Reliability of Hong Kong Chinese version of the Patient-rated Forearm Evaluation Questionnaire for lateral epicondylitis

Objective. To determine the reliability and validity of a dedicated assessment tool for lateral epicondylitis after translation into Hong Kong Chinese.

Design. Cross-sectional study.

Setting. District hospital, Hong Kong.

Patients. Seventy-four patients, 12 of whom were bilingual, were recruited (total of 82 elbows).

Main outcome measures. Translation equivalence and reliability were measured with the test-retest method. Validity was measured against the Roles and Maudsley outcome score and mean maximal grip strength.

Result. The Patient-rated Forearm Evaluation Questionnaire had high English-Chinese equivalence (Spearman’s rho correlation coefficient = 0.926; P<0.001). It was also very reliable (intraclass correlation coefficient = 0.99; P<0.001). Validity according to the Roles and Maudsley outcome score and mean of maximal grip strength was significant (P<0.01).

Conclusion. The Hong Kong Chinese version of Patient-rated Forearm Evaluation Questionnaire is a reliable and valid assessment tool for chronic lateral epicondylitis. Its equivalence to the original English version makes outcome assessment across cultural barrier feasible.

Introduction

Lateral epicondylitis, also known as tennis elbow, was first described by Rungue in 1873.1 It is a prevalent disorder characterised by an insidious onset of pain over the lateral aspect of the elbow, which is aggravated by wrist extension and gripping. The incidence and prevalence varies, partly because more than half of patients do not seek medical advice.2 However, about four in 1000 adults are affected each year.3 Since 1966, more than 185 articles have been published on the subject and more than 40 different treatments proposed. Nevertheless, many of these treatments are not supported by strong scientific evidence.4 In addition, comparison on different treatment modalities is difficult, if not impossible, when taking into account that
there is no consensus on the measurement of treatment outcome.\(^5\)

The Roles and Maudsley outcome score grades lateral epicondylitis into four categories of severity—namely, excellent, good, acceptable, and poor. 'Excellent' refers to cases with no pain, full movement, and full activity. 'Good' implies occasional discomfort but full movement and activity. 'Acceptable' denotes some discomfort after prolonged activities. 'Poor' signifies pain that is severe enough to limit activities.\(^6\) Stratification into only four levels of severity, however, lacks sensitivity to document changes in clinical condition. In contrast, measuring the strength of the handgrip or measuring forearm endurance are objective methods, although they cannot document the impact of the condition on daily function.

The current trend is for assessment to take into consideration patients’ standpoint; this approach can assist the clinician in understanding the extent of patient suffering. Such an assessment tool should have a high degree of reliability, which is fundamental to determining its effectiveness in clinical use.\(^7\) Reliability can be regarded as the consistency of repeated measurement under the same conditions, and can be documented by a number of established methodologies. Furthermore, validity is considered as the accuracy or the truth of measurement. However, pain is difficult to quantify and there is no gold standard. To the best of our knowledge, there is only one assessment tool that is simple and reliable, and which has been scientifically scrutinised—the Patient-rated Forearm Evaluation Questionnaire (PRFEQ), which was designed specifically for patients with lateral epicondylitis.

The questionnaire provides a brief (it takes 5 minutes to complete), uncomplicated, and standardised quantitative description of pain and functional disability. Moreover, it has been validated.\(^8\)

Translating the questionnaire into Hong Kong Chinese not only provides a simple and uniform medical record documentation, but also allows cross-cultural and international comparisons in regard of this disease entity.\(^9\) Translation on a word-for-word basis definitely addresses the language barrier problem. However, failure to consider local cultural factors may result in significant bias.\(^10\) Hence, when a survey instrument is translated into another language, it must undergo stringent validation before use.\(^9\)

