The McGill Pain Questionnaire, Japanese version, reconsidered: Confirming the reliability and validity

Mamoru Hasegawa MD, Suguru Hattori MA, Keiji Ishizaki MD PhD, Shosuke Suzuki MD PhD, Fumio Goto MD PhD

The McGill Pain Questionnaire, Japanese version, reconsidered: Confirming the reliability and validity.

BACKGROUND: The McGill Pain Questionnaire (MPQ) is one of the most widely used instruments in the world to evaluate patients with chronic pain. However, differences in languages and cultural backgrounds have hindered its standardization in Japan.

OBJECTIVE: To standardize the MPQ in Japan.

DESIGN: The reliability and validity of a Japanese version of the MPQ (JMPQ) were examined using a translation-based methodology, which followed a format similar to the original MPQ.

SETTING: Multidisciplinary pain treatment centre of a university hospital in Japan.

PATIENTS: Consecutive out-patients (n=152) with chronic pain.

METHODS: Each patient completed the JMPQ, other pain rating scales (visual analogue scale, verbal rating scale, numerical rating scale) and the state-trait anxiety inventory. A subset of these patients (n=40) were tested again two weeks later.

RESULTS: Acceptable levels of reliability and validity of the JMPQ, and independence of the JMPQ subscales from other pain rating scales were confirmed by principal component analysis. Chronic pain patients did not show marked levels of anxiety as might have been expected.

CONCLUSIONS: The findings suggest that the JMPQ possesses sufficient merits as a pain rating scale from the standpoint of its reliability and validity. Furthermore, it is suggested that the JMPQ can be used to measure qualitative aspects of pain transcending differences in linguistic characteristics.

Key Words: Japanese language, McGill Pain Questionnaire, Reliability, Validity

Le questionnaire McGill sur la douleur, version japonaise revue : confirmation de sa fiabilité et de sa validité

HISTORIQUE : Le questionnaire McGill sur la douleur (MPQ) est un instrument des plus utilisés au monde pour évaluer les patients atteints d’une douleur chronique. Toutefois, les différences linguistiques et culturelles ont retardé sa normalisation au Japon.

OBJECTIF : Normaliser le MPQ au Japon.

MÉTHODE : La fiabilité et la validité de la version japonaise du MPQ (JMPQ) ont été étudiées en utilisant une méthodologie basée sur la traduction, qui était calquée sur le modèle d’origine du MPQ.

CONTEXTE : Un centre de traitement de la douleur pluridisciplinaire dans un hôpital universitaire au Japon.

SUJETS : Suite de patients ambulatoires (n=152) atteints d’une douleur chronique.

MÉTHODES : Chaque patient a complété le JMPQ, d’autres échelles d’évaluation de la douleur (échelle visuelle analogique, échelle d’évaluation verbale, échelle d’évaluation numérique) et l’inventaire de l’anxiété State-Trait. Un sous-groupe de ces patients (n=40) ont passé un autre test deux semaines plus tard.


CONCLUSION : Dans les limites de la présente étude, ces résultats nous incitent à penser que le JMPQ a des qualités suffisantes comme échelle d’évaluation de la douleur du point de vue de sa fiabilité et de sa validité. De plus, on encourage l’utilisation du JMPQ pour mesurer les aspects qualitatifs de la douleur qui transcendent les différences linguistiques.
The measurement of pain, particularly chronic pain, is difficult due to its complexity. One of the most important roles of clinicians is to evaluate – precisely and objectively – the subjective experience of pain. It also has become important to use universally accepted scales for evaluating pain so that data can be compared internationally.

In English-speaking countries many pain rating scales have been employed. Of these, the McGill Pain Questionnaire (MPQ), which involves repeated psychometric assessment as a measure, is one of the most widely used and sensitive clinical tools for the verbal assessment of pain. However, differences in languages and cultural backgrounds have hindered its wide use and standardization in Japan. The MPQ has now been translated and reconstructed into many languages worldwide, including German, Italian, French and Finnish. In each translation alterations have been made where necessary to make the questionnaire most relevant to the language (1-10).

In the cross-cultural translation of psychometric tests, the ‘decentering’ concept of the back-translation technique is important (11-13). Decentering, a translation concept first outlined by Werner and Campbell (13), refers to a process by which one set of materials is not translated with as little change as possible into another language, but changed to produce a smooth, natural-sounding version in the second language. Thus, using the back-translation technique, the language is changed more freely so that the final image depicted in the translated version becomes as close as possible to that of the original. The major advantage of back-translation is that it gives researchers some control over the instrument development stage because they can examine original and back-translated versions and make inferences about the quality of the translation. Because of such advantages, we used the back-translation technique in this study.

