PHYSIOTHERAPY THEORY AND PRACTICE
https://doi.org/10.1080/09593985.2025.2450090

Taylor & Francis
Taylor & Francis Group

REPORT

OPEN ACCESS Check for updates

Clinimetric analysis of the numeric pain rating scale, patient-rated tennis elbow evaluation, and tennis elbow function scale in patients with lateral elbow tendinopathy

Ian Young PT DSc, OCS, RMSK[a], James Dunning DPT, PhD, FAAOMPT[b,c], Firas Mourad PT, PhD[d,e], James Escaloni DPT, FAAOMPT, RMSK[b,f], Paul Bliton DPT, OCS, FAAOMPT[b,g], and César Fernández-de-Las-Peñas PT, PhD[h]

[a]Department of Physical Therapy, Tybee Wellness & Osteopractic, Tybee Island, USA; [b]Department of Physical Therapy, American Academy of Manipulative Therapy Fellowship in Orthopaedic Manual Physical Therapy, Montgomery, AL, USA; [c]Department of Physical Therapy & Acupuncture, Montgomery Osteopractic Physical Therapy & Acupuncture, Montgomery, AL, USA; [d]Department of Physiotherapy, LUNEX International University of Health, Exercise and Sports, Differdange, Luxembourg; [e]Department of Physiotherapy, Luxembourg Health & Sport Sciences Research Institute, Differdange, Luxembourg; [f]Department of Physical Therapy, Wellward Regenerative Medicine, Lexington, KY, USA; [g]Department of Physical Therapy, William Middleton VA Hospital, Madison, WI, USA; [h]Department of Physical Therapy, Occupational Therapy, Rehabilitation and Physical Medicine, Universidad Rey Juan Carlos, Madrid, Spain

ABSTRACT

Background: Currently, there is conflicting clinimetric data on the patient-rated tennis elbow evaluation (PRTEE) and a paucity of evidence regarding the reliability, validity, and responsiveness of the numeric pain rating scale (NPRS), and tennis elbow function scale (TEFS) in patients with lateral elbow tendinopathy.

Objective: Perform a comprehensive clinimetric analysis of the NPRS, PRTEE, and TEFS in a sample of patients ($n = 143$) with lateral elbow tendinopathy.

Methods: Establish the reliability, construct validity, responsiveness, meaningful clinically important difference (MCID), and minimal detectable change (MDC_{90}) values for the NPRS, PRTEE, and TEFS at the 3-month follow-up.

Results: The NPRS [intraclass correlation coefficient ($ICC_{2,1}$): 0.54, 95% confidence interval (CI): 0.17–0.78], PRTEE ($ICC_{2,1}$: 0.62, 95% CI: 0.21–0.86), and the TEFS ($ICC_{2,1}$: 0.71, 95% CI: 0.14–0.90) exhibited moderate reliability. All three outcomes exhibited excellent responsiveness [NPRS: area under the curve (AUC): 0.94, 95% CI: 0.89–0.98]; PRTEE: (AUC: 0.96, 95% CI 0.93–0.99); TEFS: (AUC: 0.95, 95% CI: 0.91–0.98). The MCID and MDC_{90} were 2.3 and 1.4 for the NPRS, 14.8 and 9.7 for the PRTEE, and 7.5 and 5.7 for the TEFS, respectively. All three patients reported outcome measures also demonstrated strong construct validity (Pearson's r from 0.71 to 0.83, $p < .001$).

Conclusion: The NPRS, PRTEE, and TEFS are clinimetrically sound patient reported outcome measures for patients with lateral elbow tendinopathy at a 3-month follow-up.

Registration at ClinicalTrials.gov: NCT03167710.

ARTICLE HISTORY
Received 4 November 2024
Revised 31 December 2024
Accepted 31 December 2024

KEYWORDS
Elbow tendinosis; psychometric; patient-reported outcome measures

Introduction

Lateral elbow tendinopathy (LET) is a common presentation of pain and disability in the elbow (Di Filippo, Vincenzi, Pennella, and Maselli, 2022; Lucado et al., 2022), with an overall incidence of 1–3%, and as high as 10% in the United States (Sanders et al., 2015). Lateral elbow tendinopathy has been identified as an internal pathology of the tendon at or close to its enthesis (Kraushaar and Nirschl, 1999; Lucado et al., 2022; Scott, Backman, and Speed, 2015). Repeated mechanical loading in this area results in chronic histopathological changes within the tendon (Jomaa et al., 2020; Kraushaar and Nirschl, 1999; Lucado et al., 2022; Schneeberger and Masquelet, 2002; Scott, Backman, and Speed, 2015). Multimodal physical therapy treatment programs including local mobilization/manipulation, exercise, dry needling, and spinal manipulation (Bisset et al., 2006; Coombes et al., 2013; Dunning et al., 2024) have shown favorable effects in reducing pain and disability in this patient population. However, the level of improvement noted with short-term self-report outcomes of less than 6-weeks may not capture the true magnitude of structural and mechanical

CONTACT Ian Young tybeewellness@gmail.com Department of Physical Therapy, Tybee Wellness & Osteopractic, Tybee Island, USA

The original randomized controlled trial was approved by the ethics committee at Universidad Rey Juan Carlos, Madrid, Spain (URJC-DPTO 11–2017).

