Published Online: 2026 March 29 **Research Article**

Translation and Psychometric Characteristics of the Post-liver Transplant Quality of Life Instrument to Persian for Liver Transplantation Recipients

Mohammad Javaherian [1], Nader Abazari [2], Ali Akbar Nejatisafa [3,4,5], Mohsen Nasiri-Toosi [3], Ali Jafarian [3], Mehdi Soleimani [6], Behrouz Attarbashi Moghadam [7,*]

[1] Neuromusculoskeletal Rehabilitation Research Center, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran
[2] Department of Clinical Psychology, Shahed University, Tehran, Iran
[3] Liver Transplantation Research Center, Imam Khomeini Hospital Complex, Tehran University of Medical Sciences, Tehran, Iran
[4] Psychosomatic Research Center, Tehran University of Medical Sciences, Tehran, Iran
[5] Research Center for Cognitive and Behavioral Sciences, Tehran University of Medical Sciences, Tehran, Iran
[6] Department of Psychiatry, Roozbeh Hospital, Tehran University of Medical Sciences, Tehran, Iran
[7] Department of Physiotherapy, School of Rehabilitation, Tehran University of Medical Sciences, Tehran, Iran

[*] ***Corresponding Author:*** Department of Physiotherapy, School of Rehabilitation, Pich e Shemiran, Enghelab Ave., 1981983857, Tehran, Iran. Tel: +98-9121883095; Fax: +98-2177534133; Email: attarbashi@tums.ac.ir

Received: 3 February, 2024; **Revised:** 16 November, 2025; **Accepted:** 6 December, 2025

Abstract

Background: The Post-Liver Transplant Quality of Life (PLTQ) instrument is a specific, disease-based questionnaire to evaluate health-related quality of life (HRQOL) among liver transplantation (LTx) recipients.

Objectives: This study focused on the translation, cross-cultural adaptation, and validation of the psychometric properties of the PLTQ questionnaire to provide the Persian version of the Post-Liver Transplant Quality of Life (PV-PLTQ) instrument.

Methods: This investigation is a translation and cross-cultural adaptation study followed by evaluation of validity and reliability of PV-PLTQ. Between February and June 2021, a total of 175 LTx recipients completed the PV-PLTQ and Short-Form 36 (SF-36) questionnaires at the Liver Transplantation Research Center of Tehran University of Medical Sciences. Sixty-seven randomly selected patients were asked to complete the PV-PLTQ two to three months later. Face validity, content validity, acceptability, ceiling and floor effects, internal consistency, test-retest reliability, responsiveness to change, convergent validity, and minimal detectable change (MDC) were evaluated. Internal consistency was assessed using Cronbach's alpha (≥ 0.7); test-retest reliability using ICC; and responsiveness using effect size (ES), mean difference (MD), and standardized mean difference (SMD). Convergent validity was tested using Spearman's correlation due to nonlinear relationships.

Results: The results of face validity, content validity, and missing data proportion indicated that the PV-PLTQ questionnaire is acceptable and easy to understand. Cronbach's alpha coefficient for the PV-PLTQ questionnaire was 0.97. The results of test-retest reliability showed a moderate to good intraclass correlation coefficient from 0.6 to 0.86 ($P < 0.05$). PV-PLTQ also demonstrated strong responsiveness to HRQOL changes, with a very large effect size (Cohen's d = 3.82; 95% CI: 2.51 - 5.11), indicating substantial improvement over time among participants who reported improvement by SF-36. The correlations between PV-PLTQ domains and physical ($r = 0.65$) and mental ($r = 0.63$) components of SF-36, calculated using Spearman's correlation, presented good convergent validity ($P < 0.05$). No floor effect was observed. The amount of 0.9 was calculated as the MDC of the total PV-PLTQ score.

Conclusions: The PV-PLTQ questionnaire is a valid, reliable, and responsive to change instrument to evaluate HRQOL in LTx recipients.