Some pain rating scales, including the MPQ, originally constructed in Europe and North America have been translated into Japanese and applied generally. However, to evaluate subjective experience such as pain, it should be possible to make allowances for differences in the structure of answers owing to differences in cultural background and language. We thought it necessary to evaluate the validity and reliability of a Japanese version of the MPQ because translation may cause problems pertaining to the psychometric properties of a rating scale.

Although Satow and colleagues (14) made a full translation of the MPQ using the back-translation technique, this full version was never used. They reduced the number of pain descriptors and made their version of the Japanese MPQ (JMPQ-S); this version was not similar to the original format, eg, it was shortened. Furthermore, although they evaluated its reliability and validity, the JMPQ-S was performed only on patients with a limited number of diseases (14). Satow et al pointed out that when the JMPQ-S is used on controls without pain, the pain intensity corresponding to each of the pain descriptors has a tendency to be estimated as the midpoint in the visual analogue scale (VAS), and argued that therefore the JMPQ-S is an unreliable pain rating scale. The JMPQ-S is rarely used in Japan. However, we regard the Japanese translation of each pain descriptor in the JMPQ-S as appropriate, smooth and natural-sounding. Thus, we adopted all the expressions formulated by Satow et al.

Studies thus far on the JMPQ-S have been performed only on patients with limited diseases, using reconstructed versions different from the original format (14,15). We used a Japanese version of the MPQ (JMPQ) that followed a format that parallels the original MPQ in order to enable more direct comparison of study findings internationally, so that the psychometric properties of the JMPQ may reach cross-cultural equivalence. Accordingly, we restored the JMPQ format to its original form and used it on patients suffering from a far wider range of disease in an attempt to increase the external validity of the study.

The aim of this study was to confirm the psychometric properties, especially reliability and validity, of a JMPQ that follows a format parallel to the original MPQ. By doing so, we attempted to confirm the cross-cultural usefulness of the MPQ as an assessment tool for chronic pain patients.

PATIENTS AND METHODS

Patients were 152 consecutive patients with chronic pain (69 male; 83 female; mean age 55.8±14.2 years) who consulted the pain clinic affiliated with Gunma University School of Medicine, Maebashi, Japan, from January to May 1995. Average duration of pain was 15.2 months. In order to minimize the effects of treatment, patients were selected at their initial visit. Regions of pain described by patients were lower back (n=67), neck (n=42), face and/or head (n=15), shoulder (n=13) and other (n=15).

The JMPQ and other pain rating scales, such as the VAS, the verbal rating scale (VRS) and the numerical rating scale (NRS), were administered in a private interview involving patients awaiting treatment. The JMPQ used in this study, which followed a format similar to the original MPQ, comprised 78 pain descriptors in 20 subclasses and a five-point intensity scale (the present pain intensity [PRI] scale) (15). The descriptors described sensory (subclasses 1 to 10), affective (subclasses 11 to 15), evaluative (subclass 16) and mixed sensory, affective-evaluative (subclasses 17 to 20) aspects of pain. Previous studies employing factor analysis have revealed that these descriptors can be classified into one of three main subscales (sensory, affective and evaluative descriptors) or a miscellaneous fourth subscale (16,17). The pain rating index (PRI) is the sum of the rank values selected by the patient in each subclass. Pain descriptors used in this study were adopted as translated by Satow et al. The score of each subscale is briefly described as sensory score of the PRI (PRI-S), affective score of the PRI (PRI-A), evaluative score of the PRI (PRI-E) and miscellaneous score of the PRI (PRI-M). Total score of the PRI (PRI-T) was determined as the sum total of each subscale score.

To test the reliability of the JMPQ, internal consistency was evaluated for each subscale using the alpha coefficient. Pearson product-moment correlations were used to evaluate the stability of the JMPQ subscale and subclass scores over a two-week period in a subset of patients (n=40) (18,19). Patients for the stability study were selected from patients whose pain intensity was almost unchanged for the two weeks based on the scores of the PRI, VAS, VRS and NRS. Because data in this study were obtained by private interview, inter-rater reliability was assessed in a subset of patients (n=40) (19) who were selected from patients who did not participate in the stability study. Two board-certified clinical psychologists were used as raters for the study of inter-rater reliability.

To test the validity of the JMPQ, concurrent validity was assessed for each subscale as follows (19). First, the JMPQ subscale...
and total scores for all patients (n=152) were correlated with scores on three well known and established pain rating scales including the 100 mm VAS, the five-point VRS and the 0 to 100 scale NRS. Second, using the correlation matrix derived from the above measures and the JMPQ scales, principal components analysis (principal axis factoring followed by varimax rotation) was used to simplify the task of assessing the inter-relationships of these scales (20). Factors with eigenvalues greater than 1.0 were extracted. Items with factor loading scores greater than 0.70 were used to identify the factors derived. Thus, construct validity was assessed by means of factorial and concurrent validity.