© 2025 The Author(s). Published with license by Taylor & Francis Group, LLC.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

changes that signify chronic tendon healing (Khan, Cook, Taunton, and Bonar, 2000; Voleti, Buckley, and Soslowsky, 2012). Clinimetrically established outcome measures with longer follow-up, in conjunction with diagnostic imaging via musculoskeletal ultrasound, may provide the best option to assess the combination of clinical and histological improvement (Dones et al., 2014; Evans et al., 2019; Macdermid and Silbernagel, 2015).

The numeric pain rating scale (NPRS) (Farrar et al., 2001; Jensen, Karoly, and Braver, 1986), patient-rated tennis elbow evaluation (PRTEE) (Evans et al., 2019; Macdermid and Silbernagel, 2015), and the tennis elbow function scale (TEFS) (Lowe, 1999) are patient reported outcome measures (PROMs) commonly used in patients with LET. In concert, these PROMs should give clinicians and researchers a more comprehensive understanding of the patient response (pain, function, and disability) to an intervention (Mercieca-Bebber et al., 2018). Unfortunately, there is a paucity of evidence and/or inconsistencies in some of the established clinimetric properties of these instruments. Importantly, no data exist specifically for the NPRS as a self-report outcome in patients with LET, and only one prior study provided limited clinimetric analyses of the TEFS (Lowe, 1999). Further, only three studies with low to moderate evidence support small sample-sized anchor-based assessment of the meaningful clinically important difference (MCID) of the PRTEE at short-term follow-up (Cacchio et al., 2012; Poltawski and Watson, 2011; Shafiee et al., 2022). Thorough clinimetric analyses of these outcomes, with updated clinical interventions, and a longer term follow-up (3-months) may bolster the limited evidence of these commonly used PROMs and reduce the likelihood of over/under estimation of clinical improvement after an intervention targeting tendon pathology.

Therefore, the purpose of this study was to examine the reliability, construct validity and responsiveness of the NPRS, PRTEE, and TEFS at the 3-month follow-up, in a cohort of patients successfully treated for LET.

Methods

A secondary clinimetric analysis of a prior multicenter randomized clinical trial was performed on consecutive individuals ($n = 143$) with LET from 13 outpatient physical therapy clinics in 9 different US states (Dunning et al., 2024). Patients were screened for eligibility criteria and recruited over a 45-month period (June 2017 to March 2021). To be eligible, patients had to be between 18 and 60 years old and meet the following criteria: (1) a clinical diagnosis of lateral elbow tendinopathy—i.e., defined as two of more of the following: (a) pain on palpation over the lateral epicondyle and the associated common extensor unit, (b) pain on gripping a hand dynamometer 22, (c) pain with stretching or contraction of the wrist extensor muscles, (2) lateral elbow and forearm symptoms for longer than 6 weeks, and (3) an intensity of lateral elbow pain of at least 2 on the numeric pain rating scale (NPRS, 0–10). The inclusion and exclusion criteria and treatments in each group have previously been described in detail (Dunning et al., 2024). Patients were appropriately randomized and treated across all clinical sites with a standardized multi-modal physical therapy program (soft tissue mobilization, joint mobilization, exercise, and therapeutic ultrasound to the elbow), with and without the addition of percutaneous tendon dry needling and extremity/spinal thrust-manipulation (Dunning et al., 2024).

Outcome measures

The NPRS is an 11-point (0, no pain; 10, worst imaginable pain) used to assess the intensity of pain (Jensen, Karoly, and Braver, 1986). The NPRS is a reliable and valid instrument to assess pain intensity (Farrar et al., 2001; Salaffi et al., 2004; Young et al., 2018, 2019). The MCID for the NPRS has been shown to be 1.74 in patients with a variety of chronic pain conditions; thus, a change of 2 points or a 30% decrease in pain from baseline can be considered as MCID in individuals with chronic musculoskeletal pain (Farrar et al., 2001; Salaffi et al., 2004). The NPRS is reported to have acceptable reliability using intraclass correlation coefficient (ICC) (ICC = 0.74), strong construct validity with the patient-specific functional scale ($p < .001$), and excellent responsiveness (area under the curve: 0.90) in patients with upper extremity conditions (Hefford, Abbott, Arnold, and Baxter, 2012). Similarly, the minimal detectable change (MDC) of 5.7 points and MCID of 2.5 points on the NPRS have only been reported in patients with upper extremity problems, but not specifically in patients with LET (Hefford, Abbott, Arnold, and Baxter, 2012). Although the visual analog scale (VAS) has been supported in patients with LET (Rompe, Overend, and MacDermid, 2007), no prior clinimetric analyses have been done on the NPRS as a stand-alone or combined self-report outcome in patients with LET.

The PRTEE is the PROM most commonly used in patients with LET for assessing function and has been found to be reliable, valid, and responsive in capturing function in individuals with LET (Evans et al., 2019; Macdermid and Silbernagel, 2015; Shafiee et al., 2022). The questionnaire consists of two parts, including both pain and function. The first part consists of five questions

scored from 0 (no pain) to 10 (most severe pain). The scores for the five pain questions are summed, and a total score out of 50 is reported. The function part of the questionnaire comprises 10 questions, the scores of which are summed and divided by 2, for a total score out of 50. Scores on the pain and function subscales are summed for a total score out of 100 (Vincent and MacDermid, 2014). Lower scores indicate a higher function. Acceptable reliability (ICC) has been reported in patients with LET, ranging from 0.78 to 0.99 (Shafiee et al., 2022). A wide range of values for the MDC (0.58 points to 22.9 points) and MCID (7 points and 21 points) have been reported for the PRTEE (Shafiee et al., 2022).

