



# VALIDITY AND RELIABILITY OF PAKISTANI URDU VERSION SHORT FORM McGill PAIN QUESTIONNAIRE IN POST CARDIAC SURGERY PATIENTS

Dr Mubashra Malik<sup>1</sup>, Dr. Muhammad Tahir Hafeez<sup>1</sup>, Dr Fahad Tanveer<sup>2</sup>, Habib Ur Rehman<sup>3</sup>, Fareeha Chaudhary<sup>4</sup>, Arslan Saeed<sup>5</sup>, Sardar Abdul Hanan Saeed<sup>6</sup>, Dr Rafia Imtiaz<sup>\*7</sup>

<sup>1</sup>Senior Lecturer Iqra University Chak Shahzad Campus Islamabad

<sup>2</sup>HOD Rehabilitation Sciences, Green International University Lahore

<sup>3</sup>Physiotherapist, Masood Hospital Lahore

<sup>4</sup>Clinical Physiotherapist, Capital Diagnostic Center Islamabad

<sup>5</sup>Johar Institute of Professional Studies Lahore

<sup>6</sup>Shifa International hospital Islamabad

<sup>7</sup>Government College University Faisalabad



**\*Corresponding Author:** Dr Rafia Imtiaz, Government College University Faisalabad

Email: [rafia\\_imtiaz@yahoo.com](mailto:rafia_imtiaz@yahoo.com)



## CONFLICT OF INTEREST

None declared.



## GRANT SUPPORT

None declared.



## FINANCIAL SUPPORT

None Declared.



## ABSTRACT

**Background:** The Short Form McGill Pain Questionnaire (SF-MPQ) is a multidimensional pain measure that captures pain intensity together with sensory and affective qualities. Linguistically appropriate validation is important when a patient-reported outcome measure is used in a population whose primary clinical language differs from the source version, particularly after cardiac surgery when pain changes rapidly during recovery.

**Objective:** To evaluate the validity and reliability of the Pakistani Urdu version of the SF-MPQ in patients after cardiac surgery through median sternotomy.

**Methods:** This observational psychometric validation study included 265 clinically stable post-cardiac-surgery patients recruited by purposive sampling from the Armed Forces Institute of Cardiology between June 2021 and June 2022. Ten experts reviewed the Urdu version for face and content validity. Concurrent validity was examined by correlating Urdu and English SF-MPQ scores on postoperative day (POD) 1. Construct validity was examined against the Numeric Pain Rating Scale (NPRS) on POD 4, 6, and 15 and against the 12-item Short Form Health Survey (SF-12) on POD 6 and 15. Internal consistency and test-retest reliability were assessed using Cronbach alpha and intraclass correlation coefficients (ICC). Exploratory factor analysis (EFA) was used to examine factor structure.

**Results:** The mean age was 58 +/- 9.4 years and 206/265 participants (77.7%) were male. Expert agreement for face validity was 100%, and item- and scale-level content validity indices were 1.0. Concurrent validity between the Urdu and English versions was very strong ( $r = 0.980$ ,  $p < 0.001$ ). Correlations with NPRS were strong across POD 4, 6, and 15 ( $r = 0.796$ - $0.850$ ,  $0.809$ - $0.846$ , and  $0.811$ , respectively; all  $p < 0.001$ ). Correlation with the SF-12 was weakly negative on POD 6 ( $r = -0.315$ ,  $p < 0.001$ ) and negligible on POD 15 ( $r = 0.021$ ,  $p = 0.734$ ). Overall Cronbach alpha was 0.913 on POD 4 and 0.861 on POD 6; corresponding ICCs were 0.841 and 0.755. Domain ICCs were heterogeneous: sensory 0.098, affective 0.907, intensity 0.957, and present pain intensity 0.915. EFA retained five components explaining approximately 65% of the variance; the Kaiser-Meyer-Olkin value was 0.538.

**Conclusion:** The Pakistani Urdu SF-MPQ demonstrated strong concurrent validity with the English version, strong convergence with NPRS, and good overall reliability in this post-cardiac-surgery sample. However, the weak association with SF-12, poor sensory-domain test-retest reliability, and marginal factorability indicate that the psychometric performance was not uniformly strong across all domains. The instrument may support Urdu-language pain assessment in similar clinically stable patients, with cautious interpretation of subscale scores and further multicentre validation.

## Keywords

Cardiac surgery; McGill Pain Questionnaire; Numeric Pain Rating Scale; Pain; Reliability; Short Form Health Survey; Urdu; Validity.

## Acknowledgement

The authors acknowledge the patients who participated in the study and the clinical teams at the Armed Forces Institute of Cardiology who facilitated the data-collection process.

