During the first assessment validity testing, a total of 39 consecutive patients with cubital tunnel syndrome completed the PRUNE, Michigan Hand Outcome Questionnaire, Disabilities of the Arm, Shoulder, and Hand questionnaire, and Patient Evaluation Measure in conjunction with the grip and key pinch tests and pain score (by Visual Analogue Scale). Cronbach's alpha (CA), intraclass correlation coefficient (ICC), and the Bland-Altman plot were used to evaluate internal consistency, test-retest reliability, and agreement, respectively. Analysis of variance compared the PRUNE score with the McGowan clinical stages.

Results: After a 1-day interval, 19 patients completed the PRUNE for the second time. The total PRUNE score was 44.4 ± 20.4 , CA = 0.93, and ICC = 0.921. The total PRUNE score limits of agreement varied from -9.87 to 7.55 points. PRUNE subscale CA ranged from 0.79 to 0.92; the ICC varied from 0.738 to 0.911. The construct validity revealed a strong association with Michigan Hand Outcome Questionnaire ($R = -0.83$; $P < .000$), and moderate with Disabilities of the Arm, Shoulder, and Hand ($R = 0.75$; $P < .000$), Patient Evaluation Measure ($R = 0.75$; $P < .000$), and Visual Analogue Scale ($R = 0.69$; $P < .000$). The grip and pinch tests had low and no correlation with the total PRUNE score, respectively.

Conclusion: The Polish version of PRUNE showed good psychometric properties for use in both clinical and research practice in patients with cubital tunnel syndrome of varying intensity.

© 2017 Hanley & Belfus, an imprint of Elsevier Inc. All rights reserved.

Introduction

Cubital tunnel syndrome (CuTS) involves compression and irritation of the ulnar nerve at the elbow, which causes significant paresthesia, pain, and disability in the area innervated by that nerve.¹ The epidemiology of this second most common neuropathy of the upper limb has not been investigated in Poland yet. In Siena, the mean annual crude incidence was 24.7 cases per 100,000 person-years, and the standardized incidence was 20.9. For men,

the incidence was 32.7, and for women, it was 17.2.² The prevalence of CuTS in a US metropolitan cohort with a more sensitive case definition (lax criteria) was 5.9%, whereas a more specific case definition (strict criteria) was 1.8%.³

An appropriate tool for monitoring disease severity is necessary.^{4,5} A recent Cochrane review examined the treatment approaches for ulnar nerve neuropathy.⁶ Most studies included in this review assessed the recovery by either assessing the neurophysiological and/or imaging characteristics in conjunction with clinical characteristics. In particular, clinical characteristics included generic hand or upper limb self-reported measures.⁷⁻¹² Patients with CuTS experience unique set of signs and symptoms which may not be adequately captured by the generic hand or upper limb measures. The condition such as CuTS requires that the condition-specific measure used for a comprehensive assessment of impairments that result from the CuTS.¹²

Conflict of interest: All named authors hereby declare that they have no conflicts of interest to disclose.

* Corresponding author. Second Department of General Surgery, Jagiellonian University Medical College, Kopernika 50, Kraków, Poland. Tel.: +48 12 351 66 61; fax: +48 12 351 67 50.

E-mail address: mt.ujcm@gmail.com (M. Trybus).

The Patient-Rated Ulnar Nerve Evaluation (PRUNE) is an open access, patient-reported assessment consisting of 20 items about ulnar nerve-related sensory/motor deficits and their associated functional disability.¹³ In the analysis, PRUNE demonstrated appropriate psychometric properties for patients with CuTS, with an intraclass coefficient for a total score equaling 0.98 and a standardized response mean of 1.55.

Despite the preliminary evidence on the psychometric properties of the English version of the PRUNE, the PRUNE is not available for use in Polish language. According to available literature, approximately 40 million people worldwide speak Polish as their first language; therefore, there is a value in developing a Polish language version of the PRUNE.¹⁴ The purpose of this study was to translate and adapt the English PRUNE into Polish language (PRUNE-P). The second objective was to examine whether the PRUNE-P has adequate reliability and validity for assessing impairments in patients with CuTS who are scheduled to undergo surgery. In particular, the study examined the test-retest reliability, internal consistency, and concurrent validity of the PRUNE-P.

Material and methods

The following study used a cross-sectional design, whereas its methodology consisted of two parts: (1) the first part translated and culturally adapted the English version of PRUNE into Polish and (2) the second part assessed reliability and validity of the PRUNE-P by testing it on patients who were scheduled to undergo surgical interventions for the condition. The translation agreement from PRUNE authors was obtained. The study was approved by the Bioethical Committee of the Jagiellonian University Medical College, Kraków (#122.6120.107.2015).