The TEFS is a 10-item, 5-point response self-report scale designed to measure elbow discomfort during the performance of personal care, household, work, and recreational activities (Lowe, 1999). To date, there are limited psychometric/clinimetric data on the TEFS for evaluating patients with LET. Acceptable reliability (ICC 0.92), construct validity, and responsiveness of the TEFS have been reported in a single preliminary study with a smaller sample size (Lowe, 1999), while the MCID and MDC have not yet been established in patients with LET.

Patients also completed a 15-point Global Rating of Change (GROC) scale described by Jaeschke, Singer, and Guyatt, (1989) to rate their own perception of improved function. The scale ranges from −7 (a very great deal worse) to 0 (about the same) to +7 (a very great deal better). The MCID for the GROC has not been specifically reported, but scores of +4 and +5 have typically been indicative of moderate changes in patient status (Jaeschke, Singer, and Guyatt, 1989) Scores of +3 to +5 are commonly used to identify "improved" versus "stable" patients for psychometric/clinimetric analyses (Young et al., 2018, 2019; Young, Cleland, Michener, and Brown, 2010).

Data analysis

We categorized patients into two mutually exclusive groups at the 3-month follow-up based on their GROC scores. Those scoring from −2 to +2 were considered clinically "'stable'" (minimal to no change), and those scoring +3 to +7 were considered clinically "improved" (at least somewhat better). Our analysis focused on patients who were "'stable'" and those who reported being "improved" at the 3-month follow-up.

Test–retest reliability was examined for the NPRS, PRTEE, and TEFS using "stable" patients by comparing scores at the initial examination with those at the 3-month follow-up. The intra-class correlation coefficient (ICC) was calculated and rated according to procedures described by Shrout and Fleiss ($ICC_{2,1}$) (Shrout and Fleiss, 1979). Values <0.5 indicate poor reliability, while values between 0.50–0.75, 0.75–0.90, and >0.90 moderate, good, and excellent agreement, respectively (Koo and Li, 2016).

Construct validity of the NPRS, PRTEE, and TEFS was examined by comparing the change in outcome scores for the "stable" and "improved" groups using separate, two-way analyses of variance for the repeated measures at baseline and reevaluation. We hypothesized that "stable" patients in each group would have NPRS, PRTEE, and TEFS intake values that did not change, whereas patients classified in the improved categories would demonstrate a significant change in their values. This would be represented by a significant group × time interaction. Pearson's correlation coefficient was also examined between all outcome measures.

Responsiveness, the ability of a measure to recognize change when change has occurred, was assessed for the NPRS, PRTEE, and TEFS using the clinically "stable" and "improved" groups at the 3-month follow-up point. Receiver operator characteristic (ROC) curves (Hanley and McNeil, 1982) were constructed by plotting sensitivity values (true-positive rate) on the y-axis and 1-specificity values (false-positive rate) on the x-axis for each level of change score. Separate ROC curves were constructed for the NPRS, PRTEE, and TEFS. The area under the curve (AUC) and the 95% CI were obtained as a method for determining the ability of each measure to distinguish improved patients from stable patients in each category. An AUC of 0.50 indicates that the measure has no diagnostic accuracy beyond chance, whereas a value of 1 suggests perfect accuracy (Hanley and McNeil, 1982). MCID, the smallest difference that patients perceive as beneficial, was calculated by identifying the point on the ROC curve nearest to the upper left-hand corner, which is considered to be the best cutoff score for distinguishing improved and stable patients.[22] Sensitivity and specificity values for the selected cutoff scores were also calculated. MDC, the amount of change that must be observed before the change can be considered to have exceeded measurement error, was calculated by determining the standard error of measurement (SEM) for the NPRS, PRTEE, and TEFS in the stable group (n = 23) from the ICC reliability analysis. The SEM was calculated using the formula SD/∏n, where SD is the standard deviation of the change score values, and n = the sample size. The SEM was multiplied by 1.65 to determine the 90% CI (MDC_{90}) and then multiplied by the ∏2 to account for the errors taken with repeated measurements.

Table 1. Baseline characteristics.

	N = 143 Mean ± SD
Gender: male/female	71/72
Weight: kg	81.9 ± 14.5
Age	42.7 ± 10.5
Months with elbow pain	9.3 ± 15.5
Medication use ≥1×/week (%)	128 (90%)
Numeric Pain rating Scale (0–10)	5.0 ± 1.6
Patient-Rated Tennis Elbow Evaluation (0–100)	42.4 ± 13.6
Tennis Elbow Function Scale (0–40)	19.6 ± 6.6
Number of treatment visits	7.0 ± 1.7

kg = kilograms, SD = standard deviation

Results

One hundred forty-three patients satisfied the inclusion and exclusion criteria, completed the study, and were included in the data analysis. Baseline characteristics are located in Table 1. The mean GROC score for all patients included in the analysis at the 3-month follow-up was +4.3 (SD + 2.6). The mean GROC score for the improved vs. stable groups was +5.3 (SD + 1.4) and +0.7 (SD + 1.5), respectively. At the 3-month follow-up, 117 (81.8%) patients were classified as improved, and 23 (16.1%) remained stable. There was a significant difference ($p < .001$) in mean change scores between stable and improved patients for the NPRS, PRTEE, and TEFS at the 3-month follow-up (Table 2). Additionally, all three outcome measures exhibited strong construct validity (Pearson's correlation coefficient ranging from 0.71 to 0.83, Table 3, Figure 2).