## Malik M et al. Short Title

Validity of Pakistani Urdu SF-MPQ After Cardiac Surgery





## INTRODUCTION

Pain is a multidimensional experience that includes intensity as well as sensory and affective qualities. The McGill Pain Questionnaire (MPQ), developed by Melzack, introduced a structured vocabulary for characterizing these dimensions and became an influential instrument for clinical pain assessment (1,2). To reduce administration time while retaining the core descriptive domains, the Short Form McGill Pain Questionnaire (SF-MPQ) was developed in 1987 (3). The original SF-MPQ contains 15 pain descriptors, including 11 sensory and four affective descriptors, each rated from 0 (none) to 3 (severe), together with a Present Pain Intensity (PPI) rating and a visual analogue measure of pain intensity (3). Pain after cardiac surgery is clinically important because sternotomy, chest wall retraction, tissue dissection, drains, coughing, mobilization, and other perioperative factors can produce substantial acute pain and may contribute to persistent postsurgical pain and impaired recovery (4). Accurate assessment is therefore essential for communication, treatment planning, respiratory rehabilitation, mobilization, and evaluation of analgesic response. Contemporary standards for patient-reported outcome measures also emphasize transparent evaluation of validity, reliability, and interpretability in the population and language in which the instrument will be used (5).

The SF-MPQ and related versions have been evaluated in multiple languages and clinical populations. Cross-cultural studies have reported psychometric evidence for Arabic (6), Korean (7), Norwegian (8), Iranian (9,23), Sinhala (10), E-san (11), Turkish (12), Chinese (14), and other versions. The revised SF-MPQ-2, which is distinct from the original SF-MPQ, has also undergone extensive validation (17). More recently, an Urdu SF-MPQ-2 was validated in people with low back pain in Pakistan (13), and a Cantonese Chinese SF-MPQ-2 demonstrated high internal consistency and good test-retest reliability in chronic pain populations (24). These studies reinforce the need to evaluate each language version in the clinical population for which it is intended rather than assuming that psychometric properties transfer unchanged across languages or diagnoses. Post-sternotomy pain is dynamic during the first postoperative days, and pain descriptors may differ from those observed in chronic musculoskeletal or neuropathic pain. Studies after cardiac surgery have shown that SF-MPQ-type instruments can provide clinically meaningful information on both intensity and pain quality (22,25). Recent evidence also links greater post-sternotomy pain with poorer subsequent quality of life (26). Despite the availability of an authorized Pakistani Urdu SF-MPQ, evidence regarding its psychometric performance specifically in post-cardiac-surgery patients has been limited. The present study therefore aimed to evaluate the validity and reliability of the Pakistani Urdu version of the SF-MPQ in patients recovering from cardiac surgery through sternotomy.

## MATERIALS AND METHODS

### Study design, setting and duration

This was an observational psychometric validation study conducted at the Armed Forces Institute of Cardiology. Participants were recruited using non-probability purposive sampling. Data collection was conducted from June 2021 to June 2022. The study was designed to examine face and content validity, concurrent and construct validity, internal consistency, test-retest reliability, and exploratory factor structure of the Pakistani Urdu SF-MPQ. Psychometric outcomes are reported using terminology consistent with established questionnaire-validation methods and contemporary measurement-property reporting recommendations (5,15,16).

### Participants and eligibility criteria

Eligible participants were clinically stable patients after cardiac surgery performed through a sternotomy who were able to read and understand both Urdu and English, were physically independent, and had no associated postoperative complications that prevented participation. Patients were excluded if they had undergone eye surgery within the preceding 6 months, had significant cognitive impairment or neurologic injury, psychomotor dysfunction, or physical disability. Written informed consent was obtained before study procedures.

### Sample size

The study protocol used an item-based psychometric sampling approach for exploratory factor analysis. Approximately 12 participants were planned for each of 22 evaluated questionnaire elements, yielding a minimum of 264 observations; 265 patients were ultimately enrolled. This sample also provided a large dataset for correlation, internal-consistency, and test-retest analyses (16).

### Study instruments

Pakistani Urdu Short Form McGill Pain Questionnaire (SF-MPQ). The SF-MPQ includes 15 pain descriptors: 11 sensory descriptors (throbbing, shooting, stabbing, sharp, cramping, gnawing, hot-burning, aching, heavy, tender, and splitting) and four affective descriptors (tiring-exhausting, sickening, fearful, and punishing-cruel). Each descriptor is rated from 0 (none) to 3 (severe), producing a descriptor-based pain rating score from 0 to 45. The instrument also includes the PPI and a visual analogue assessment of pain intensity (3). Permission to use the copyrighted Pakistani Urdu version was obtained from Mapi Research Trust. The original SF-MPQ and translated versions have demonstrated acceptable measurement properties across several pain populations (6,11,12,23).

Numeric Pain Rating Scale (NPRS). The NPRS is a unidimensional 0-10 rating scale in which 0 represents no pain and 10 represents the worst possible pain. It is widely used to quantify pain intensity and has demonstrated strong agreement with other unidimensional pain-intensity measures and high repeatability in clinical populations (18,19). In the present study, NPRS scores were used to examine convergent construct validity of the Urdu SF-MPQ.