Keywords: Quality of Life, Liver Transplantation, Reliability, Questionnaire, Validity, Persian

How to Cite: Javaherian M, Abazari N, Nejatisafa AA, Nasiri-Toosi M, Jafarian A, et al. Translation and Psychometric Characteristics of the Post-liver Transplant Quality of Life Instrument to Persian for Liver Transplantation Recipients. Middle East J Rehabil Health Stud. 2026;13(2):e145622. doi: https://doi.org/10.5812/mejrh-145622

1. Background

Liver transplantation (LTx) is a well-established life-saving intervention for patients with end-stage liver disease. The one-year survival rate of liver transplant recipients in Iran is approximately 85%, reflecting significant advancements in transplantation techniques and post-operative care (1). Primarily, the survival rate of patients after transplantation was a criterion for the evaluation of clinical and research outcomes of this procedure (2). The survival rate of LTx among recipients has grown significantly in the past two decades owing to advancements in treatments, drugs, and concentration on multidisciplinary approaches in post-transplant care (3).

Nowadays, recovering patients' social functions and improving the Quality of Life (QOL) are the most essential outcomes of this intervention (4). Health-Related Quality of Life (HRQOL), as a component of QOL, can be affected by diseases and interventions through alteration of lifestyle, mental balance, and well-being among individuals with various conditions, according to their own judgments and perceptions (5). Due to the growing number of LTx recipients, their special conditions, and the importance of post-treatment QOL, it is necessary to assess their HRQOL via a standard and specific instrument (6). An accurate assessment of HRQOL depends on the psychometric features of the instrument developed to measure this construct (7). Because differences in language, cultural beliefs about illness, and healthcare organization can affect how patients interpret questionnaires, simple literal translation is insufficient. Rigorous cross-cultural adaptation is therefore required to ensure that HRQOL instruments remain conceptually equivalent and valid in Iranian liver transplant recipients (8-10). In Iran, LTx has grown significantly over the past two decades, with specialized centers now performing hundreds of procedures annually. However, regional disparities in access, variations in follow-up care, and differences in patient education may influence post-transplant recovery and QOL.

Health-related QOL is a critical outcome for LTx recipients and it has been assessed in several studies using different instruments including generic and disease-related questionnaires. Generic instruments such as the short form-36 (SF-36) are widely utilized because they allow comparisons across different diseases and populations; however, they may not fully address transplant-specific issues such as immunosuppressive medication side effects, graft rejection concerns, or the financial burdens often accompanying LTx. Disease-specific instruments have also been used to capture unique challenges faced by LTx recipients. For instance, the Chronic Liver Disease Questionnaire (CLDQ) focuses on symptoms relevant to patients with chronic liver disease, while the Diabetes Digestive and Kidney Diseases Liver Transplantation Database (NIDDK-LTD) form has 4 domains, including liver disease symptoms, physical function, health satisfaction, and overall wellbeing (11). Despite these advances, there remained a need for a comprehensive, transplant-targeted questionnaire that would thoroughly assess the physical, emotional, and social impacts unique to post-transplant life. Generic questionnaires, such as the SF-36, are designed for broad applicability but often lack the sensitivity to detect subtle, transplant-specific changes in HRQOL. In contrast, disease-specific tools like the CLDQ target chronic liver disease symptoms but may not address post-transplant realities such as complex medication regimens, frequent follow-up visits, or long-term psychosocial adjustments. This gap prompted the development and subsequent cross-cultural validation of more LTx-specific tools, one of the most prominent being the Post-Liver Transplant Quality of Life (PLTQ) instrument (8-10).