The test–retest ($ICC_{2,1}$) values and MDC_{90} calculated from the stable patients are reported in Table 4. At the 3-month follow-up, the NPRS ($ICC_{2,1}$: 0.54, 95% CI 0.17–0.78) had fair reliability, while the PRTEE ($ICC_{2,1}$: 0.62, 95% CI 0.21–0.86), and TEFS ($ICC_{2,1}$: 0.71, 95% CI

Table 2. Difference between change scores from baseline to 3-month follow-up on self-report outcomes.

	Improved GROC (+3 to +7) *N* = 117 Mean (SD)	Stable GROC (−2 to +2) *N* = 23 Mean (SD)	Mean Difference (95% CI)	*P*
NPRS	3.59 (1.5)	0.74 (1.4)	2.9 (2.2; 3.5)	*p*<.001
PRTEE	30.47 (12.4)	7.9 (7.1)	22.6 (17.3; 27.9)	*p*<.001
TEFS	14.38 (6.3)	3.5 (3.7)	10.9 (8.2; 13.6)	*p*<.001

NPRS = numeric pain rating scale (0–10), PRTEE = patient-rated tennis elbow evaluation (total score: 0–100), TEFS = tennis elbow functional scale (0–40), GROC = global rating of change scale (−7 to +7), CI = confidence interval.

Table 3. Pearson's correlation coefficient (r).

Outcome Measure	PRTEE r (95% CI)	TEFS r (95% CI)	GROC r (95% CI)
NPRS	0.82 (0.76; 0.87) *p <.001*	0.71 (0.62; 0.79) *p <.001*	0.79 (0.72; 0.84) *p <.001*
PRTEE	—–	0.83 (0.77; 0.88) *p <.001*	0.76 (0.68; 0.82) *p <.001*
TEFS	—–	—–	0.72 (0.63; 0.79) *p <.001*

NPRS = numeric pain rating scale, TEFS = tennis elbow functional scale, PRTEE = patient-rated tennis elbow evaluation (total score), GROC = global rating of change scale, CI = confidence interval.

Table 4. Reliability: 3-month follow-up.

Outcome Measure	$ICC_{2,1}$ (95% CI) Stable (GROC −2 to +2) *N* = 23	MDC_{90}
Numeric Pain Rating Scale	0.54 0.17, 0.78	1.4
Patient-Rated Tennis Elbow Evaluation	0.62 0.21; 0.86	9.7
Tennis Elbow Function Scale	0.71 0.14; 0.90	5.7

ICC = intraclass correlation coefficient, CI = confidence interval.
GROC = global rating of change.

Table 5. Area under the curve: 3-month follow-up.

Outcome Measure	AUC 95% CI Improved (GROC +3 to +7) *N* = 117
Numeric Pain Rating Scale	0.94 0.89; 0.98
Patient-Rated Tennis Elbow Evaluation	0.96 0.93; 0.99
Tennis Elbow Function Scale	0.95 0.91; 0.98

AUC = area under the curve, CI = confidence interval.
GROC = global rating of change.

Table 6. Meaningful clinically important difference values.

Outcome Measure	MCID Sn; Sp Improved (GROC +3 to +7) *N* = 117
Numeric Pain Rating Scale	2.3 0.74; 0.96
Patient-Rated Tennis Elbow. Evaluation	14.8 0.93; 0.86
Tennis Elbow Function Scale	7.5 0.85; 0.85

MCID = minimally clinically important difference, Sn = sensitivity.
Sp = specificity, GROC = global rating of change.