Short Form Health Survey (SF-12). The SF-12 is a 12-item patient-reported measure of health-related quality of life derived from the SF-36 and is commonly summarized into physical and mental health components. Psychometric studies have demonstrated acceptable reliability and construct validity across clinical and general populations (20,21). In the present study, SF-12 scores were used as a broader health-status comparator for construct validity.

### Face and content validity

A panel of 10 experts from cardiology, physiotherapy, and anaesthesiology reviewed the Pakistani Urdu SF-MPQ. Face validity was assessed using a dichotomous yes/no judgment regarding the acceptability and comprehensibility of each evaluated element. Content validity was rated for relevance, clarity, simplicity, and ambiguity on a 4-point scale. For content-validity calculations, ratings of 3 or 4 were treated as acceptable. Item-level content validity index (I-CVI) was calculated as the proportion of experts providing an acceptable rating. Scale-level CVI was summarized both as the average of I-CVIs (S-CVI/Ave) and as universal agreement (S-CVI/UA). One expert proposed a wording change for item 13; because the translated version was supplied under copyright permission, no wording was altered without approval from the instrument custodian.

### Data-collection procedure

On POD 1, once patients were clinically stable and met the eligibility criteria, the Pakistani Urdu SF-MPQ was administered in the morning. The original English SF-MPQ was administered 6-7 hours later to examine concurrent validity. On POD 4, the Urdu SF-MPQ was administered in the morning and repeated 7-8 hours later; the later assessment also included NPRS. On POD 6, the Urdu SF-MPQ was again administered in the morning and repeated after 7-8 hours together with NPRS and SF-12. On POD 15, during follow-up, the Urdu SF-MPQ, NPRS, and SF-12 were administered once. Same-day repeated Urdu SF-MPQ measurements on POD 4 and POD 6 were used for test-retest reliability, whereas correlations with NPRS and SF-12 were used for construct validity.

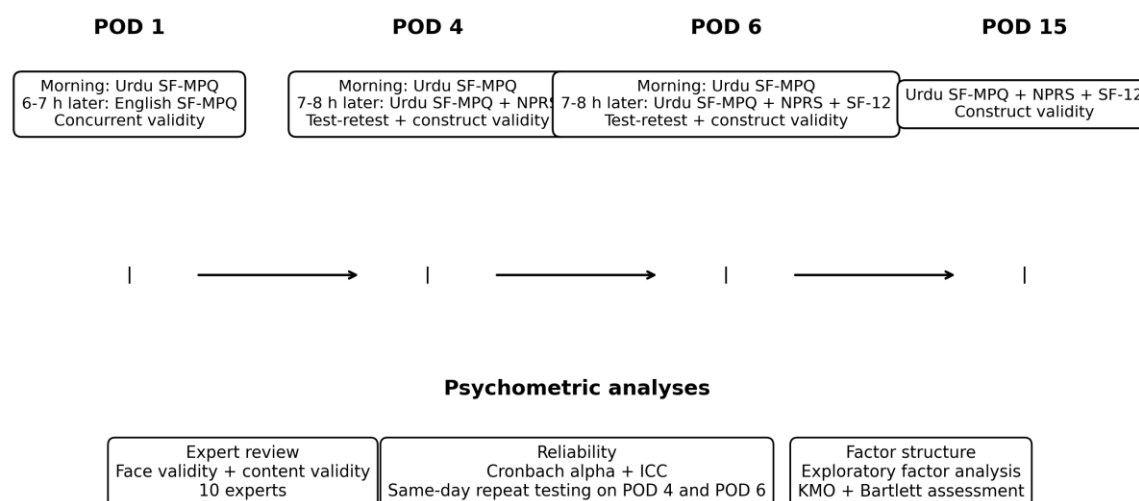


Figure 1. Study schedule and psychometric assessment plan.

### Statistical analysis

Data were analysed using SPSS version 25.0. Continuous demographic variables are presented as mean  $\pm$  standard deviation, and categorical variables as frequency and percentage. Face and content validity were summarized as expert agreement and CVI values. Pearson correlation coefficients ( $r$ ) were used for the reported concurrent and construct validity analyses. Correlations were interpreted by magnitude and direction, with two-sided  $p < 0.05$  considered statistically significant. Internal consistency was assessed using Cronbach alpha, and test-retest reliability was assessed using ICC with 95% confidence intervals. EFA was performed to examine latent factor structure; the Kaiser-Meyer-Olkin (KMO) statistic and Bartlett test were used to assess the suitability of the correlation matrix for factor analysis. The number of retained components and cumulative explained variance were reported from the available EFA output.



### Ethical considerations

Written informed consent was obtained from participants before data collection. Permission was also obtained from the relevant clinical staff before each assessment session, and patients were assessed only when clinically stable and able to respond. Permission to use the copyrighted Pakistani Urdu SF-MPQ was obtained from Mapi Research Trust.

## RESULTS

### Participant characteristics

A total of 265 participants were included. The mean age was 58 +/- 9.4 years and the mean body mass index (BMI) was 22 +/- 3.1 kg/m<sup>2</sup>. There were 206 males (77.7%) and 59 females (22.3%). Most participants underwent coronary artery bypass grafting (CABG; 233/265, 87.9%), while 32 (12.1%) underwent valvular replacement surgery.