Translation of the polish PRUNE

The translation was in accordance with the American Academy of Orthopedic Surgeons Outcome Committee¹⁵ Forward translation from the original PRUNE was done by 2 independent researchers (T1 and T2), both of whom were native Polish speakers and fluent in English, along with medical degrees as well. T2 was not informed about the study concept. Subsequently, a consensus version was developed in which all minor cultural discrepancies were resolved (T1 and T2). Two professional (native) English translators who speak Polish as a second language (BT1 and BT2) did a backup

translation as the next step. In addition to BT1 also having medical training, eventually, all translations were compared by the expert committee, which consisted of T1, T2, BT1, BT2, a methodologist, and a developer. During this process, all reports were reviewed, as consensus about discrepancies was clarified (semantic, idiomatic, experiential, and conceptual equivalences). This led to a preliminary version of the PRUNE-P.

We worked on pretesting, in which the preliminary version was distributed to 5 patients with CuTS (scheduled for surgery), so that they could read the questionnaire for comprehension and interpretation of specific points. Those patients were asked to indicate any ambiguous questions. Collected data allowed us to compose the final version of the PRUNE-P and evaluate it on a larger group for test-retest reliability, internal consistency, and validity.

Patients

The study was conducted in the Second Department of General Surgery, Jagiellonian University Medical College from January 2015 to July 2016. A total of 39 patients were included in the study. Inclusion criteria consisted of several parameters: (1) CuTS symptoms; (2) nerve conduction study confirmation; (3) no surgery on hand/wrist related to ulnar nerve entrapment; (4) >18 years old; (5) ability to read and answer the complete questionnaire independently; (6) Polish was their first language; (7) no expected or significant changes in symptoms between the first and second measurements in 1 day; and (8) no history of neurologic and/or mental disease. Written informed consent was received from all participants.

Study participants were classified according to the McGowan Scale modified by Goldberg¹⁶: (1) I—subjective symptoms without objective findings; (2) IIA—motor weakness with good intrinsic strength, no atrophy; (3) IIB—presenting atrophy of intrinsic muscles with fair strength; and (4) III—intense sensory and motor disturbance with marked intrinsic atrophy.

PRUNE

Questions in the PRUNE are divided into 4 groups: (1) pain (6 questions); (2) sensory/motor symptoms (4 questions); (3) specific activity (6 questions); and (4) typical activity (4 questions).¹³ Each item is recorded on a 0–10 scale, where 0 is none/no difficulty and 10 is the worst possible/completely disabled.

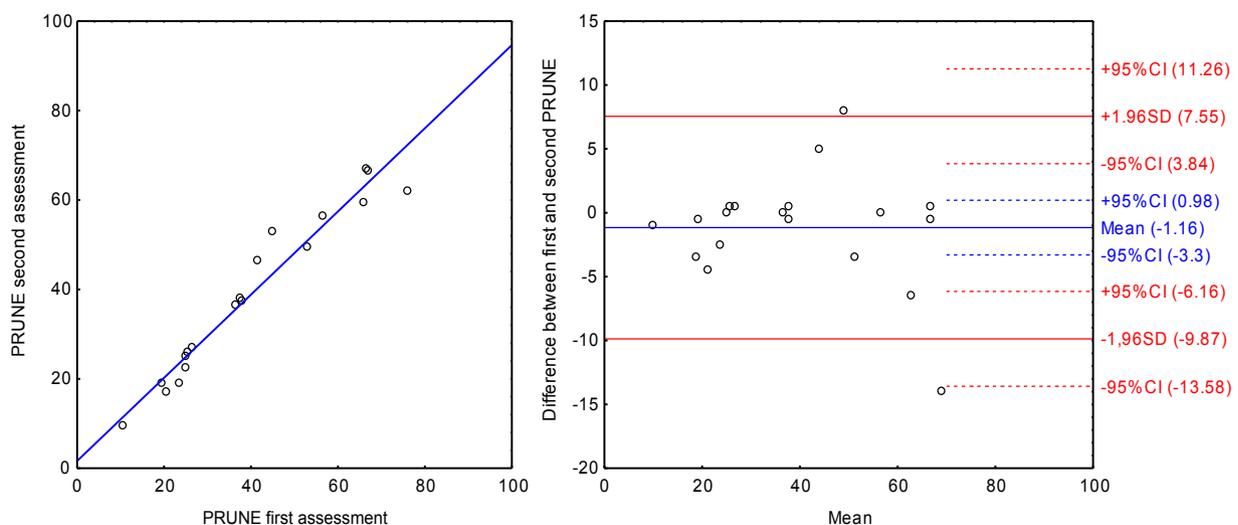


Figure 1. The Bland-Altman plot for results after a 1- to 2-day interval for PRUNE score. CI = confidence interval; PRUNE = Patient-Rated Ulnar Nerve Evaluation; SD = standard deviation.

The scoring system's subscale is counted by adding the component questions (pain: 60, sensory/motor symptoms: 40, specific activity: 60, and usual activity: 40). The overall score is calculated from 0 (no symptoms or difficulty) to 100 (worst symptoms and completely incapable of doing all functional activities). The total score equally weighs the 10 symptom items and the 10 functional items by dividing the total by 2.