**Materials and methods**

**Questionnaire**

The original PRFEQ tool was developed from two sources. The first was from a study by Stratford et al,\(^7\) who assessed the reliability, validity, and sensitivity of visual analogue scales of pain and function, as well as a set of items on pain-free function among patients with lateral epicondylitis. The instrument was highly reliable, moderately valid, and sensitive to changes. The second source was a wrist evaluation questionnaire adopted in the Hand and Upper Limb Centre of St Joseph’s Health Centre in London, Ontario, Canada. Questionnaire items were generated from an international survey, expert and patient opinion, and review and research findings. The test-retest reliability was 0.90, and the instrument had high validity with respect to the Short Form 36 Health Survey.\(^11\)

The PRFEQ assesses the average pain and function of the affected arm during the preceding week. This time frame allows an accurate memory recall, while avoiding effects from acute fluctuations in symptoms. The questionnaire consists of two parts: part 1 deals with pain and part 2 deals with function. Each of the five items in part 1 is scored using a 10-cm visual numeric rating scale, ranging from 0 (no pain) to 10 (worst pain imaginable).\(^12,13\) Similarly, the 10 items of part 2 use a scale of 0 (no difficulty) to 10 (unable to perform an activity) to rate function. In both parts, patients are requested to place a mark along a line ranging from 0 to 10, with no other descriptors placed along the line.\(^14\) The total score is the combined score for all items from both parts, ranging from 0 (no pain and no functional impairment) to 150 (worst pain imaginable with a very significant function deficit).

**Translation**

Written copyright permission was obtained from Dr TJ Overend,\(^8\) the chief investigator of the original article describing the PRFEQ. We translated the questionnaire with the assistance of bilingual orthopaedic specialists, specialist nurses, and occupational therapists. All of them were of southern Chinese ancestry and had been able to speak, read, and write English and Cantonese fluently since childhood. All translators were informed of the purpose of the study. Forward translation was performed by two orthopaedic surgeons, and the initial translation was modified according to suggestions of the other translators. Each item was then back-translated, with near equivalence to the original. The English and the Chinese versions are shown in Appendices 1 and 2, respectively.

**Study participants**

We recruited 74 patients (82 elbows) who attended the General Orthopaedics Clinic at Kwong Wah Hospital between July 2001 and March 2002. Participants’ ages ranged from 28 to 69 years. All patients had a clinical diagnosis of lateral epicondylitis on the basis on stringent inclusion and exclusion criteria.

The inclusion criteria were as follows:

1. Six-month duration of symptoms, and
2. Pain could be induced by two or more of the following tests:
   a. Palpation of the lateral epicondyle;
   b. Resisted wrist extension (Thomsen test). With the shoulder flexed to 60°, elbow extended, forearm pronated, and wrist extended at 30°, the patient is asked to extend and radial deviate the hand.

Hong Kong Med J Vol 10 No 3 June 2004 173
while resistance is applied over the second and third metacarpal bones;
(c) Resisted finger extension. With the shoulder flexed at 60°, elbow extended, forearm pronated, and wrist extended to 30°, the patient is asked to extend a finger with resistance applied over the second to fifth proximal phalanges; and
(d) Chair test. The patient is asked to lift a 3.5-kg chair with the shoulder flexed at 60° and elbow extended.

The following exclusion criteria were also observed strictly:
(1) Arthritic elbow suggested either by a limitation on range of motion, system arthritis, or abnormal X-ray;
(2) Soft tissue infection;
(3) Clinical radial tunnel syndrome (pain distal to lateral epicondyle);
(4) Cervical radiculopathy or myelopathy suggested by sensory disturbance;
(5) Thoracic outlet syndrome;
(6) History of fracture or dislocation at elbow;
(7) History of elbow surgery;
(8) Local injection of steroid or local anaesthetic in the past 6 months, or
(9) Illiterate, poor mental state.

Twelve patients were bilingual. To prove their proficiency, they were requested to read aloud the introductory paragraph printed on the questionnaire. All enrolled patients provided their written consent. The protocol had been approved by the Kwong Wah Hospital Ethics Committee.