To place the PLTQ in context, it is useful to compare its psychometric performance with other liver- and transplant-specific HRQOL instruments. The CLDQ and its language adaptations have consistently shown very good internal consistency, with total Cronbach's alpha coefficients typically around 0.90 - 0.95 and satisfactory domain-level alphas in patients with chronic liver disease (8). In validation studies of the PLTQ, including the Brazilian Portuguese version, the total score has shown excellent internal consistency (Cronbach's alpha $\approx$ 0.91) and high reproducibility (ICC $\approx$ 0.90), together with acceptable convergent validity and limited floor and ceiling effects. These data indicate that the PLTQ is psychometrically robust and support the need to further evaluate its performance in new linguistic and cultural settings, such as Persian-speaking liver transplant recipients (10).

Post-Liver Transplant Quality of Life was originally developed to address specific issues faced by LTx recipients. This questionnaire comprises 32 items distributed across eight domains: Emotional Function, Worry, Medications, Physical Function, Healthcare, Graft Rejection Concern, Financial, and Pain. Each item is scored on a 1 - 7 Likert scale, where higher values indicate better perceived QOL. The PLTQ's domain-specific structure allows for a targeted evaluation of transplant-related factors not typically captured by non-specific HRQOL measures (8).

2. Objectives

The present study focused on the translation, cross-cultural adaptation, and validation of PLTQ to provide the Persian version of this questionnaire (PV-PLTQ). The validity of this instrument was assessed by evaluating the face validity, content validity, acceptability, ceiling and floor effects, responsiveness to change, and convergent validity. Internal consistency and test-retest reliability were also used to assess the reliability.

3. Methods

3.1. Participants

The sample size was calculated as approximately five participants per item, which is a commonly used rule-of-thumb for validation studies of multi-item questionnaires. Given the 32 items of the PLTQ, we aimed to recruit at least 160 participants and ultimately included 180 adult liver transplant recipients. Eligible participants were ≥ 18 years old, had undergone LTx at least three months before enrolment, had a functioning graft, were fluent in Persian, and were able and willing to provide informed consent. We excluded patients with severe cognitive impairment, acute rejection episodes, or other unstable medical conditions that prevented them from completing the questionnaire. After explaining the study procedures, informed consent was obtained from all participants. This study was approved by the Ethics Committee of Tehran University of Medical Sciences (approval ID: IR.TUMS.IKHC.REC.1397.068). Informed consent was obtained from each participant.

3.2. Procedure

This study contains three phases: (1) Translation and cross-cultural adaptation of the PV-PLTQ instrument; (2) Analysis of convergent validity and ceiling and floor effects of the instrument by a cross-sectional design; (3) Assessment of the reliability and responsiveness to change of the questionnaire by a prospective cohort design.

3.3. Instruments

3.3.1. Post-Liver Transplant Quality of Life Instrument

The PLTQ instrument is a specific, disease-targeted questionnaire for measuring HRQOL in LTx recipients. This instrument was developed by Saab et al. (8). It contains 32 items in eight domains, including emotional function (four items), worry (seven items), medications (four items), physical function (six items), healthcare (four items), graft rejection concern (two items), financial (two items), and pain (three items). Each item is scored on a Likert scale from 1 (person's experience with the item all the time) to 7 (person's experience with the item none of the time). The total score is expressed as the average of all items, with higher values presenting better HRQOL.

3.3.2. Short-Form Health Survey (SF-36)

Short-Form Health Survey is a well-known, generic HRQOL questionnaire. This instrument has 36 questions with different scoring systems in eight health-related domains: physical functioning (10 items), role limitations due to physical problems (4 items), bodily pain (2 items), general health perceptions (5 items), vitality (4 items), social functioning (2 items), role limitations due to emotional problems (3 items), and perceived mental health (5 items). Additionally, it includes a single item measuring perceived changes in general health status over a one-year period (health transition). These domains form two distinct subscales: Mental Component Summary Score (MCS) and Physical Component Summary (PCS) Score. This survey has been translated and cross-culturally validated to a Persian version (12).