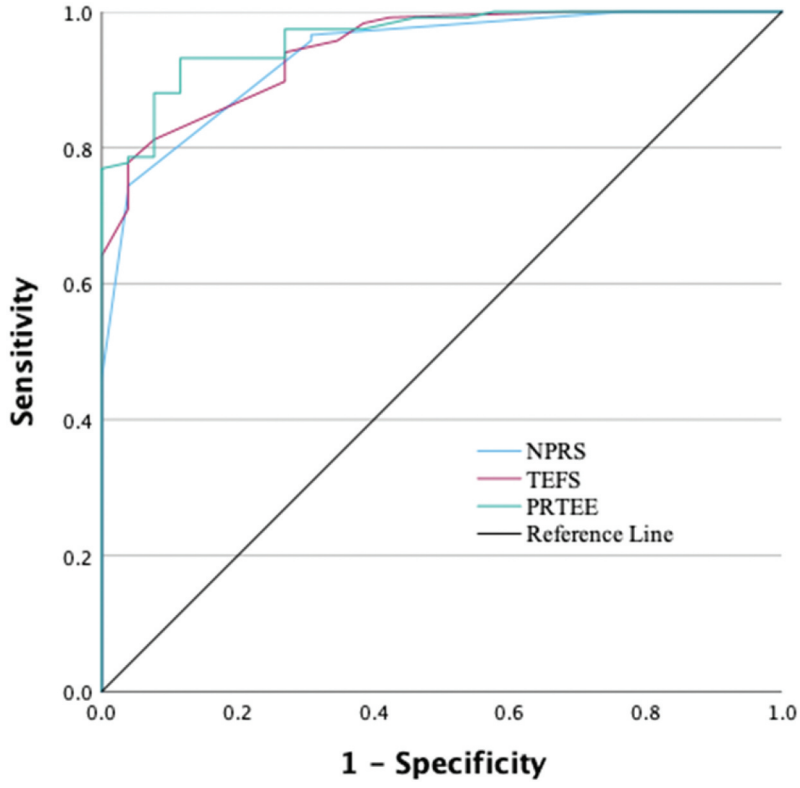

Figure 1. Receiver operating curves for all outcome measures. NPRS = numeric pain rating scale, TEFS = tennis elbow functional scale, PRTEE = patient-rated tennis elbow evaluation.

0.14–0.90) exhibited moderate reliability at the 3-month follow-up. At 3-months, the MDC_{90} was 1.4, 9.7, and 5.7 for the NPRS, PRTEE, and TEFS, respectively.

The NPRS, PRTEE, and TEFS all demonstrated excellent responsiveness (AUC range from 0.94 to 0.96) and are reported in Table 5. The MCID threshold and the sensitivity/specificity associated with each cutoff score are also located in Table 6. At 3-months, the MCID was 2.3, 14.8, and 7.5 for the NPRS, PRTEE, and TEFS, respectively.

Discussion

To date, this is the first and largest study to comprehensively examine the clinimetric properties of three different PROMs in patients diagnosed with LET at 3-month follow-up. The NPRS, PRTEE, and TEFS exhibited acceptable reliability, strong construct validity, and excellent responsiveness in this patient population. See Tables 2–6 and Figures 1, 2.

The PRTEE is the primary outcome measure recommended for patients with LET (Evans et al., 2019; Macdermid and Silbernagel, 2015). Prior analyses of the PRTEE have reported conflicting results in comparison to the current study (Shafiee et al., 2022). The PRTEE exhibited only moderate reliability (ICC 0.62) in this cohort of patients with LET. In a recent high-quality meta-analysis, Shafiee and MacDermid (Shafiee et al., 2022) pooled data from seven studies on chronic LET and found excellent test–retest reliability (ICC 0.97). This is a considerable difference compared to our study; however, our data included patients with an average of 9.3 months symptom duration. Furthermore, we examined patients who remained "stable" at the "3-month" follow-up, and our data is derived from the strict analysis of the English version, using absolute agreement between test–retest measures. It should also be noted that the average follow-up time frame was shorter (30 min) (Kaux et al., 2016; Leung, Yen, and Tse, 2004), 2 h (Altan, Ercan, and Konur, 2010), 4 weeks (Chung and Wiley, 2010), or unknown (Nilsson, Baigi, Marklund, and Månsson, 2008) in previous published studies, and only one of the seven studies identified the reliability cohort as being "stable" using a GROC (Chung and Wiley, 2010). It has been recommended that 1–2-week follow-up time frames are most appropriate for test–retest reliability (Streiner, Norman, and Cairney, 2014). The current analysis did not include a 1–4-week clinimetric analysis; however, in conditions like LET, longer retest time frames may be more appropriate, in order to evade recall bias (Shafiee et al., 2022). Further, time to full recovery for tendinopathy has been reported to be 3–6 months (Khan, Cook, Taunton, and Bonar, 2000) and can be much longer. Therefore, the current authors suggest addressing a longer-term retest follow-up in tendinopathy, in order to capture a more realistic combination of clinical and histological change.

In the current study, the MDC of the PRTEE was 9.7 points. Valuable pooled data from Shafiee and Macdermid (Shafiee et al., 2022) suggest that the estimated MDC for chronic LET is 6.5 (ranging from 0.58 to 11.0) (Shafiee et al., 2022). Considering our reliability was only moderate, with a longer follow-up, and a larger test–retest SD (±20.1) in the stable patients (*n*