**Table 1. Demographic and clinical characteristics (n = 265)**

| Variable                 | Category / summary   | Value       |
|--------------------------|----------------------|-------------|
| Age (years)              | Mean +/- SD          | 58 +/- 9.4  |
| BMI (kg/m <sup>2</sup> ) | Mean +/- SD          | 22 +/- 3.1  |
| Gender                   | Male                 | 206 (77.7%) |
| Gender                   | Female               | 59 (22.3%)  |
| Surgery                  | CABG                 | 233 (87.9%) |
| Surgery                  | Valvular replacement | 32 (12.1%)  |

### Face and content validity

The expert panel comprised four cardiologists, four physiotherapists, and two anaesthesiologists. All 10 experts agreed with the face-validity items. For content validity, all experts rated the evaluated elements as relevant, simple, and acceptable with respect to ambiguity. For clarity, 21 of 22 evaluated elements received the highest rating from all experts; item 13 received a rating of 3 from one expert and 4 from the remaining nine. Because ratings of 3 and 4 were treated as acceptable for CVI calculations, the I-CVI for all evaluated elements was 1.0, with S-CVI/Ave and S-CVI/UA both equal to 1.0.

**Table 2. Face and content validity of the Pakistani Urdu SF-MPQ**

| Validity parameter                                  | Result                                                                                      |
|-----------------------------------------------------|---------------------------------------------------------------------------------------------|
| Expert panel (n = 10)                               | 4 cardiologists, 4 physiotherapists, 2 anaesthesiologists                                   |
| Face validity                                       | 100% agreement across the 22 evaluated elements                                             |
| Content validity - relevance, simplicity, ambiguity | All experts provided acceptable ratings across the 22 evaluated elements                    |
| Content validity - clarity                          | 21/22 elements rated 4/4 by all experts; item 13 rated 3/4 by 1 expert and 4/4 by 9 experts |
| Item-level CVI (I-CVI)                              | 1.0 for all evaluated elements                                                              |
| Scale-level CVI (S-CVI/Ave)                         | 1.0                                                                                         |
| Scale-level CVI (S-CVI/UA)                          | 1.0                                                                                         |

### Concurrent and construct validity

Concurrent validity between the Pakistani Urdu and English SF-MPQ scores on POD 1 was very strong ( $r = 0.980$ ,  $p < 0.001$ ). Convergent construct validity with NPRS was strong at each postoperative assessment: POD 4 morning  $r = 0.796$  and evening  $r = 0.850$ ; POD 6 morning  $r = 0.809$  and evening  $r = 0.846$ ; and POD 15  $r = 0.811$  (all  $p < 0.001$ ). The relationship with SF-12 was weaker. On POD 6, SF-MPQ and SF-12 scores showed a weak inverse correlation ( $r = -0.315$ ,  $p < 0.001$ ), whereas the POD 15 correlation was negligible and not statistically significant ( $r = 0.021$ ,  $p = 0.734$ ).

**Table 3. Concurrent and construct validity**

| Comparison                            | N   | r     | p value | Interpretation       |
|---------------------------------------|-----|-------|---------|----------------------|
| SF-MPQ Urdu vs SF-MPQ English (POD 1) | 265 | 0.980 | <0.001  | Very strong positive |
| SF-MPQ Urdu vs NPRS - POD 4 AM        | 265 | 0.796 | <0.001  | Strong positive      |





| Comparison                     | N   | r      | p value | Interpretation  |
|--------------------------------|-----|--------|---------|-----------------|
| SF-MPQ Urdu vs NPRS - POD 4 PM | 265 | 0.850  | <0.001  | Strong positive |
| SF-MPQ Urdu vs NPRS - POD 6 AM | 265 | 0.809  | <0.001  | Strong positive |
| SF-MPQ Urdu vs NPRS - POD 6 PM | 265 | 0.846  | <0.001  | Strong positive |
| SF-MPQ Urdu vs NPRS - POD 15   | 265 | 0.811  | <0.001  | Strong positive |
| SF-MPQ Urdu vs SF-12 - POD 6   | 265 | -0.315 | <0.001  | Weak inverse    |
| SF-MPQ Urdu vs SF-12 - POD 15  | 265 | 0.021  | 0.734   | Negligible      |