3.4. Translation and Cross-Cultural Adaptation

Initially, permission for translation and cross-cultural adaptation was obtained from Sammy Saab. The processes of translation and cross-cultural adaptation were followed based on the published guideline (13). Accordingly, two independent native Persian language

speakers with good English language skills (one physiotherapist and one psychologist) translated the English version of the PLTQ questionnaire into Persian. The two forward translations were reviewed and synthesized into a single version by the research team through discussion. Any discrepancies were resolved by consensus before the next step. Two professional English translators independently translated this version back to the English version. These translators were blinded to the original questionnaire. Finally, an expert panel (including a psychiatrist, a psychologist, two transplant surgeons, two physiotherapists, and two hepatologists) carefully reviewed all versions of the questionnaire. They evaluated semantic, idiomatic, experiential, and conceptual equivalence of each item, discussed any discrepancies, and considered whether examples or terms required adaptation to fit Persian cultural norms and healthcare context. Equivalent terms were selected considering patients' beliefs about disease and QOL in the Persian cultural context. Decisions were made by consensus, and after incorporating the agreed changes, the pre-final PV-PLTQ questionnaire was prepared (Appendix 1 in Supplementary File). A full explanation regarding any change in the terms of PV-PLTQ based on Persian culture is provided in Appendix 2 in Supplementary File.

3.5. Face Validity and Content Validity

In the presence of one of the authors, the pre-final PV-PLTQ questionnaire was read and completed by 15 LTx recipients to evaluate the difficulties in understanding and the levels of clarity of instruction, items, and response format. Each participant received a paper-based copy of the questionnaire along with a brief written instruction sheet. The participants completed the questionnaire independently in a quiet consultation room within the clinic to minimize distractions. 87% of participants stated that all contents were totally clear and understandable. Any questions or difficulties the participants encountered were documented on a standardized feedback form. We also recorded the approximate time taken to complete the questionnaire. Participants were encouraged to make written or verbal notes on any items they found ambiguous or culturally irrelevant. After reviewing the comments by the expert panel and making some corrections, the final PV-PLTQ instrument was developed. Furthermore, we asked the expert panel members to score the content equivalency including semantic equivalency, conformity with Persian culture, and ease of understanding and responding of instruction, items, and scales by a five-point Likert scale (from very poor to very strong). Each expert independently rated the clarity and relevance of every item on a five-point Likert scale (1 = very poor, 5 = very strong). These ratings were then discussed in a face-to-face meeting. Items that received low ratings or raised concerns in terms of wording, cultural appropriateness, or redundancy were revised and re-evaluated until full consensus was reached. No items were deleted, and the final version reflected both the numerical ratings and qualitative comments of the panel. Although we did not calculate formal content validity indices (CVI or CVR), this process provided a structured, consensus-based assessment of content validity.

To reduce selection bias, participants were recruited consecutively during routine follow-up visits at the transplant clinic. All questionnaires were administered by trained interviewers who were not involved in the patients' clinical care, using a standardized script and neutral wording. Patients were assured that their responses would remain confidential and would not influence their treatment, in order to minimize social desirability bias.

3.6. Reliability

The reliability of the PV-PLTQ questionnaire was tested by the test-retest analysis. Also, Cronbach's alpha was provided for assessing its internal consistency. To carry out test-retest reliability, 70 LTx recipients were randomly selected for a longitudinal assessment and were asked to complete the PV-PLTQ questionnaire for the second time two to three months later. Also, they were asked to answer the global rating of change (GRC) questionnaire in the second test to measure the reproducibility and responsiveness of the PV-PLTQ instrument to change. This questionnaire is a single-item measure designed to capture a patient's overall perceived change in health status over a specified period. Participants are typically asked to compare their current condition (e.g., pain, function, well-being) to their status at a previous time point, and then rate this comparison on a 7-point Likert scale ranging from "much worse" to "much better." In our study, LTx recipients were asked: "Compared to your condition [two to three months ago], how would you describe your

overall status now?" Response options ranged from −3 (indicating "much worse") to +3 (indicating "much better"), with 0 meaning "no change." We used this instrument to categorize individuals into three groups: improved, deteriorated, or unchanged. Those who reported "no change" (0 on the GRC scale) were included in the test-retest reliability analysis to ensure that any differences in scores were due to measurement error rather than actual clinical changes (14). Based on the GRC score, patients were divided into three categories: (1) patients with improvement (+1 to +3); (2) patients with deterioration (-1 to -3); and (3) patients with no change (0). Of the 70 randomly selected participants for the longitudinal survey, two of them were not willing to continue the study due to personal reasons and one was hospitalized.