Additional assessment tools

The PRUNE-P was compared with other objective/subjective hand measure instruments. The most commonly used tools in Poland (and described in the available literature) were followed with the chosen tools.^{4,17}

Disabilities of the Arm, Shoulder, and Hand (DASH) was chosen as a reference questionnaire, for which scores could be compared to PRUNE-P.¹⁸ DASH is a 30-item, patient-reported questionnaire that measures physical function and symptoms with any (or several) musculoskeletal disorders of the upper limbs. The score for each item is added and then converted to the scale of 0–100, where the higher number indicated greater disability.

The second questionnaire was the Polish Michigan Hand Outcome Questionnaire (MHQ), which is a hand-specific outcome assessment that measures patient results with conditions or disabilities of the hand/wrist or related injury.^{19,20} An official license from the Regents of the University of Michigan was obtained to use this tool in the study. It consisted of 37 core questions, some of which addressed the left and right hand individually. MHQ contains 6 distinct subscales: (1) overall hand function; (2) activities of daily living; (3) work performance; (4) pain; (5) esthetics; and (6) patient satisfaction with hand function.

Each question was answered from 1 to 5, with each of the 6 scales being scored and converted from 0 to 100 according to the specific scoring system. A higher outcome indicated better results, except for the pain scale, where a higher outcome was associated with more intense pain. The final score was calculated individually for both the left and right hands.

The Patient Evaluation Measure (PEM) was composed of 3 parts. In this analysis, the second part starts with, "How is your hand now," consisting of 10 questions, each of which addressed different aspects of hand symptoms and function.²¹ Each question was scored with a 7-point scale (1–7), where 1 indicated a normal state and 7 indicated severe pathology. The rating ranged from 10 to 70 points, but the final result was expressed as a percent. The intensity of pain during activity was evaluated with the 10-cm Visual Analogue Scale (VAS). The range of assessment was 0–10. A higher score indicated more pain.

The grip and key pinch strengths of both hands were measured with the Jamar hand hydraulic dynamometers (Patterson Medical Kit, Warrenville, IL). Three measurements were taken, and the final outcome was presented as the average score \pm 0.1 kg. Grip strength was assessed with the elbow flexed at 90° and the forearm in neutral rotation. For key pinch testing, patients were seated with the elbow at 90°, the forearm in a neutral position, and the wrist in neutral deviation, in which the pad of the thumb was placed against the middle radial aspect of the index finger.²²

Procedure

Patients admitted to the hand surgery and orthopedic outpatient clinic with confirmed CuTS were informed about the project. PRUNE-P, MHQ, DASH, and PEM were given next, and the VAS scale, along with the grip and pinch strengths were measured. All participants were asked to complete the PRUNE-P for the second

time after at least 1 day (24–48 hours). Demographic data, including sex, age, duration of symptoms, and dominant hand, were collected.

Questionnaires were considered valid if they reached adequate completion, depending on guidance and scoring instructions. For MHQ scales with less than 50% of the questions unanswered, the average of the existing scale items may be substituted for the missing scale(s).¹⁹ DASH was not considered valid if more than 3 items were missing.¹⁸ PEM was only valid when entirely completed. This study focused on PRUNE-P, so the assessment had to be fully completed in order for the results to be valid.

Statistical analysis

Descriptive statistics (mean with standard deviation [SD] and frequency) were calculated for all demographic and outcome variables.

Reliability, a measure of overall consistency, was assessed using internal consistency, reproducibility, and measurement error. To evaluate internal consistency, Cronbach's alpha (CA) was used, which is a function of the number of items in a given test, the average covariance between item pairs, and the variance of the total score.²³ A higher CA (ranging from 0.7 to 0.95) is preferable.²⁴

Reproducibility describes the grade of similar answers for questions used in repeated measurements for a stable person (test-retest). It was assessed by 2 analyses. The first was the intraclass correlation coefficient (ICC_{2,1}: a 2-way random effects absolute agreement model). The outcome was within the range of 0–1. Values > 0.7 indicated a good test-retest reliability.²⁴ The second analysis was the Bland-Altman method, which consists of plotting mean scores for both the occasions vs differences in the scores for both the occasions.²⁵ The plot should illustrate mean difference in the scores for both occasions plotted along with limits of agreements (LoA) which are defined as ± 2 SD for the mean difference. The plot for each hand's PRUNE-P score was paired together with a scatter plot between the first and second assessments. According to Bland and Altman, the Pearson coefficient could be misleading as an analysis for establishing agreement between measures, given that data can present high and significant positive correlation but still be in poor agreement.²⁵

The standard error of measurement ($SEM = SD * \sqrt{1 - ICC}$) is the level of error that can be considered an actual measurement error. The minimal detectable change ($MDC_{95} = 1.96 * SEM * \sqrt{2}$) is the lowest change in a patient's PRUNE-P score, ensuring the change was not a result of measurement error.