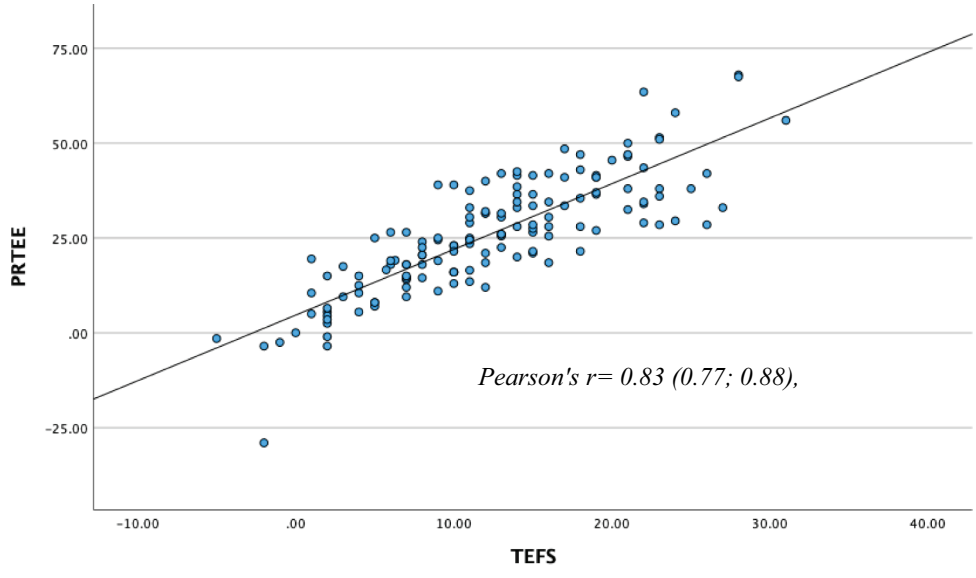

Figure 2. Scatter plot: 3-month change scores: PRTEE and TEFS. PRTEE = patient-rated tennis elbow evaluation (0–100), TEFS = tennis elbow function scale (0–40). *Pearson's r* = 0.83 (0.77; 0.88).

= 23), our results may be comparable to those of shorter-term reliability outcomes noted. Using the ROC curve based on the change scores of improved patients (GROC +3 to +7; $n = 117$), the MCID of the PRTEE was a 14.8-point change in the current study. Again, this finding is somewhat inconsistent in comparison to prior studies on the PRTEE. For example, only two prior studies used anchor-based analyses (GROC) and ROC curves, reporting MCID values of 8 points (Cacchio et al., 2012) and 7 points (GROC = a little better) to 11 points (GROC = much better) (Poltawski and Watson, 2011). A subgroup analysis also revealed an MCID of 21 points in patients with baseline PRTEE scores greater than 40 (Poltawski and Watson, 2011). The average baseline scores in the current study for the "improved" vs "not improved" groups were 43.1 (±13.1) and 39.1 (±15.5), respectively. This higher baseline score seems to be in-line with our higher threshold for the MCID (14.8 points) (Poltawski and Watson, 2011). Unfortunately, the two aforementioned anchor-based studies had much smaller sample sizes [$n = 49$ (Cacchio et al., 2012), $n = 57$ (Poltawski and Watson, 2011)] than the current study ($n = 117$). Additionally, the sample size was not reported in the subgroup analysis on PRTEE baseline scores >40 (Poltawski and Watson, 2011). Overall, the MCID found in the current study (14.8 points) encompasses the best anchor-based ROC combination of sensitivity/specificity in patients, ranging from "somewhat better" to "a very great deal better," and lies strategically between the 7 and 21 point estimates previously reported (Cacchio et al., 2012; Poltawski and Watson, 2011). Finally, our estimation of the MCID in patients who report improvement in a time frame (3-months) that supports a more realistic range for clinical/histological improvement of tendinopathy, seems to be of paramount importance. Future studies should consider even longer follow-up time frames with the addition of diagnostic ultrasound imaging for assessing tendon health.

The TEFS or NPRS are not currently recommended as one of the preferred PROMs for patients with LET (Evans et al., 2019; Macdermid and Silbernagel, 2015). This argument can be supported, with the TEFS having only a single study from 1999 reporting its most fundamental clinimetric properties (Lowe, 1999), and the NPRS without clinimetric analyses as a stand-alone or adjunct "self-report" outcome in patients with LET (Evans et al., 2019; Macdermid and Silbernagel, 2015). To date, the current study is the first to perform thorough clinimetric analyses of both the TEFS, and the NPRS in a large cohort of patients with LET. Based on our results, the TEFS and NPRS demonstrated moderate reliability and excellent anchor-based responsiveness in this patient population (Tables 4–5). The TEFS and NPRS also demonstrated strong construct validity with the other primary and secondary self-report outcomes (PRTEE, GROC) used in this study (Table 3, Figure 2).

The current analysis does not come without limitations. First, we did not provide any subgroup analyses, including age, gender, or baseline scoring. Second, clinicians and researchers should interpret the moderate reliability found in all three PROMs with caution, secondary to wider confidence intervals. Third, the results of this analysis are short term in nature, not expanding beyond 3-month follow-up. Future research should address the above limitations, and any innovative assessment tools used for patients with LET.

Conclusion

This study examined the clinimetric properties of three PROMs in a large sample of patients treated for LET. The NPRS, PRTEE, and TEFS all exhibited moderate reliability and strong construct validity at the 3-month follow-up. Additionally, the NPRS, PRTEE, and TEFS exhibited a high level of responsiveness over time. At the 3-month follow-up, clinicians and researchers should expect a 2.3-point change on the NPRS, a 14.8-point change on the PRTEE, and a 7.5-point change on the TEFS to be considered clinically meaningful after 4 weeks of intervention. The NPRS, PRTEE, and TEFS seem to be well suited as self-reported outcome measures for patients with LET. Future trials and clinimetric analyses of LET should incorporate a 3- to 6-month follow-up, including diagnostic ultrasound of the healing tendon, to further enhance the validity of our self-report and clinical outcomes.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The author(s) reported that there is no funding associated with the work featured in this article.