In addition, standard error of measurement (SEM) and minimal detectable change (MDC) were calculated for the total PV-PLTQ score and each domain. SEM presents the precision of the individual measurement. In this formula, SD of patients with no changes was considered. MDC is the minimal amount of change in an instrument's score that should occur in an individual to ensure that this change is not related to measurement error (14). MDC was calculated as 1.96 * sqrt (2) * SEM (15).

3.7. Convergent and Divergent Validity

At the time of the first assessment, participants completed both the PV-PLTQ and the Persian version of the SF-36. We assessed construct validity by examining Spearman's correlation coefficients between PV-PLTQ scores (total and domains) and the SF-36 PCS and Mental Component Summary (MCS) scores. Based on a priori hypotheses, we expected moderate-to-strong positive correlations ($r \geq 0.50$) between the PV-PLTQ total score and both SF-36 component summaries, and between conceptually related domains (e.g., emotional function and worry with MCS; physical function and pain with PCS). In contrast, we expected low-to-moderate correlations ($r < 0.40$) between clearly distinct constructs (e.g., the cost domain versus PCS/MCS), which would support divergent validity. Correlations were interpreted as low (< 0.30), moderate (0.30 - 0.49), or strong (≥ 0.50).

3.8. Ceiling and Floor Effects

It is essential to assess whether a substantial portion of participants achieve the lowest (floor) or highest (ceiling) possible scores on the instrument (16). For analysis of ceiling and floor effects, the percentage of participants with the lowest (i.e., 1) and highest (i.e., 7) possible score in total PV-PLTQ and each domain score are taken into account. It is considered that the ceiling and floor effects are present if more than 15% minimal detectable difference of patients scored the highest or lowest total score. Identifying ceiling and floor effects ensures that the questionnaire is appropriately sensitive across the spectrum of patient experiences, which is crucial for both clinical and research applications (16).

3.9. Statistical Analysis

All data were analyzed using the Statistical Package for Social Science version 18.0 (SPSS, INC, Chicago, IL). A P-value < 0.05 was considered statistically significant. The Shapiro-Wilk test was established to evaluate the normal distribution of data.

Acceptability and simplicity of the PV-PLTQ questionnaire were assessed by calculating the missing data proportion for each item and domain. A missing data proportion of ≤ 4% was considered satisfactory. Missing data were handled using multiple imputation by chained equations (MICE) with an ordered-logit (proportional-odds) model (m = 20). To improve the plausibility of missing at random and enhance precision, model for end-stage liver disease (MELD) score was included as the main predictor. Domain and total PLTQ scores were computed as passive functions of the items after imputation (17). To estimate internal consistency for the total instrument and each subscale, Cronbach's alpha was computed. Cronbach's alpha ≥ 0.7 was considered satisfactory (18). The intraclass correlation coefficient (ICC) (2, 1) was used to evaluate test-retest reliability in patients who showed "no change" between the two assessment phases (15). The ICC values < 0.5, 0.5 - 0.75, 0.75 - 0.9, and > 0.9 were considered poor, moderate, good, and excellent reliability measures, respectively (15). We calculated ICC using this formula: (MSR - MSE) / (MSR + (K - 1) * MSE).