Construct validity describes convergence between tested instruments and other outcomes that measure similar constructs and measures features for which it was created.^{24–26} For concurrent validity, the correlation coefficient was used to assess the association of PRUNE-P with Polish MHQ, DASH, PEM, and VAS during activity and during the grip and pinch tests.²⁷ In this analysis, only hands with CuTS of MHQ outcomes were taken into account. Several hypotheses were designed: (1) tools measuring the same construct should indicate a high correlation ($R > 0.75$) and (2) hand-specific questionnaires should demonstrate a moderate to high correlation ($R > 0.50$). With MHQ, higher grip and pinch test scores reflected better results, although an inverse correlation for this particular test should similarly exist. In the development of PRUNE, Short Form-36 (SF-36) was considered in this analysis, which is only a generic health questionnaire.¹³ Ceiling and floor effects and the number of respondents who achieved the lowest or highest possible scores were calculated using the 15% criterion.²⁸

To compare the PRUNE score and symptom duration between the McGowan grades, analysis of variance (ANOVA) was used with Tukey post hoc comparison, preceded by checking the normal

distribution (Shapiro-Wilk test) and homogeneity of data (the Levine's test).

A *P* value of < .05 was seen as statistically significant. Analyses were performed with STATISTICA v12 (StatSoft Inc, Tulsa, OK).

Results

Translation and adaptation

We observed no linguistic issues during translation. Only one activity was renamed due to cross-cultural adaptation. On the “eat” item, the term “chopsticks” is not popular in Poland, so it was changed to “spoon.” Pretesting a group of 5 patients demonstrated very minor difficulties in the preliminary draft, which did not occur in testing of the final version. Examples were added to a “repeated reaching” item to make the question more understandable (ie, hanging laundry and curtains). No floor or ceiling effects were identified, and no missing items were found in PRUNE.

Demographic

The study group consisted of 19 men (48.7%) and 21 women (51.3%), with a mean age of 54.0 ± 15.3 years (Table 1). All participants were Polish (Caucasian). The left hand was affected slightly more often than the right one (56.4% vs 43.6%, respectively). Most of the patients were right handed (92.3%) and lived in the city (74.4%). McGowan grades for the patients consisted of several scores: (1) I:4 (10.3%); (2) IIA: 9 (23.1%); (3) IIB: 16 (41.0%); and (4) III: 10 (25.6%). The average duration of symptoms was 27.1 ± 20.1 months, and there were no significant differences between clinical groups (ANOVA *F* = 1.99, *P* = .163).

Analysis of reliability and validity

Table 2 presents study outcomes for subjective and objective measures. Nineteen patients (48.7%) filled in the form for the second time. The PRUNE-P scores for the first assessment were not different between those who were retested vs those who were not. The average difference between the first (40.0 ± 19.3) and second (38.3 ± 17.8) assessments for 19 patients was -1.68 (95% confidence interval [CI]: -1.34 to 4.71). The internal consistency

Table 1
Patients' demographic data (total *n* = 39)

Characteristics	Value	SD, %
Gender		
Male	19	48.7
Female	20	51.3
Nationality		
Polish	100%	
Age (y)	54.0	15.3
Affected hand		
Right	17	43.6
Left	22	56.4
Dominant hand		
Right	36	92.3
Left	3	7.7
Place of residence		
Village	10	25.6
City	29	74.4
Symptom duration (mo)	27.1	20.1
McGowan scale (modified)		
I	4	10.3
IIA	9	23.1
IIB	16	41.0
III	10	25.6

Table 2
Average values with standard deviation of measured parameters

Variable	#1 Assessment	#2 Assessment
PRUNE (<i>n</i> = 39; <i>n</i> = 19)	44.4 ± 20.4	38.3 ± 17.8
MHQ injury hand (<i>n</i> = 39)	52.4 ± 17.0	
PEM (<i>n</i> = 39)	45.1 ± 18.8	
DASH (<i>n</i> = 39)	38.6 ± 18.8	
VAS (<i>n</i> = 39)	4.8 ± 3.2	
Grip test—injury hand (<i>n</i> = 32)	28.5 ± 12.3	
Pinch test—injury hand (<i>n</i> = 32)	5.8 ± 2.8	

DASH = Disabilities of the Arm, Shoulder, and Hand; MHQ = Michigan Hand Outcome Questionnaire; PRUNE = Patient-Rated Ulnar Nerve Evaluation; PEM = Patient Evaluation Measure; VAS = pain Visual Analogue Scale during activity.

for the final score and the subscales were in the desirable range (CA = 0.79-0.93). The ICC of the final PRUNE-P score was 0.921, SEM = 15.9, and MDC95 = 5.7 (Table 3). The LoA for the PRUNE-P total score varied from -9.87 (95% CI: -13.58 to -6.16) to 7.55 (95% CI: 3.84-11.26). The Bland-Altman plot, with the scatter plot graph, is presented in Figure 1.