ORCID

Ian Young PT DSc, OCS, RMSK http://orcid.org/0000-0001-6559-993X
James Dunning DPT, PhD, FAAOMPT http://orcid.org/0000-0002-1194-0108
Firas Mourad PT, PhD http://orcid.org/0000-0002-8981-2085

James Escaloni DPT, FAAOMPT, RMSK http://orcid.org/0000-0002-1646-8594

References

Altan L, Ercan I, Konur S 2010 Reliability and validity of Turkish version of the patient rated tennis elbow evaluation. Rheumatology International 30: 1049–1054. 10.1007/s00296-009-1101-6

Bisset L, Beller E, Jull G, Brooks P, Darnell R, Vicenzino B 2006 Mobilisation with movement and exercise, corticosteroid injection, or wait and see for tennis elbow: Randomised trial. The British Medical Journal 333: 939. 10.1136/bmj.38961.584653.AE

Cacchio A, Necozione S, MacDermid JC, Rompe JD, Maffulli N, di Orio F, Santilli V, Paoloni M 2012 Cross-cultural adaptation and measurement properties of the Italian version of the patient-rated tennis elbow evaluation (PRTEE) questionnaire. Physical Therapy 92: 1036–1045. 10.2522/ptj.20110398

Chung BPH, Wiley JP 2010 Validity, responsiveness and reliability of the patient-rated tennis elbow evaluation. Hand Therapy 15: 62–68. 10.1258/ht.2010.010014

Coombes BK, Bisset L, Brooks P, Khan A, Vicenzino B 2013 Effect of corticosteroid injection, physiotherapy, or both on clinical outcomes in patients with unilateral lateral epicondylalgia: A randomized controlled trial. Journal of the American Medical Association 309: 461–469. 10.1001/jama.2013.129

Di Filippo L, Vincenzi S, Pennella D, Maselli F 2022 Treatment, diagnostic criteria and variability of terminology for lateral elbow pain: Findings from an overview of systematic reviews. Healthcare (Basel) 10: 1095. 10.3390/healthcare10061095

Dones VC 3rd, Grimmer K, Thoirs K, Suarez CG, Luker J 2014 The diagnostic validity of musculoskeletal ultrasound in lateral epicondylalgia: A systematic review. BMC Medical Imaging 14: 10. 10.1186/1471-2342-14-10

Dunning J, Mourad F, Bliton P, Charlebois C, Gorby P, Zacharko N, Layson B, Maselli F, Young I, Fernández-de-Las-Peñas C 2024 Percutaneous tendon dry needling and thrust manipulation as an adjunct to multimodal physical therapy in patients with lateral elbow tendinopathy: A multicenter randomized clinical trial. Clinical Rehabilitation 38: 1063–1079. 10.1177/02692155241249968

Evans JP, Porter I, Gangannagaripalli JB, Bramwell C, Davey A, Smith CD, Fine N, Goodwin VA, Valderas JM 2019 Assessing patient-centred outcomes in lateral elbow tendinopathy: A systematic review and standardised comparison of English language clinical rating systems. Sports Medicine - Open 5: 10. 10.1186/s40798-019-0183-2

Farrar JT, Young JP Jr., LaMoreaux L, Werth JL, Poole RM 2001 Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. Pain 94: 149–158. 10.1016/S0304-3959(01)00349-9

Hanley JA, McNeil BJ 1982 The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143: 29–36. 10.1148/radiology.143.1.7063747

Hefford C, Abbott JH, Arnold R, Baxter GD 2012 The patient-specific functional scale: Validity, reliability, and responsiveness in patients with upper extremity musculoskeletal problems. Journal of Orthopaedic & Sports Physical Therapy 42: 56–65. 10.2519/jospt.2012.3953

Jaeschke R, Singer J, Guyatt GH 1989 Measurement of health status. Ascertaining the minimal clinically important difference. Controlled Clinical Trials 10: 407–415. 10.1016/0197-2456(89)90005-6

Jensen MP, Karoly P, Braver S 1986 The measurement of clinical pain intensity: A comparison of six methods. Pain 27: 117–126. 10.1016/0304-3959(86)90228-9

Jomaa G, Kwan CK, Fu SC, Ling SK, Chan KM, Yung PS, Rolf C 2020 A systematic review of inflammatory cells and markers in human tendinopathy. BMC Musculoskeletal Disorders 21: 78. 10.1186/s12891-020-3094-y

Kaux JF, Delvaux F, Schaus J, Demoulin C, Locquet M, Buckinx F, Beaudart C, Dardenne N, Van Beveren J, Croisier JL, et al. 2016 Cross-cultural adaptation and validation of the patient-rated tennis elbow evaluation questionnaire on lateral elbow tendinopathy for French-speaking patients. Journal of Hand Therapy 29: 496–504. 10.1016/j.jht.2016.06.007

Khan KM, Cook JL, Taunton JE, Bonar F 2000 Overuse tendinosis, not tendinitis: Part 1: A new paradigm for a difficult clinical problem. The Physician and Sportsmedicine 28: 38–48. 10.3810/psm.2000.05.890

Koo TK, Li MY 2016 A guideline of selecting and reporting intraclass correlation coefficients for reliability research. Journal of Chiropractic Medicine 15: 155–163. 10.1016/j.jcm.2016.02.012

Kraushaar BS, Nirschl RP 1999 Current concepts review - tendinosis of the elbow (tennis elbow). Clinical features and findings of histological, immunohistochemical, and electron microscopy studies. Journal of Bone and Joint Surgery 81: 259–278. 10.2106/00004623-199902000-00014

Leung HB, Yen CH, Tse PY 2004 Reliability of Hong Kong Chinese version of the patient-rated forearm evaluation questionnaire for lateral epicondylitis. Hong Kong Medical Journal 10: 172–177.