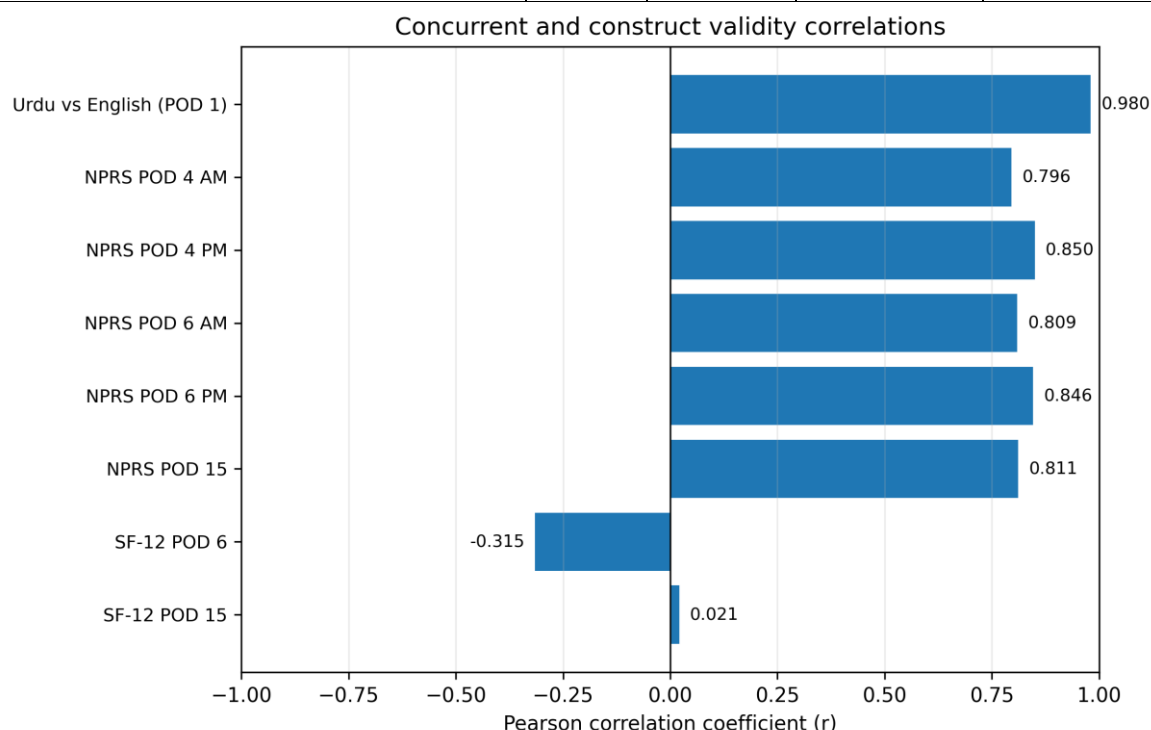


Figure 2. Correlation coefficients for concurrent and construct validity.

## Reliability

Overall internal consistency was high on POD 4 (Cronbach alpha = 0.913) and POD 6 (alpha = 0.861). The corresponding overall ICCs were 0.841 (95% CI 0.801-0.873) and 0.755 (95% CI 0.698-0.803), indicating good overall repeatability. Domain-level reliability was heterogeneous. The affective, intensity, and present pain intensity domains showed high ICCs (0.907, 0.957, and 0.915, respectively), whereas the sensory domain showed low internal consistency (alpha = 0.63) and poor test-retest reliability (ICC = 0.098, 95% CI 0.076-0.125).

Table 4. Reliability of the Pakistani Urdu SF-MPQ

| Domain / time point    | Cronbach alpha | ICC   | 95% CI      |
|------------------------|----------------|-------|-------------|
| Overall - POD 4        | 0.913          | 0.841 | 0.801-0.873 |
| Overall - POD 6        | 0.861          | 0.755 | 0.698-0.803 |
| Sensory                | 0.63           | 0.098 | 0.076-0.125 |
| Affective              | 0.95           | 0.907 | 0.883-0.926 |
| Intensity              | 0.97           | 0.957 | 0.945-0.966 |
| Present Pain Intensity | 0.95           | 0.915 | 0.893-0.933 |



### Pain descriptors

Among the most frequently reported descriptors were tender pain in 47 participants (17.7%), heavy pain in 39 (14.7%), aching pain in 35 (13.2%), sharp pain in 20 (7.5%), and stabbing pain in 19 (7.2%). These frequencies describe the most commonly endorsed pain qualities reported in the study rather than an exhaustive mutually exclusive distribution.

**Table 5. Most frequently reported SF-MPQ pain descriptors**

| Pain descriptor | n  | % of participants |
|-----------------|----|-------------------|
| Tender          | 47 | 17.7              |
| Heavy           | 39 | 14.7              |
| Aching          | 35 | 13.2              |
| Sharp           | 20 | 7.5               |
| Stabbing        | 19 | 7.2               |

### Exploratory factor analysis

The exploratory factor analysis retained five components. The cumulative extracted variance was approximately 65%. The KMO statistic was 0.538, indicating marginal sampling adequacy for factor analysis. The study output reported nine eigenvalues and a Bartlett test statistic of 103. Taken together, the EFA findings provide preliminary rather than definitive evidence regarding the latent structure of the Pakistani Urdu SF-MPQ in this postoperative population.