Table 4 shows the correlation between the PRUNE-P and other analyzed instruments. Significant correlations were found with MHQ (*R* = -0.83, *P* = .000), DASH (*R* = 0.75, *P* = .000), and PEM (*R* = 0.75, *P* = .000). The VAS moderately correlated with PRUNE (*R* = 0.69, *P* = .000). The grip test showed a low correlation with PRUNE-P (*R* = -0.36, *P* = .043), and the pinch test did not show any correlation with PRUNE-P (*P* > .05).

PRUNE-P scores were compared between each McGowan grade for the total score and 4 subscales (Table 5). Significant differences were observed between grades I and III in the total scores and the sensory motor symptom scale (*P* < .05). The Specific Activities Scale score was statistically different between I and IIB and I and III.

Discussion

Hand surgeons and hand therapists are increasingly interested in the problems that afflict patients; the choice of a proper questionnaire is crucial for monitoring a patient's state.⁴ This study provides the translation and adaptation of PRUNE to Polish. It only assessed patients prior to surgical intervention for CuTS, as sensitivity of the measures to minimally clinically important differences was not analyzed. The standardized response mean for improvement in the English PRUNE's total score from baseline to 2 years was also presented, equaling 1.55 points (mean change 34.3 ± 19.5 points). The significant measure changes (standardized response mean > 0.9) observed over time support its responsiveness. The study group in the Polish study was ethnically homogenous, which allowed for a reliable analysis.

Table 3
Internal consistency and test-retest reliability of the Polish version of the PRUNE

PRUNE	#1 Assessment		#2 Assessment		Cronbach's ICC	SEM	MDC95
	<i>(n</i> = 39)		<i>(n</i> = 19)				
	Mean score	SD	Mean score	SD	alpha		
Pain	37.7	26.7	33.5	23.3	0.92	0.895	8.7 24.0
Sensory/motor symptoms	56.7	22.8	52.0	22.7	0.79	0.738	11.7 32.3
Specific activity	46.8	23.4	39.1	22.1	0.87	0.840	9.4 25.9
Usual activity	38.4	26.3	30.5	20.2	0.92	0.911	7.8 21.7
Total	44.4	20.4	38.3	17.8	0.93	0.921	5.7 15.9

ICC = intraclass correlation coefficient; MDC95 = minimal detectable change 95%; PRUNE = Patient-Rated Ulnar Nerve Evaluation; SD = standard deviation; SEM = standard error of measurement.

Table 4

Correlation of PRUNE with MHQ, DASH, PEM, VAS, grip strength, and pinch strength for hand with diagnosed cubital tunnel syndrome

PRUNE	MHQ		DASH		PEM		VAS		Grip test		Pinch test	
	R ^a	P	R	P	R	P	R	P	R ^a	P	R ^a	P
Pain	−0.56	.000	0.49	.002	0.49	.003	0.85	.000	−0.23	.197	−0.06	.728
Sensory/motor symptoms	−0.65	.000	0.55	.000	0.70	.000	0.37	.02	−0.31	.086	−0.29	.104
Specific activity	−0.74	.000	0.69	.000	0.70	.000	0.44	.006	−0.32	.073	−0.22	.23
Usual activity	−0.74	.000	0.71	.000	0.67	.000	0.41	.009	−0.31	.08	−0.13	.489
Total	−0.83	.000	0.75	.000	0.75	.000	0.69	.000	−0.36	.043	−0.20	.275

DASH = Disabilities of the Arm, Shoulder, and Hand; MHQ = Michigan Hand Outcome Questionnaire; PEM = Patient Evaluation Measure; PRUNE = Patient-Rated Ulnar Nerve Evaluation; VAS = pain Visual Analogue Scale during activity.

^a Inverse correlation: in MHQ, grip, and pinch tests, a better score indicates a better result.

The translation steps did not exhibit any major cultural or linguistic discrepancies. Only one item from the original questionnaire required some modification; the use of chopsticks in Poland is very unpopular, so that the corresponding question was changed to the use of a spoon, given that eating soup is much more common in Poland. Despite this change, the main strength of the translation maintained its experimental equivalence in the 2 versions, such that the results of the Polish PRUNE are comparable with observations using the English PRUNE.