To evaluate the responsiveness to change of the PV-PLTQ instrument, the effect size (ES) was measured by calculating the mean difference (MD) and standardized mean difference (SMD) using Cohen's d (19). For this purpose, the changes in the PV-PLTQ questionnaire total scores were compared between patients who reported "improvement" and "deterioration" during a longitudinal period (20). The Spearman correlation

Table 1. Demographic and Clinical Characteristics of Participants [a]

Characteristics	Cross-Sectional Survey	Longitudinal Survey
Total number	175	67
Age (y)	46.2 ± 12.4	47.6 ± 12.8
Gender		
Male	115 (66)	48 (72)
Female	60 (34)	19 (28)
Level of education		
Under diploma	35 (20)	13 (19.4)
Diploma	103 (58.8)	38 (56.7)
Undergraduate	26 (14.9)	11 (16.4)
Postgraduate	11 (6.3)	5 (7.5)
Time since transplantation at initial survey (mo)		
< 6	33 (18.8)	15 (22.4)
6 - 12	48 (27.4)	16 (23.9)
> 12	94 (53.8)	36 (53.7)
MELD score at transplant time	19.7 ± 5.6	18.9 ± 5.9
PV-PLTQ score in the first assessment time	5.18 ± 1.49	5.89 ± .84
PV-PLTQ score in the second assessment time	NA	5.9 ± 1.03
SF-36 score in the first assessment time		
Physical component summary score	71.3 ± 17.3	75.5 ± 14.6
Mental component summary score	72.7 ± 18.7	78.4 ± 16.9

Abbreviation: MELD, model for end-stage liver disease; PV-PLTQ, Persian version of post-liver transplant quality of life instrument; SF-36, short form-36; NA, not applicable.

[a] Values are expressed as mean ± SD or No. (%).

coefficient between PV-PLTQ questionnaire total and each domain scores and MCS and PCS scores were established to test convergent validity. This type of correlation coefficient test was selected as the scatterplot showed nonlinear relationships between PV-PLTQ questionnaire total and each domain scores and MCS and PCS scores. MCS and PCS were calculated by the principal investigator of translation and cross-cultural adaptation Persian version of SF-36 (12).

4. Results

4.1. Patients Characteristics

Out of 180 included patients, five did not complete more than 50% of the items and therefore were excluded. Table 1 summarizes the demographic characteristics of the participants in the cross-sectional and longitudinal assessment phases. In the cross-sectional survey, there were 175 participants, with an average age of 46.2 ± 12.4 years, while the longitudinal survey included 67 patients of them randomly, with an average age of 47.6 ± 12.8 years. After follow-up duration, 51, 12, and four presented no change, improved, and worsened, respectively, based on GRC scores. Gender distribution showed that 66% of participants in the cross-sectional survey and 72% in the longitudinal survey were male. Regarding education, 20% of the cross-sectional participants had less than a diploma, 58.8% held a diploma, 14.9% were undergraduate, and 6.3% had a postgraduate degree. In the longitudinal survey, 19.4% had less than a diploma, 56.7% held a diploma, 16.4% were undergraduate, and 7.5% had a postgraduate degree.

Time since transplantation varied in both surveys, with 18.8% of the cross-sectional group transplanted less than 6 months ago, 27.4% between 6 - 12 months, and 53.8% more than 12 months. In the longitudinal survey, 22.4% were transplanted less than 6 months ago, 23.9% between 6 - 12 months, and 53.7% more than 12 months. The mean MELD score at transplant time was 19.7 ± 5.6 in the cross-sectional survey and 18.9 ± 5.9 in the longitudinal survey. The PV-PLTQ score was 5.18 ± 1.49 in the first assessment of the cross-sectional survey and 5.89 ± 0.84 in the longitudinal survey. Since the cross-sectional group was not followed, the second PV-PLTQ evaluation was not applicable, but for the longitudinal

Table 2. Reliability of Persian Version of Post-Liver Transplantation Quality of Life Questionnaire