**Table 6. Summary of exploratory factor analysis**

| Parameter                     | Result            |
|-------------------------------|-------------------|
| Kaiser-Meyer-Olkin statistic  | 0.538             |
| Bartlett test statistic       | 103               |
| Reported eigenvalues          | 9                 |
| Retained components           | 5                 |
| Cumulative explained variance | Approximately 65% |

## DISCUSSION

This study evaluated the Pakistani Urdu SF-MPQ in 265 patients recovering from cardiac surgery through sternotomy. The main findings were: complete expert agreement for face and content validity; very strong correlation with the English SF-MPQ; strong correlations with NPRS across several postoperative time points; good overall internal consistency and test-retest reliability; heterogeneous reliability across subdomains, with particularly poor sensory-domain ICC; and a five-component EFA solution with marginal KMO sampling adequacy. These findings support several important measurement properties of the Urdu instrument, but they do not support an unqualified claim that every domain has excellent validity and reliability. The demographic profile was broadly comparable with other post-cardiac-surgery pain studies. Alharbi et al. evaluated 74 adults after median sternotomy using an Arabic SF-MPQ and reported a mean age of approximately 57 years, close to the mean age of 58 years in the present sample (22). Katijjahbe et al. similarly used a multidimensional McGill-based pain measure in adults after cardiac surgery and demonstrated that postoperative pain is characterized by both intensity and qualitative descriptors that change during recovery (25). In the present cohort, tender, heavy, aching, sharp, and stabbing were the most frequently reported descriptors. Differences in the ranking of descriptors across studies are plausible because postoperative pain is influenced by the surgical procedure, timing of assessment, analgesic exposure, language, and cultural interpretation of pain descriptors (22,25).

Expert review supported strong content relevance and comprehensibility, with I-CVI and scale-level CVI values of 1.0. This is consistent with previous cross-cultural adaptations of the SF-MPQ that have reported good acceptability and content validity (6,11,12,23). Nevertheless, expert agreement and patient comprehensibility represent related but distinct aspects of content validity. Contemporary measurement-property guidance emphasizes the importance of relevance, comprehensiveness, and comprehensibility from the perspective of the intended respondent population (5). Accordingly, future work should complement expert assessment with formal cognitive interviewing of Urdu-speaking cardiac-surgery patients. Concurrent validity with the English version was very strong ( $r = 0.980$ ). This suggests close correspondence between the Urdu and English administrations in bilingual participants. The magnitude is higher than some previous cross-language SF-MPQ studies, which may reflect the use of the same underlying instrument in the same participants. However, the two language versions were administered 6-7 hours apart and in a fixed sequence during the immediate postoperative period, when pain may change with analgesia, mobilization, chest-tube management, and recovery. The correlation therefore supports cross-language agreement but should be interpreted in the context of possible time and order effects.





The strong correlations with NPRS ( $r = 0.796-0.850$  on POD 4,  $r = 0.809-0.846$  on POD 6, and  $r = 0.811$  on POD 15) provide convergent evidence that higher SF-MPQ scores track higher pain intensity. Comparable convergent validity has been reported across Arabic, Iranian, Norwegian, Chinese, and more recent Urdu or Cantonese versions of McGill-derived instruments (6,8,9,13,14,23,24). In contrast, the relationship with SF-12 was weakly inverse on POD 6 and negligible on POD 15. This is not necessarily contradictory: SF-12 reflects broad physical and mental health status rather than pain alone, and the association between postoperative pain and global quality of life can vary according to recovery stage and other clinical factors. Recent cardiac-surgery evidence nevertheless indicates that greater early post-sternotomy pain can be associated with poorer subsequent quality of life (26). Reliability findings require a domain-specific interpretation. Overall Cronbach alpha values of 0.913 and 0.861 and overall ICCs of 0.841 and 0.755 indicate good reproducibility at the scale level. Affective, intensity, and present pain intensity domains also had high ICCs. In contrast, the sensory domain showed  $\alpha = 0.63$  and  $ICC = 0.098$ , indicating weak internal consistency and poor repeatability for that domain in this sample. This differs from several SF-MPQ-2 validation studies that reported higher domain ICCs and internal consistency (9,14,24), although direct comparison is limited because SF-MPQ-2 is a revised instrument with different content and scoring. The low sensory reliability may also reflect genuine short-term changes in acute post-sternotomy sensations during the 6-8-hour retest interval rather than measurement error alone.

The factor-analytic findings also warrant caution. Five components were retained and approximately 65% of variance was explained, but the KMO value of 0.538 indicates only marginal adequacy of the correlation matrix for factor analysis. In contrast, Iranian and Chinese SF-MPQ-2 studies reported stronger sampling-adequacy indices and factor solutions closer to the structure of the revised instrument (9,14). More recent SF-MPQ-2 work has similarly supported multidimensional structures using confirmatory approaches (13,24). The present five-component result should therefore be considered exploratory and population-specific until replicated with a larger multicentre sample and a fully specified factor-loading analysis. Clinically, the Pakistani Urdu SF-MPQ may be useful when clinicians need a brief multidimensional description of pain in Urdu-speaking post-cardiac-surgery patients. Its strong association with NPRS suggests that the overall score reflects pain intensity while also preserving qualitative descriptors that a unidimensional 0-10 scale cannot provide. The results are most supportive for overall scoring and the affective/intensity-related domains; the sensory subscale should be interpreted more cautiously until its repeatability is confirmed.