The original study for the English version of PRUNE was based on 89 patients with ICC subscales scores greater than 0.87 and a total score equal to 0.98.¹³ Outcomes based on ICC in the Polish study demonstrated high reliability of the measures and yielded scores similar to those obtained in the English study, even though the ICC method used in the English study was not clearly stated. The total score of English PRUNE in the first assessment was 34.4 ± 21.5 . In the present study, the final score in the Polish version was 44.4 ± 20.4 , which indicates that patients in the Polish cohort had more severe symptoms than those in the English cohort. It is also worth mentioning that the mean duration of symptoms for Polish patients lasted longer, with a mean value of approximately 2.3 years. Similarly, we found a difference between the sensory/motor symptom subscales; they were 12.6 ± 7.9 and 56.7 ± 22.8 in the English and Polish PRUNE, respectively. In addition, the higher values of PRUNE-P scores were associated with greater SDs, which explain the notably higher SEM and MDC95 scores for every subclass, as well as the total score compared to those in the English version (7.2 and 3.1, respectively). The higher scores for PRUNE-P could be due to: (1) that patients had more severe symptoms than patients in the English study and (2) the smaller sample size of the Polish study.

The strength of our study is the Bland-Altman analysis, which provides LoA values for the total score (−9.87 to 7.55) and has not been conducted so far for the English PRUNE. Changes within the LoA are likely due to random error, whereas changes outside the LoA are due to systematic error. The measurements outside of LoA range are not considerably different between the English and Polish

Table 5

Values of PRUNE depending on the McGowan grade clinical classification

PRUNE	McGowan grade (modified)							
	I		IIA		IIB		III	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Pain	22.1	23.3	35.7	23.4	35.9	29.6	48.7	25.2
Sensory motor symptoms	28.8	16.6	50.6	22.1	60.0	22.2	68.0	17.2
Specific activities	18.8	12.7	37.6	17.5	48.0	24.6	64.5	14.1
Usual activities	20.6	12.6	28.6	22.7	38.1	27.1	54.8	25.6
Total score	22.1	12.6	37.8	17.0	44.8	21.2	58.5	14.3

PRUNE = Patient-Rated Ulnar Nerve Evaluation; SD = standard deviation.

I = subjective symptoms and without objective findings; IIA = motor weakness with good intrinsic strength, no atrophy; IIB = presenting atrophy of intrinsic muscles with fair strength; III = intense sensory and motor disturbance with marked intrinsic atrophy.

versions and are seen on both sides of the distributions (upper and lower limits). This suggests acceptable agreement between the first and second assessments. Moreover, in the English and Polish studies, the McGowan clinical stage was used to classify patients' symptoms; however, the 2 studies cannot be compared because the English study applied the standard McGowan score, whereas the present study applied the modified version of the McGowan score. In the latter, stage II was divided into 2 substages A and B, depending on both atrophy and strength of hand muscles (Table 1).^{16,29} We conducted an ANOVA analysis, which demonstrated that the PRUNE-P total score could clearly distinguish only between extreme grades of the McGowan score modified by Goldberg. We concluded, therefore, that it is not possible to determine the explicit point ranges between grades based on the PRUNE questionnaire. We found that specific activity subgroup outcomes reflect the McGowan score most closely. Interestingly, the pain and usual activities of PRUNE components did not show statistically significant differences. This suggests that the McGowan clinical stage, as modified by Goldberg, cannot distinguish between these components, such that the amplitude of intensity is wider for each clinical stage of pain and a person's usual activities. Applying the PRUNE questionnaire along with the clinical scale provides much more useful information, which is essential in planning rehabilitation after surgery and which could improve outcomes of general treatment. It should also be mentioned that the Dellon scale was used in research conducted by Song et al in the preoperative classification of CuTS: (1) mild; (2) moderate; or (3) severe neuropathy. Similar to PRUNE-P, no significant differences in pain subscales and activities of daily living between each grade were found as assessed by MHQ.^{30,31}

This is the first study which compared a disease-specific PRUNE with general hand/wrist questionnaires. In the original study of the English PRUNE, correlation coefficients were significantly related to physical health domains on the SF-36 in contrast to mental health domains. The SF-36 is a generic instrument for measuring overall health but may not capture the interest of patients with upper extremity injuries and conditions.^{32,33}

In the construct validity of our translation, it was observed that MHQ, DASH, and PEM were significantly correlated with PRUNE-P. The most appropriate questionnaire for CuTS would require a specifically designed study to assess the changes in total PRUNE scores (and other forms) over time. Currently, the authors are conducting a study to answer this question. Recently, Song et al reported that the MHQ and Carpal Tunnel questionnaires were more focused than the DASH for isolated CuTS, treated with decompression.³⁰

In this study, the relationship to PRUNE's total score and medical instrument measurements are comparable to findings by Szekeres et al.³⁴ They observed a significant correlation of PRUNE with the grip test ($R = -0.38, P < .001$), an insignificant association of PRUNE with the pinch test ($R = -0.22, P = .06$), and a 2-point discrimination test ($R = 0.09, P = .45$) at baseline. They reported that changes in grip and pinch strengths, as well as sensation, accounted for 20% of the

variance in PRUNE changes at a 2-year follow-up. The insignificant pinch test association with PRUNE scores did not indicate irrelevance; however, since pinching involves fewer muscles than gripping, it is more difficult to detect in a single PRUNE measurement.³⁴