It is well known that patients with chronic pain are generally more anxious than controls; thus it was anticipated that anxiety would influence the JMPQ scores (21). The state-trait anxiety inventory (STAI), which measures anxiety stemming from traits and particular psychological state at the time, was administered to the out-patients as a psychometric test to assess psychological change and/or anxiety induced by the pretreatment state in a relatively short period of time. The JMPQ subscale scores and the state and trait scores from the STAI were correlated to assess the relationship of concurrent pain and anxiety. Data obtained in this study were analyzed using the program package STATISTICA (StatSoft Inc, Oklahoma). P≤0.05 was considered statistically significant.

RESULTS
Average ± SD JMPQ scores for all patients were PRI-S 13.6±7.4, PRI-A 1.9±2.3, PRI-E 2.0±2.0, PRI-M 3.4±3.2 and PRI-T 20.9±12.4. The average ± SD JMPQ PPI score was 2.4±1.2, and was 47.7±23.3 for the VAS, 48.2±22.8 for the VRS and 47.4±24.0 for the NRS. Average ± SD scores on the simultaneously administered STAI were 41.5±8.4 and 40.6±6.7 for each state and trait form, respectively (n=152).

Regarding the reliability of the JMPQ, the internal consistency of all subscale scores and the total score on the JMPQ, ranging from 0.52 to 0.81, are listed in Table 1. The PRI-T appeared to be satisfactory. The alpha coefficient of the PRI-E could not be calculated because it consisted of only one subclass. As expected, the intrarater reliability was fairly high: 0.93 (n=40). The stability (test-retest reliability) coefficients of the JMPQ subscales were in the 0.80 to 0.91 range (Table 1). Table 2 shows the stability (test-retest reliability) coefficients among the 20 JMPQ subclasses, ranging from 0.61 to 0.95. These values indicated a significant correlation and stability over time for all subclasses and subscales.

Concurrent validity of the JMPQ is shown in Table 3. Significant intercorrelations among the JMPQ scores and other pain rating scale scores (VAS, VRS and NRS) indicate the validity of the JMPQ as a tool for assessing pain.

A combination of the Kaiser and Scree criteria indicated that a two-factor solution was most appropriate for intercorrelating the six JMPQ scales with the three previously validated scales (VAS, VRS and NRS). This solution accounted for 78% of the common variance.

The factor structure resulting from the varimax rotation is listed in Table 4. Factor 1 appeared to be related to pain severity. The PPI of the JMPQ and the three other pain rating scales that measured pain severity had major correlations with this factor. Factor 2 appeared to represent verbal expression of pain. Major loadings on this factor included the JMPQ subscales that consisted of sensory, affective and miscellaneous aspects of pain expressions. The PRI-E scale also was moderately correlated with this factor.

In summary, these results suggest that the JMPQ scales were independent of the other pain rating scales (VAS, VRS and NRS) that represent completely different aspects of pain. Mean values on the STAI indicated moderate anxiety concerning both state and trait forms. As can be seen in Table 3, a significant but relatively weak association was found between the JMPQ subscales and each of the STAI.

DISCUSSION
There are two types of methodology for introducing the MPQ into countries where English is not spoken. Some countries have used a translation-based method while others have used a reconstruction-
equivalent to the specific intensity and pain attributes of each scriptors. It is unlikely that any translation will give a satisfactory difficulty being that pain descriptors have semantic connotations specifically.

JMPQ adapted from the original by Satow et al that involved translation.

Concurrent validity of the JMPQ (correlation coefficients among the JMPQ subscales and other scales)