Lowe KA 1999 The test retest reliability, construct validity, and responsiveness of the tennis elbow function scale.

Lucado AM, Day JM, Vincent JI, MacDermid JC, Fedorczyk J, Grewal R, Martin RL 2022 Lateral elbow pain and muscle function impairments. Journal of Orthopaedic & Sports Physical Therapy 52: Cpg1–cpg111. 10.2519/jospt.2022.0302

Macdermid JC, Silbernagel KG 2015 Outcome evaluation in Tendinopathy: Foundations of assessment and a summary of selected measures. Journal of Orthopaedic & Sports Physical Therapy 45: 950–964. 10.2519/jospt.2015.6054

Mercieca-Bebber R, King MT, Calvert MJ, Stockler MR, Friedlander M 2018 The importance of patient-reported outcomes in clinical trials and strategies for future optimization. Patient related outcome measures 9: 353–367. 10.2147/PROM.S156279

Nilsson P, Baigi A, Marklund B, Månsson J 2008 Cross-cultural adaptation and determination of the reliability and validity of PRTEE-S (Patientskattad Utvärdering av Tennisarmbåge), a questionnaire for patients with lateral epicondylalgia, in a Swedish population. BMC Musculoskeletal Disorders 9: 79. 10.1186/1471-2474-9-79

Poltawski L, Watson T 2011 Measuring clinically important change with the Patient-rated Tennis Elbow Evaluation. Hand Therapy 16: 52–57. 10.1258/ht.2011.011013

Rompe JD, Overend TJ, MacDermid JC 2007 Validation of the Patient-rated Tennis Elbow Evaluation Questionnaire. Journal of Hand Therapy 20: 3–10. 10.1197/j.jht.2006.10.003

Salaffi F, Stancati A, Silvestri CA, Ciapetti A, Grassi W 2004 Minimal clinically important changes in chronic musculoskeletal pain intensity measured on a numerical rating scale. European Journal of Pain 8: 283–291. 10.1016/j.ejpain.2003.09.004

Sanders TL, Maradit Kremers H, Bryan AJ, Ransom JE, Smith J, Morrey BF 2015 The Epidemiology and Health Care Burden of Tennis Elbow:A Population-Based Study. The American Journal of Sports Medicine 43: 1066–1071. 10.1177/0363546514568087

Schneeberger AG, Masquelet AC 2002 Arterial vascularization of the proximal extensor carpi radialis brevis tendon. Clinical Orthopedics and Related Research 398: 239–244. 10.1097/00003086-200205000-00033

Scott A, Backman LJ, Speed C 2015 Tendinopathy: Update on Pathophysiology. Journal of orthopaedic & sports physical therapy 45: 833–841. 10.2519/jospt.2015.5884

Shafiee E, MacDermid JC, Walton D, Vincent JI, Grewal R 2022 Psychometric properties and cross-cultural adaptation of the Patient-Rated tennis elbow evaluation (PRTEE); a systematic review and meta-analysis. Disability and Rehabilitation 44: 5402–5417. 10.1080/09638288.2021.1938248

Shrout PE, Fleiss JL 1979 Intraclass correlations: uses in assessing rater reliability. Psychological Bulletin 86: 420–428. 10.1037/0033-2909.86.2.420

Streiner DL, Norman GR, Cairney J 2014 Health Measurement Scales: A practical guide to their development and use. Oxford, United Kingdom: Oxford University Press.

Vincent J, MacDermid JC 2014 Patient-rated tennis elbow evaluation questionnaire. Journal of Physiotherapy 4: 240. 10.1016/j.jphys.2014.08.002

Voleti PB, Buckley MR, Soslowsky LJ 2012 Tendon healing: repair and regeneration. Annual review of biomedical engineering 14: 47–71. 10.1146/annurev-bioeng-071811-150122

Young IA, Cleland JA, Michener LA, Brown CJ 2010 Reliability, Construct Validity, and Responsiveness of the Neck Disability Index, Patient-Specific Functional Scale, and Numeric Pain Rating Scale in Patients with Cervical Radiculopathy. American Journal of Physical Medicine and Rehabilitation 89: 831–839. 10.1097/PHM.0b013e3181ec98e6

Young IA, Dunning D, Butts J, R P, Mourad F, Cleland JA 2018 Reliability, construct validity, and responsiveness of the neck disability index and numeric pain rating scale in patients with mechanical neck pain without upper extremity symptoms. Physiotherapy Theory and Practice 35: 1328–1335. 10.1080/09593985.2018.1471763

Young IA, Dunning J, Butts R, Cleland JA, Fernandez-de-Las-Penas C 2019 Psychometric properties of the Numeric Pain Rating Scale and Neck Disability Index in patients with cervicogenic headache. Cephalalgia 39: 44–51. 10.1177/0333102418772584