PV-PLTQ Score/Domain	Mean ± SD	Cronbach's Alpha	ICC [a] [95% CI] (P-Value) [b]	SEM	MDC	Proportion (%) of Missingdata
Total Score	5.17 ± 1.5	0.97	0.87 [0.76, 0.93]	0.32	0.9	-
Healthcare	5.7 ± 1.9	0.95	0.73 [0.55, 0.85]	0.42	1.17	2.8
Emotional	4.9 ± 1.54	0.9	0.82 [0.69, 0.9]	0.52	1.45	3.4
Medication	5.3 ± 1.7	0.88	0.8 [0.65, 0.89]	0.54	1.49	4
Physical	5.2 ± 1.6	0.89	0.78 [0.61, 0.87]	0.49	1.36	6.2
Cost	4.8 ± 2	0.82	0.6 [0.36, 0.77]	1.14	3.15	4

Abbreviations: PV-PLTQ, Persian Version of Post-Liver Transplantation Quality of Life Instrument; ICC, Intraclass correlation coefficient; SEM, Standard error of measurement; MDC, minimal detectable change.

[a] ICC was measured in the participants presented "no change" during two assessment phases based on Global Rating Scale (n = 40).

[b] Statistically significant.

group, it was 5.9 ± 1.03. Subgroup analyses of PV-PLTQ total scores according to demographic and clinical characteristics are presented in Appendix 3 in Supplementary File. There were no statistically significant differences in PV-PLTQ total scores between patients younger than 50 years and those aged ≥ 50 years, between men and women, or across levels of education (all $P > 0.05$). Likewise, PV-PLTQ total scores were comparable across categories of time since transplantation and MELD score (all $P > 0.05$), indicating broadly similar post-transplant HRQOL in the subgroups examined.

4.2. Acceptability

All of the items had less than 4% missing data, except for items 20 (learning to walk after surgery) and 21 (patients' concerns about the ability to drive) with 6% missing data (Table 2). The patients in the face validity phase recognized these items to be totally clear. This low amount of missing data indicates good acceptance and understanding of the items.

4.3. Reliability

The values for Cronbach's alpha exceed the satisfactory level value of more than 0.7 in the total instrument and each domain (Table 2). The ICC values for the total PV-PLTQ questionnaire, healthcare, emotional, medication, physical, and cost domains were 0.88, 0.73, 0.82, 0.8, 0.78, and 0.6, respectively ($P < 0.001$). Although the Cost domain showed a relatively low ICC (0.60), it still meets the threshold for moderate reliability. Participants in the longitudinal survey presented change of PV-PLTQ scores of 0.006 ± 1.84, -0.02 ± 0.45, and 0.9 ± 0.83 in patients presented no change, got worse, and better. The result of responsiveness analysis shows that PV-PLTQ can be sensitive to changes in patients' overall status during time with a very strong ES (MD: 2.25 [95% CI: 1.87, 2.72]; Cohen's d: 3.82 [95% CI: 2.51, 5.11]).

The amount of SEM and MDC for the total PV-PLTQ score was calculated 0.32 and 0.9, respectively. The results of SEM and MDC calculation for each domain are presented in Table 2.

4.4. Convergent and Divergent Validity

The result of the Spearman test presented a moderate to high correlation between the total and all items of the PV-PLTQ questionnaire and MCS and PCS ($P < 0.05$). The highest Spearman correlation value was related to the physical domain score of the PV-PLTQ questionnaire and the PCS of SF-36 ($r = 0.65$, $P < 0.01$, Table 3). In the same correlation matrix, we also explored divergent (discriminant) validity. None of the correlations between the PV-PLTQ total score or domains and the SF-36 component summaries exceeded 0.70 (r range = 0.34 - 0.65), indicating that the PV-PLTQ is associated with the generic HRQOL construct. As expected, the Cost domain, which reflects transplant-related financial burden rather than physical or mental health status, showed the lowest correlations with both the SF-36 Mental and Physical Component Summary scores ($r = 0.37$ and $r = 0.34$, respectively), supporting its divergent validity with respect to these more general health components.