Methodology
Subjects were requested not to take any analgesics or non-steroidal anti-inflammatory drugs for 2 days, before attending a short briefing about the questionnaire’s background, format, and the time frame under question. An orthopaedic surgeon was available to answer queries while participants were answering the questionnaire. All patients had to complete the PRFEQ twice, with a half-hour period in between, during which handgrip strength and the Roles and Maudsley outcome score were recorded. At no time were allowed to refer to the first questionnaire while they were completing the second.

Maximal isometric grip strength was measured with a Jamar hand dynamometer (JA Preston Corp, Jackson, United States) with its display covered and handle locked at the second closest position. The dynamometer has five preset positions for its handle. Locking the handle at the second closest position is most appropriate for Chinese people. Patients were requested to take the measurement for both sides. Two positions were used. In the first, the patient stood and held the elbow in 90° flexion beside the body with the forearm unsupported and in neutral rotation (thumb-up position), with the wrist in neutral position (approximately 30° of extension). In the second position, the elbow was held in full extension with the shoulder in 60° of flexion and neutral abduction. Positions were unrestrained once effort had been made. A total of three trials of maximal grip strength were made for both sides alternatively.

Questionnaire equivalence, reliability, and validity
Equivalence of the English and Hong Kong Chinese versions were measured by analysing the responses from the 12 bilingual patients. They were asked randomly to complete the questionnaires in either language.

Data analysis
Data were analysed using the Statistical Package for Social Science (Windows version 10.0; SPSS Inc, Chicago, United States). The probability cut-off level for significance was set at P<0.05. Reliability was judged by the test-retest method and expressed as an intraclass correlation coefficient.15 Criterion validity was measured using the Roles and Maudsley outcome score; both the outcome score and the grip strength were the external standard. One-way analysis of variance (ANOVA) and Spearman’s rho correlation test were used when appropriate.

Results
The total score ranged from 15 to 137. The mean score was 75.5 while the standard deviation was 32. Further descriptive information was given in Table 1.

Validation of translation
Validation of the translation from English to Hong Kong Chinese was performed by Spearman’s rho correlation test. The correlation coefficient was 0.916, 0.926, and 0.963 for pain, function, and total scores, respectively (P<0.001 for each). The difference in total score between the two translations was -16 to 13, with a mean of 0.33 points higher.
for the Hong Kong Chinese version, compared with the English.

Reliability and validity

Test-retest reliability for pain, function subscales, and the overall PRFEQ scores was high (Table 2). Criterion validity was measured against the Roles and Maudsley outcome score and the maximal grip strength. Patients scored higher in pain and function subscales, and had higher total scores when they were classified in progressively poorer outcome according to the Roles and Maudsley outcome score (Table 3). By means of ANOVA, the scores were once again validated against the Roles and Maudsley outcome scores (P values were <0.001 for pain, function, and total scores). The PRFEQ score showed a statistically significant negative correlation with the maximal grip strength at different elbow positions (Table 4).

Discussion

The PRFEQ Hong Kong Chinese version seems to be a reliable tool for assessing pain and function in patients with chronic lateral epicondylitis. There is no consensus on how high a correlation coefficient should be to show clinical significance. Fleiss interpreted intraclass correlation coefficients of 0.00 to 0.40 as ‘poor’, 0.40 to 0.75 as ‘fair to good’, and higher than 0.75 as ‘excellent’. The intraclass correlation coefficients for the PRFEQ Hong Kong Chinese version remained higher than 0.90, which is comparable to the result for original English version. Therefore, both versions of the PRFEQ should be regarded as a highly reliable tool.

The standard error of measurement (SEM) provides an estimate of absolute reliability. In clinical terms, a single score estimates the true score, and the SEM can be viewed as the standard deviation of a score. In this study, the SEM of the PRFEQ total score was 3.28. Thus, when an individual scores 75, the true score might lie between 68 and 82 ie a score within two SEMs. If one is confident that a real change has occurred, the participant should report a score outside this range. Despite the high intraclass correlation coefficients, the confidence intervals are modest.