<table>
<thead>
<tr>
<th>Variable</th>
<th>PRI-S</th>
<th>PRI-A</th>
<th>PRI-E</th>
<th>PRI-M</th>
<th>PRI-T</th>
<th>PPI</th>
<th>VAS</th>
<th>VRS</th>
<th>NRS</th>
<th>STA-1</th>
<th>STA-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRI-S</td>
<td>1.00</td>
<td>0.58*</td>
<td>0.51*</td>
<td>0.60*</td>
<td>0.94*</td>
<td>0.52*</td>
<td>0.56*</td>
<td>0.55*</td>
<td>0.55*</td>
<td>0.17*</td>
<td>0.11*</td>
</tr>
<tr>
<td>PRI-A</td>
<td>0.58*</td>
<td>1.00</td>
<td>0.43*</td>
<td>0.55*</td>
<td>0.74*</td>
<td>0.49*</td>
<td>0.45*</td>
<td>0.47*</td>
<td>0.46*</td>
<td>0.28*</td>
<td>0.19*</td>
</tr>
<tr>
<td>PRI-E</td>
<td>0.51*</td>
<td>0.43*</td>
<td>1.00</td>
<td>0.46*</td>
<td>0.66*</td>
<td>0.69*</td>
<td>0.54*</td>
<td>0.54*</td>
<td>0.53*</td>
<td>0.24*</td>
<td>0.19*</td>
</tr>
<tr>
<td>PRI-M</td>
<td>0.60*</td>
<td>0.55*</td>
<td>0.46*</td>
<td>1.00</td>
<td>0.79*</td>
<td>0.53*</td>
<td>0.47*</td>
<td>0.48*</td>
<td>0.47*</td>
<td>0.20*</td>
<td>0.12*</td>
</tr>
<tr>
<td>PRI-T</td>
<td>0.94*</td>
<td>0.74*</td>
<td>0.66*</td>
<td>0.79*</td>
<td>1.00</td>
<td>0.64*</td>
<td>0.63*</td>
<td>0.62*</td>
<td>0.62*</td>
<td>0.24*</td>
<td>0.16</td>
</tr>
<tr>
<td>PPI</td>
<td>0.52*</td>
<td>0.49*</td>
<td>0.68*</td>
<td>0.52*</td>
<td>0.64*</td>
<td>1.00</td>
<td>0.72*</td>
<td>0.73*</td>
<td>0.72*</td>
<td>0.21*</td>
<td>0.14</td>
</tr>
<tr>
<td>VAS</td>
<td>0.56*</td>
<td>0.45*</td>
<td>0.54*</td>
<td>0.47*</td>
<td>0.63*</td>
<td>0.72*</td>
<td>1.00</td>
<td>0.98*</td>
<td>0.99*</td>
<td>0.16</td>
<td>0.10</td>
</tr>
<tr>
<td>VRS</td>
<td>0.55*</td>
<td>0.47*</td>
<td>0.54*</td>
<td>0.48*</td>
<td>0.62*</td>
<td>0.73*</td>
<td>0.98*</td>
<td>1.00</td>
<td>0.99*</td>
<td>0.17*</td>
<td>0.10</td>
</tr>
<tr>
<td>NRS</td>
<td>0.55*</td>
<td>0.46*</td>
<td>0.53*</td>
<td>0.47*</td>
<td>0.62*</td>
<td>0.27*</td>
<td>0.99*</td>
<td>0.99*</td>
<td>1.00</td>
<td>0.15</td>
<td>0.07</td>
</tr>
<tr>
<td>STA-1</td>
<td>0.17*</td>
<td>0.28*</td>
<td>0.24*</td>
<td>0.20*</td>
<td>0.24*</td>
<td>0.21*</td>
<td>0.16</td>
<td>0.17*</td>
<td>0.15</td>
<td>1.00</td>
<td>0.43*</td>
</tr>
<tr>
<td>STA-2</td>
<td>0.11</td>
<td>0.19*</td>
<td>0.19*</td>
<td>0.12</td>
<td>0.16</td>
<td>0.14</td>
<td>0.10</td>
<td>0.10</td>
<td>0.07</td>
<td>0.43*</td>
<td>1.00</td>
</tr>
</tbody>
</table>

*P<0.05 (n=152), JMPQ: Japanese version of the McGill Pain Questionnaire; NRS: Numerical rating scale; PPI: Present pain intensity; PRI-A: Affective score of the pain rating index (PRI); PRI-E: Evaluative score of the PRI; PRI-M: Miscellaneous score of the PRI; PRI-S: Sensory score of the PRI; PRI-T: Total score of the PRI; STA-1: State trait anxiety inventory – State form; STA-2: State trait anxiety inventory – Trait form; VAS: Visual analogue scale; VRS: Verbal rating scale

Principal component analysis with varimax rotation of the JMPQ subscales and other pain rating scales