## LIMITATIONS

The study was conducted at a single cardiac centre using purposive sampling, which limits generalizability to other hospitals, regions, and clinical pathways. Eligibility required participants to understand both Urdu and English, so the findings may not fully represent Urdu-speaking patients with limited English literacy. The study focused on clinically stable and physically independent post-sternotomy patients; psychometric performance may differ in patients with complicated postoperative courses or greater functional dependency. The test-retest interval was 6-8 hours during an acute postoperative period in which pain can change because of recovery, analgesia, mobilization, and routine clinical care. This may violate the assumption that the underlying construct remains stable between repeat measurements and may partly explain the low sensory-domain ICC. The POD 1 cross-language comparison also used a fixed sequence rather than randomized or counterbalanced administration, creating potential order and time effects.

The SF-12 used for the health-related quality-of-life comparison was not an Urdu-translated version. In addition, the weak SF-12 correlations and marginal KMO value indicate that construct and factor-structure evidence was less robust than the concurrent validity and overall reliability findings. The five-component EFA solution should therefore be confirmed in an independent and more diverse sample.

## RECOMMENDATIONS

Future studies should evaluate the Pakistani Urdu SF-MPQ in multicentre samples, include Urdu-dominant participants across different educational levels, and use a retest interval in which clinical stability can be demonstrated. Formal cognitive interviews with target-language patients would strengthen evidence for comprehensibility. Confirmatory factor analysis should be used to test the factor structure suggested by the present EFA, and domain-level reliability - particularly the sensory domain - should be re-examined in independent postoperative and non-postoperative pain populations.

## CONCLUSION

The Pakistani Urdu version of the SF-MPQ showed very strong concurrent validity with the English SF-MPQ, strong convergent validity with NPRS, and good overall internal consistency and test-retest reliability in clinically stable patients after cardiac surgery. Evidence was less consistent for the SF-12 relationship, sensory-domain reliability, and factor structure. The instrument can support multidimensional Urdu-language pain assessment in similar post-cardiac-surgery patients, but subscale findings - particularly the sensory domain - should be interpreted cautiously and confirmed in broader multicentre validation studies.

## REFERENCES

1. Melzack R. The McGill pain questionnaire: from description to measurement. *Anesthesiology*. 2005;103(1):199-202. doi:10.1097/00000542-200507000-00028.
2. Melzack R. The McGill Pain Questionnaire: major properties and scoring methods. *Pain*. 1975;1(3):277-299. doi:10.1016/0304-3959(75)90044-5.