As for validity testing, when a gold standard exists, it is preferable. More typically, new instruments are compared with more established instruments. In this study, we elected to set the criterion comparators as the Roles and Maudsley outcome score and the mean of maximal grip strength.

Although the correlation was statistically significant (P=0.001) [Table 4], it was not strong enough to unerringly predict or explain the grip strength by the PRFEQ score (r = –0.40). However, this finding should not be considered as failure of the PRFEQ. Rather, it might indicate that such anthropometrical tests cannot detect the actual deficit in function. Making a powerful grip and completing a functional task are just two very different indicators. Good hand function requires dexterity and interplay of intact sensory and motor functions. Similar conclusions were drawn for knee flexor strength and gross motor ability, and for hip abductor torque and 6-minute walking distance.

Standardised patient questionnaires offer advantages in evaluation because they are client-centred and time-

<table>
<thead>
<tr>
<th>Table 2. Reliability of the Patient-rated Forearm Evaluation Questionnaire</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Intraclass</strong></td>
</tr>
<tr>
<td><strong>correlation</strong></td>
</tr>
<tr>
<td>Pain</td>
</tr>
<tr>
<td>Function</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 3. Scoring of patients according to different Roles and Maudsley outcome score</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>No. of elbows</strong></td>
</tr>
<tr>
<td>Excellent</td>
</tr>
<tr>
<td>Good</td>
</tr>
<tr>
<td>Acceptable</td>
</tr>
<tr>
<td>Poor</td>
</tr>
</tbody>
</table>

* SD standard deviation
† A higher score compared with the next better category; P<0.05

<table>
<thead>
<tr>
<th>Table 4. Mean of maximal grip strength at two different elbow positions in relation to the Patient-rated Forearm Evaluation Questionnaire</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Grip strength with flexed elbow</strong></td>
</tr>
<tr>
<td>Pain score</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Function score</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Total score</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
efficient, and they allow comparisons of clinical outcome to be made. They supplement, but do not replace, other components of clinical evaluation, such as history taking and physical examination. In this study, the Hong Kong Chinese version of the PRFEQ has been shown to be a reliable and valid assessment tool. Its use should benefit clinicians and researchers in the assessment of chronic lateral epicondyliitis.

References


Appendix 1. The validated Patient-rated Forearm Evaluation Questionnaire dedicated for lateral epicondyliitis

- The questions below help us understand how much pain and difficulty you have had with your arm in the past week. You will be describing your average arm symptoms over the past week on a scale of 0-10.
- Please provide an answer for ALL questions on both sides of the questionnaire. If you did not perform an activity listed, please provide an ESTIMATE of the pain or difficulty you would expect if you performed that activity. If you never perform the activity, draw a line completely through the answer key.

<table>
<thead>
<tr>
<th>1. PAIN with affected arm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Please note the average amount of pain in your arm over the past week by circling the number that best describes your pain on a scale from 0-10. A ’0’ means that you did not have any pain and a ’10’ means that you had the worst pain imaginable.</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td>no pain</td>
</tr>
</tbody>
</table>

Rate your difficulty (over the past week)

<table>
<thead>
<tr>
<th>When you are at rest</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>When doing a task with repeated arm movement</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>When carrying a plastic bag of groceries</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>When your pain was at its least</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>When your pain was at its worst</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
</tbody>
</table>
2. FUNCTION with affected arm

A. SPECIFIC ACTIVITIES
Rate how much difficulty you experienced performing each of the items listed below with your affected arm (over the past week) by circling the number that best describes your difficulty on a scale of 0-10. A ‘0’ means that you did not experience any difficulty with your affected arm and a ‘10’ means that it was so difficult you were unable to do it at all.