<table>
<thead>
<tr>
<th>Variable</th>
<th>Factor 1</th>
<th>Factor 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRI-S</td>
<td>0.32847</td>
<td>0.82135*</td>
</tr>
<tr>
<td>PRI-A</td>
<td>0.197865</td>
<td>0.81375*</td>
</tr>
<tr>
<td>PRI-E</td>
<td>0.586325</td>
<td>0.413947</td>
</tr>
<tr>
<td>PRI-M</td>
<td>0.354628</td>
<td>0.783629*</td>
</tr>
<tr>
<td>PRI-T</td>
<td>0.413270</td>
<td>0.883646*</td>
</tr>
<tr>
<td>PPI</td>
<td>0.782957*</td>
<td>0.364421</td>
</tr>
<tr>
<td>VAS</td>
<td>0.889765*</td>
<td>0.280037</td>
</tr>
<tr>
<td>VRS</td>
<td>0.902835*</td>
<td>0.275423</td>
</tr>
<tr>
<td>NRS</td>
<td>0.911372*</td>
<td>0.273087</td>
</tr>
</tbody>
</table>

Factor loadings >0.70 are considered significant (n=152). JMPQ: Japanese version of the McGill Pain Questionnaire; NRS: Numerical rating scale; PPI: Present pain intensity; PRI-A: Affective score of the pain rating index (PRI); PRI-E: Evaluative score of the PRI; PRI-M: Miscellaneous score of the PRI; PRI-S: Sensory score of the PRI; PRI-T: Total score of the PRI; VAS: Visual analogue scale; VRS: Verbal rating scale

In this study we used a translation-based Japanese version of the MPQ adapted from the original by Satow et al that involved translation from a bilingual English linguist and that followed a format parallel to the original MPQ (14,15). Back-translation procedures were carried out to ensure adequacy of the Japanese version (11-13). By applying our JMPQ to patients with chronic pain, we aimed to establish a pain rating scale of proven reliability and validity. However, a relatively high correlation was recognized between the subscales shown in Table 3. The original MPQ employed four mutually independent subscales to assess different qualitative aspects of pain. The subscales of the JMPQ did not display adequate discriminant validity in contrast to the original MPQ (15). The four PRI components had relatively high intercorrelations in the JMPQ, results similar to those in the English version (22). Therefore, the JMPQ had problems similar to those in the English version in assessing the different qualitative aspects of pain. These problems are likely a structural flaw of the original MPQ as opposed to structural changes caused by translation.

In addition, Satow et al stated that it is difficult to rank the intensity of Japanese pain descriptors using the VAS because almost all the descriptors are assessed as being of moderate intensity. Hence, it has been suggested that the JMPQ-S is not clinically useful as a pain rating scale in Japan.

Despite these flaws, we maintain that the JMPQ is useful clinically as a pain rating scale because it allows the assessment of qualitative aspects of pain, transcending differences in the linguistic characteristics of scales such as the VAS, VRS and NRS. The JMPQ allows this assessment because of its reliability and validity, and because of the independence of the JMPQ subscales (PRI-S, PRI-A, PRI-E, PRI-M and PRI-T) from other pain rating scales (VAS, VRS and NRS) as confirmed by principal component analysis. Moreover, the form of the original MPQ – adopting subclasses to replace subjective experiences of pain intensity with objective forms – is fully maintained in this translation-based JMPQ. However, because the subscales are not independent from each other, use of separate subscale scores is inappropriate. They are measuring the same construct. Therefore, we view the total score (PRI-T) as a good general measure of pain severity on behalf of the subscales (22). To assess the qualitative aspects of pain more clearly, we propose that reli-
sions in the current JMPQ may make it more adaptable to the Japanese language and, hence, of future clinical value.

Scores on the simultaneously administered STAI weakly correlated with the JMPQ subscale scores, revealing only slight anxiety. This was in contrast to predictions that patients with chronic pain are highly anxious, a result similar to that reported by Bourreau and colleagues (21).

CONCLUSIONS

This study examined the reliability and validity of a JMPQ that followed a format parallel to the original MPQ. Results showed acceptable levels of reliability and validity for the JMPQ and independence of the JMPQ subscales from other pain rating scales as confirmed by principal component analysis.

Within the limitations of this study, these findings suggest that the JMPQ possesses sufficient merits as a pain rating scale from the standpoint of its reliability and validity. It is hoped that this study stimulates further research interest in this area.

REFERENCES

13. Werner O, Campbell D. Translating, working through interpreters, and the JMPQ possesses sufficient merits as a pain rating scale from the standpoint of its reliability and validity. In other words, the form of the original MPQ is fully maintained in this translation-based JMPQ. Furthermore, we feel that the JMPQ has an advantage over other pain rating scales regarding qualitative aspects because the JMPQ transcends differences in linguistic characteristics. Therefore, we suggest that the JMPQ based on a format similar to the original MPQ may sufficiently aid the clinical assessment of pain in Japan, similar to the MPQ in English-speaking countries. In addition, concurrent administration of the JMPQ and STAI suggest that patients with chronic pain are less anxious than initially expected.

PAIN RES MANAGE VOL 1 NO 4 WINTER 1996 237