3. Melzack R. The short-form McGill Pain Questionnaire. *Pain*. 1987;30(2):191-197. doi:10.1016/0304-3959(87)91074-8.
4. Krakowski JC, Hallman MJ, Smeltz AM. Persistent pain after cardiac surgery: prevention and management. *Semin Cardiothorac Vasc Anesth*. 2021;25(4):289-300. doi:10.1177/10892532211041320.
5. Gagnier JJ, Lai J, Mokkink LB, Terwee CB. COSMIN reporting guideline for studies on measurement properties of patient-reported outcome measures. *Qual Life Res*. 2021;30(8):2197-2218. doi:10.1007/s11136-021-02822-4.
6. Terkawi AS, Tsang S, Abolkhair A, Alsharif M, Alswiti M, Alsadoun A, et al. Development and validation of Arabic version of the Short-Form McGill Pain Questionnaire. *Saudi J Anaesth*. 2017;11(Suppl 1):S2-S10. doi:10.4103/sja.SJA\_42\_17.
7. Park KS, Castaneda AM, Choi JK, Park SY, Jung SA, Kim BR, et al. Psychometric evaluation of the Korean versions of the McGill Pain Questionnaire-Short Form (MGPQ-SF) and Neuropathic Pain Scale (NPS) in neuropathic pain patients. *Int J Pain*. 2017;8(1):1-15.
8. Hasvik E, Haugen AJ, Haukeland-Parker S, Rimehaug SA, Gjerstad J, Grovle L. Cross-cultural adaptation and validation of the Norwegian Short-form McGill Pain Questionnaire-2 in low back-related leg pain. *Spine (Phila Pa 1976)*. 2019;44(13):E774-E781. doi:10.1097/BRS.0000000000002976.
9. Mehdizadeh M, Fereshtehnejad SM, Goudarzi S, Moshtagh A, Dehghanian Nasrabadi F, Habibi SAH, et al. Validity and reliability of Short-Form McGill Pain Questionnaire-2 (SF-MPQ-2) in Iranian people with Parkinson's disease. *Parkinsons Dis*. 2020;2020:2793945. doi:10.1155/2020/2793945.
10. Edirisinghe NP, Makuloluwa TR, Amarasekara TD, Goonewardena CSE. Psychometric properties of Sinhala version of Short-Form McGill Pain Questionnaire-2 (SF MPQ-2-Sin) among patients with cancer pain in Sri Lanka. *Pain Res Manag*. 2019;2019:5050979. doi:10.1155/2019/5050979.
11. Jorns TP, Loppanthong S, Pansaeng P, Puasiri S. Validity and reliability testing of the E-san version of the Short-form McGill Pain Questionnaire in musculoskeletal, neuropathic and odontogenic pain. *J Med Assoc Thai*. 2019;102(4):75-81.
12. Yakut Y, Yakut E, Bayar K, Uygur F. Reliability and validity of the Turkish version Short-Form McGill Pain Questionnaire in patients with rheumatoid arthritis. *Clin Rheumatol*. 2007;26(7):1083-1087. doi:10.1007/s10067-006-0452-1.
13. Sharif A, Amjad F, Arslan SA, Ahmad A. Translation and validation of Urdu version Short Form-McGill Pain Questionnaire-2. *BMC Sports Sci Med Rehabil*. 2023;15(1):102. doi:10.1186/s13102-023-00715-2.
14. Wang JL, Zhang WJ, Gao M, Zhang S, Tian DH, Chen J. A cross-cultural adaptation and validation of the Short-Form McGill Pain Questionnaire-2: Chinese version in patients with chronic visceral pain. *J Pain Res*. 2017;10:121-128. doi:10.2147/JPR.S116997.
15. Bolarinwa OA. Principles and methods of validity and reliability testing of questionnaires used in social and health science researches. *Niger Postgrad Med J*. 2015;22(4):195-201. doi:10.4103/1117-1936.173959.
16. Arafat SMY. Validation study can be a separate study design. *Int J Med Sci Public Health*. 2016;5(11):2421-2422. doi:10.5455/ijmsph.2016.19042016471.
17. Dworkin RH, Turk DC, Revicki DA, Harding G, Coyne KS, Peirce-Sandner S, et al. Development and initial validation of an expanded and revised version of the Short-form McGill Pain Questionnaire (SF-MPQ-2). *Pain*. 2009;144(1-2):35-42. doi:10.1016/j.pain.2009.02.007.
18. Rodriguez CS. Pain measurement in the elderly: a review. *Pain Manag Nurs*. 2001;2(2):38-46. doi:10.1053/jpmn.2001.23746.
19. Ferraz MB, Quaresma MR, Aquino LR, Atra E, Tugwell P, Goldsmith CH. Reliability of pain scales in the assessment of literate and illiterate patients with rheumatoid arthritis. *J Rheumatol*. 1990;17(8):1022-1024.
20. Arovah NI, Heesch KC. Assessment of the validity and reliability of the Indonesian version of Short Form 12 (SF-12). *J Prev Med Hyg*. 2021;62(2):E421-E429. doi:10.15167/2421-4248/jpmh2021.62.2.1878.
21. Luo X, George ML, Kakouras I, Edwards CL, Pietrobon R, Richardson W, et al. Reliability, validity, and responsiveness of the Short Form 12-item survey (SF-12) in patients with back pain. *Spine (Phila Pa 1976)*. 2003;28(15):1739-1745. doi:10.1097/01.BRS.0000083169.58671.96.
22. Alharbi HA, Albabtain MA, Alobiad N, Aba Alhasan J, Alruhaimi M, Alnefisah M, et al. Pain perception assessment using the Short-Form McGill Pain Questionnaire after cardiac surgery. *Saudi J Anaesth*. 2020;14(3):343-348. doi:10.4103/sja.SJA\_34\_20.
23. Adelmanesh F, Arvantaj A, Rashki H, Ketabchi SM, Montazeri A, Raissi G. Results from the translation and adaptation of the Iranian Short-Form McGill Pain Questionnaire (I-SF-MPQ): preliminary evidence of its reliability, construct validity and sensitivity in an Iranian pain population. *Sports Med Arthrosc Rehabil Ther Technol*. 2011;3(1):27. doi:10.1186/1758-2555-3-27.
24. Tsui FPY, Wong SSC, Chan TCW, Lee Y, Cheung CW. A validation study of the Cantonese Chinese version of Short Form McGill Pain Questionnaire 2 in Cantonese-speaking patients with chronic pain in Hong Kong. *Pain Pract*. 2024;24(3):449-457. doi:10.1111/papr.13319.
25. Katijjahbe MA, Royse C, Granger C, Deneyh L, Ali NAM, Abdul Rahman MR, et al. Location and patterns of persistent pain following cardiac surgery. *Heart Lung Circ*. 2021;30(8):1232-1243. doi:10.1016/j.hlc.2020.12.009.
26. Viana LBR, Oliveira EJSG, Oliveira CMB, Moura ECR, Viana LHL, Nina VJS, et al. Assessment of pain and quality of life in patients undergoing cardiac surgery: a cohort study. *Rev Assoc Med Bras (1992)*. 2023;69(3):473-478. doi:10.1590/1806-9282.20221655.