<table>
<thead>
<tr>
<th>Activity</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
</table>

Rate your difficulty in the affected arm (over the past week)

Turning a doorknob 0 1 2 3 4 5 6 7 8 9 10
Carrying a plastic bag of groceries 0 1 2 3 4 5 6 7 8 9 10
Lifting a full coffee cup or glass to your mouth 0 1 2 3 4 5 6 7 8 9 10
Opening a jar 0 1 2 3 4 5 6 7 8 9 10
Pulling up pants 0 1 2 3 4 5 6 7 8 9 10
Wringing out a facelcloth or dishrag 0 1 2 3 4 5 6 7 8 9 10

B. USUAL ACTIVITIES
Rate how much difficulty you experienced performing your usual activities in each of the areas listed below (over the past week) by circling the number that best describes your difficulty on a scale of 0-10. By usual activities, we mean activities you performed before you started having a problem with your arm. A ‘0’ means that you did not experience any difficulty and a ‘10’ means that it was so difficult you were unable to do any of your usual activities.

Rate your difficulty (over the past week)

Personal care activity (ie dressing, washing) 0 1 2 3 4 5 6 7 8 9 10
Household work (maintenance, cleaning) 0 1 2 3 4 5 6 7 8 9 10
Work (your usual job) or main activity if not employed 0 1 2 3 4 5 6 7 8 9 10
Recreation or sporting activities 0 1 2 3 4 5 6 7 8 9 10

Appendix 2. Hong Kong Chinese version of the Patient-rated Forearm Evaluation Questionnaire

- 此問卷可幫助我們明白閣下因關節炎所引致的問題及困難，請你以0-10 分評估出上星期你患肢的平均病徵。
- 請回答所有問題(共15 項)。如果你上星期並沒有做以下列出的動作，請給予我們一個估計。

1. 患側手肘的痛楚

請用0-10 為規準，圈出你上星期患肢肘的平均痛楚。0' 指你沒有任何痛楚。'10'分是指痛楚的程度是超乎你所想的。

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
</table>

請圈出你上週痛楚的程度

1. 當你休息時 (睡眠時除外) 0 1 2 3 4 5 6 7 8 9 10
2. 當你重覆使用患肢時 0 1 2 3 4 5 6 7 8 9 10
3. 當你痛楚最輕微時(如沒有痛請圈0) 0 1 2 3 4 5 6 7 8 9 10
4. 當你最痛時 0 1 2 3 4 5 6 7 8 9 10
5. 請上一週患肢的選物回家 0 1 2 3 4 5 6 7 8 9 10

2. 患肢的功能

A. 指定動作

以0-10 為規準，請圈出你上星期以患肢做以下動作時的平均困難程度。0' 指你沒有任何困難。'10'分指你不能用患肢做到指定的動作。

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
</table>

請圈出你上週在下列情況時困難程度 (在使用患肢時):

6. 拔開門鎖 0 1 2 3 4 5 6 7 8 9 10
7. 提著一袋超級市場的雜物回家 0 1 2 3 4 5 6 7 8 9 10
8. 提著一杯水來飲 0 1 2 3 4 5 6 7 8 9 10
9. 提著一個籃子 0 1 2 3 4 5 6 7 8 9 10
10. 蒸繃 0 1 2 3 4 5 6 7 8 9 10
11. 扳乾洗面巾 0 1 2 3 4 5 6 7 8 9 10

B. 日常活動

以0-10 為規準，請圈出你上星期以患肢做以下活動時的平均困難程度。0' 指你沒有任何困難。'10'分指你不能用患肢做到該類活動。

請圈出你上週在下列情況時困難程度 (在使用患肢時):

12. 自我護理(如浴室更衣) 0 1 2 3 4 5 6 7 8 9 10
13. 日常家務(如洗繩) 0 1 2 3 4 5 6 7 8 9 10
14. 日常工作(如你有職的活)或平日生活(如你沒有上班的活) 0 1 2 3 4 5 6 7 8 9 10
15. 誠或運動時 0 1 2 3 4 5 6 7 8 9 10

Hong Kong Med J Vol 10 No 3 June 2004  177