The main limitation of the preliminary study was its small sample for retest analysis. Some patients could not have had a long enough interval between the first and second assessments in a stable state, due to the operation being performed after the first assessment. The “gold standard” interval for obtaining measurements in the test-retest analysis has not been determined. In this study, a 24- to 48-hour interval was used, as Marx et al did not find significant differences for the 2 intervals.³⁵ The similarly constructed Patient-Rated Wrist Hand Evaluation (PRWHE-A) was tested with an even shorter interval time (30 minutes to 24 hours).³⁶ Minimizing possible recall bias was achieved by asking patients to fill out the PRUNE questionnaire first, followed by others. The answers required a response to reduce the possibility of remembering specific questions. Another known limitation is the possibility of a selection bias due to convenience sampling; however, other validation studies use similar design methods. In future studies, subjects should be recruited not only from orthopedic/hand clinics but from other settings (family practices, physical therapy offices, and so forth). This would help analyze patients with less symptomatic disease, increase sample size, and also increase estimate precision. The specific limitation of this study was using a statistical classical approach, vs using an item response theory approach, which was not performed in this preliminary work due to the small sample size together with possible number of rating scale responses (ie 20 items × 10 rating scale steps). The authors of the English PRUNE indicated that the next step would be the Rasch analysis, executed for differential item functioning issues.¹³

Conclusion

This preliminary study translates and adapts the Polish version of the PRUNE questionnaire. Agreement analysis limits have shown that the discrepancy between the average of the 2 measures, along with their differences, was not high, suggesting acceptable agreement between them. Construct validity revealed a strong association with MHQ and a moderate one with DASH, PEM, and VAS. The grip test indicated a low correlation with PRUNE-P's total score, whereas the pinch test did not show significant correlation. Future studies should confirm factor structure and dimensionality. In addition, identifying minimally clinically important differences of the PRUNE-P would increase the authors' confidence about its measurement properties across different contexts.

Supplementary data

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.jht.2017.08.004>

References

- Thorne C, Chung KC, Gosain A, Guntner GC, Mehrara BJ. *Grabb and Smith's plastic surgery*. 7th ed. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2014.
- Mondelli M, Giannini F, Ballerini M, Ginanneschi F, Martorelli E. Incidence of ulnar neuropathy at the elbow in the province of Siena (Italy). *J Neurol Sci*. 2005;234:5–10.
- An TW, Evanoff BA, Boyer MI, Osei DA. The Prevalence of Cubital Tunnel Syndrome: A Cross-Sectional Study in a U.S. Metropolitan Cohort. *J Bone Joint Surg Am*. 2017;99:408–416.
- Zyluk A, Piotuch B. [Use of questionnaires in outcome measurement in hand surgery]. *Chir Narzadow Ruchu Ortop Pol*. 2009;74:193–201.
- Assmus H, Antoniadis G, Bischoff C, et al. Cubital tunnel syndrome - a review and management guidelines. *Cent Eur Neurosurg*. 2011;72:90–98.
- Caliandro P, La Torre G, Padua R, Giannini F, Padua L. Treatment for ulnar neuropathy at the elbow. *Cochrane Database Syst Rev*. 2016;11:CD006839.
- Kang HJ, Koh IH, Chun YM, Oh WT, Chung KH, Choi YR. Ulnar nerve stability-based surgery for cubital tunnel syndrome via a small incision: a comparison with classic anterior nerve transposition. *J Orthop Surg Res*. 2015;10:121.
- Ebersole GC, Davidge K, Damiano M, Mackinnon SE. Validity and responsiveness of the DASH questionnaire as an outcome measure following ulnar nerve transposition for cubital tunnel syndrome. *Plast Reconstr Surg*. 2013;132:81e–90e.
- Giladi AM, Gaston RG, Haase SC, et al. Trend of recovery after simple decompression for treatment of ulnar neuropathy at the elbow. *Plast Reconstr Surg*. 2013;131:563e–573e.
- Malay S, Chung KC, Group SS. The minimal clinically important difference after simple decompression for ulnar neuropathy at the elbow. *J Hand Surg Am*. 2013;38:652–659.
- Kim KW, Lee HJ, Rhee SH, Baek GH. Minimal epicondylectomy improves neurologic deficits in moderate to severe cubital tunnel syndrome. *Clin Orthop Relat Res*. 2012;470:1405–1413.
- Macadam SA, Bezuhyly M, Lefavre KA. Outcomes measures used to assess results after surgery for cubital tunnel syndrome: a systematic review of the literature. *J Hand Surg Am*. 2009;34:1482–1491.e1485.
- MacDermid JC, Grewal R. Development and validation of the patient-rated ulnar nerve evaluation. *BMC Musculoskelet Disord*. 2013;14:146.
- Lewis MP, Simons GF, Fennig CD, eds. *Ethnologue: Languages of the World*. 18th ed. Dallas, Tex: SIL International; 2015.
- Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine (Phila Pa 1976)*. 2000;25:3186–3191.
- Goldberg BJ, Light TR, Blair SJ. Ulnar neuropathy at the elbow: results of medial epicondylectomy. *J Hand Surg Am*. 1989;14:182–188.
- Sambandam SN, Priyanka P, Gul A, Ilango B. Critical analysis of outcome measures used in the assessment of carpal tunnel syndrome. *Int Orthop*. 2008;32:497–504.
- The Disabilities of the Arm, Shoulder and Hand (DASH) Outcome Measure. Available at: <http://dash.iwh.on.ca/>. Accessed September 30, 2016.
- Chung KC, Pillsbury MS, Walters MR, Hayward RA. Reliability and validity testing of the Michigan Hand Outcomes Questionnaire. *J Hand Surg Am*. 1998;23:575–587.
- Koziej M, Trybus M, Mydlowska, et al. The Polish version of the Michigan Hand Outcomes Questionnaire: Cross-cultural adaptation, reliability, construct validity, and measurement error. *J Hand Surg Eur Vol* 2017. <http://dx.doi.org/10.1177/1753193417729579> [Epub ahead of print].
- Macey AC, Burke FD, Abbott K, et al. Outcomes of hand surgery. British Society for Surgery of the Hand. *J Hand Surg Br*. 1995;20:841–855.
- MacDermid JC, Evenhuis W, Louzon M. Inter-instrument reliability of pinch strength scores. *J Hand Ther*. 2001;14:36–42.
- Cronbach L. Coefficient alpha and the internal structure of tests. *Psychometrika*. 1951;16:297–334.
- Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60:34–42.
- Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1:307–310.
- Kirshner B, Guyatt G. A methodological framework for assessing health indices. *J Chronic Dis*. 1985;38:27–36.
- Hinkle DE, Wiersma W, Jurs SG. In: *Applied statistics for the behavioral sciences*. 5th ed. Boston, Mass: Houghton Mifflin; [London: Hi Marketing] [distributor]; 2003.
- McHorney CA, Ware JE. Construction and validation of an alternate form general mental health scale for the Medical Outcomes Study Short-Form 36-Item Health Survey. *Med Care*. 1995;33:15–28.
- McGowan AJ. The results of transposition of the ulnar nerve for traumatic ulnar neuritis. *J Bone Joint Surg Br*. 1950;32-B:293–301.
- Song JW, Waljee JF, Burns PB, et al. An outcome study for ulnar neuropathy at the elbow: a multicenter study by the surgery for ulnar nerve (SUN) study group. *Neurosurgery*. 2013;72:971–981. discussion 981–972; quiz 982.
- Dellon AL. Review of treatment results for ulnar nerve entrapment at the elbow. *J Hand Surg Am*. 1989;14:688–700.
- 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;30:81–86.
- Alderman AK, Chung KC. Measuring outcomes in hand surgery. *Clin Plast Surg*. 2008;35:239–250.
- Szekeres M, MacDermid JC, King GJ, Grewal R. The relationship between the Patient-rated Ulnar Nerve Evaluation and the common impairment measures of grip strength, pinch strength, and sensation. *J Hand Ther*. 2015;28:39–44. quiz 45.
- Marx RG, Menezes A, Horovitz L, Jones EC, Warren RF. A comparison of two time intervals for test-retest reliability of health status instruments. *J Clin Epidemiol*. 2003;56:730–735.
- Hasani FN, MacDermid JC, Tang A, Kho ME. Cross-cultural adaptation and psychometric testing of the Arabic version of the Patient-Rated Wrist Hand Evaluation (PRWHE-A) in Saudi Arabia. *J Hand Ther*. 2015;28:412–419. quiz 420.

JHT Read for Credit

Quiz: # 594

Record your answers on the Return Answer Form found on the tear-out coupon at the back of this issue or to complete online and use a credit card, go to JHTReadforCredit.com. There is only one best answer for each question.

- # 1. The original PRUNE was developed by
 - a. Dellon et al
 - b. Moberg et al
 - c. MacDermid et al
 - d. Fess, Bell et al
- # 2. The following statistic was NOT used
 - a. Pearson's Coefficient
 - b. Bland-Altman Plot
 - c. Cronbach's Alpha
 - d. ICC
- # 3. The PRUNE was evaluated for its
 - a. reliability
 - b. construct validity

- c. internal consistency
- d. all of the above
- # 4. The forward translation was performed by
 - a. two UN representatives from Krakow stationed in NYC
 - b. the chairman of the foreign language department of Yale University
 - c. two medically trained native Polish speakers who were also fluent in English
 - d. two medically trained native English speakers who were also fluent in Polish
- # 5. The Polish translation of the PRUNE was shown to meet traditional standards for clinical and research use
 - a. false
 - b. true

When submitting to the HTCC for re-certification, please batch your JHT RFC certificates in groups of 3 or more to get full credit.