4.5. Ceiling and Floor Effects

Analysis of ceiling and floor effects indicated that the total PV-PLTQ score did not demonstrate any floor or ceiling effects. At the domain level, the emotional,

Table 3. Spearman Correlation of PV-PLTQ Domain and SF-36

PV-PLTQ Score/Domain	SF-36	
	Mental Component Summary (P-Value)	Physical Component Summary (P-Value)
Total Score	0.63 (0.00)	0.65 (0.00)
Healthcare	0.49 (0.00)	0.44 (0.00)
Emotional	0.6 (0.00)	0.63 (0.00)
Medication	0.55 (0.00)	0.55 (0.00)
Physical	0.55 (0.00)	0.61 (0.00)
Cost	0.37 (0.00)	0.34 (0.00)

medication, physical, cost, worry, graft rejection, and pain domains all showed no floor or ceiling effects, as fewer than 15% of participants scored at the lowest or highest possible values. However, the Healthcare domain demonstrated a ceiling effect, with 37% of participants reporting the highest possible score. This finding suggests that, while most domains of the PV-PLTQ were appropriately sensitive across the range of patient experiences, the Healthcare domain may have limited ability to discriminate among patients with better post-transplant outcomes.

5. Discussion

The aim of this study was evaluating the translation, cross-cultural adaptation, and validation of the PLTQ questionnaire to provide the PV-PLTQ instrument. Based on the results of the present study, PV-PLTQ seems to have acceptable internal consistency with moderate to good test-retest reliability, very strong sensitivity to change, high correlation to PCS and MCS, and without floor or ceiling effects. It seems that this instrument can be easily understood by all Iranian population as the expert panel tried to use simple, clear, and commonly-used words by LTx recipients. During translation, different medical specialists that face patients after transplantation discussed to reach a clinical and usable instrument.

The PV-PLTQ questionnaire seems to be relatively acceptable and simple to be completed based on the results of face validity and missing data proportion. Items 20 and 21 had higher missing data proportions. These items are related to patients' concerns about learning how to walk and drive. This may be because most patients in this center are mobilized and learn walking during hospitalization by the physiotherapist. Subsequently, they can walk and drive soon after LTx; therefore, it seems that asking about patients' concerns about walking and driving in the condition may be perceived as irrelevant under current clinical conditions. In addition, although our sample included participants with a wide range of educational levels, the results of the face validity and missing data analysis suggest that education did not substantially hinder the completion of the PV-PLTQ. Most participants, including those with less than a diploma, were able to complete the questionnaire with minimal missing responses, and 87% of respondents in the pilot testing phase stated that the contents were fully clear and understandable. Nevertheless, it should be acknowledged that lower educational levels may affect the interpretation of some items, especially those related to abstract psychosocial constructs.

The MDC for the total PV-PLTQ score was calculated at 0.9. This means that a change of greater than 0.9 must be obtained to indicate true changes whether improvement or deterioration on HRQOL. This value was calculated 3.15 for the cost domain, which means a relatively large amount of score must be obtained to show a change in liver transplant recipients' cost status.

5.1. Internal Consistency, Test-Retest Reliability, and Responsiveness to Change

Cronbach's alpha of the PV-PLTQ questionnaire was 0.97, which is greater than previously reported values by Saab et al. (8), Xiao et al. (9), and Molski et al. (10). This result presented that the PV-PLTQ instrument has satisfactory internal consistency. Although alpha values above 0.90 are sometimes interpreted as indicating possible item redundancy, this result should be interpreted with caution. The PLTQ is a disease-specific instrument with 32 items that all tap related aspects of post-liver-transplant HRQOL (emotional function, worry, medications, healthcare, financial burden, pain, etc.), and our sample was drawn from a single centre with

relatively similar clinical characteristics. Both factors can contribute to very high alpha values. In this initial cross-cultural adaptation we purposely retained all items and the original domain structure in order to preserve content validity and ensure full comparability with the original PLTQ. Nevertheless, future studies with larger and more heterogeneous samples may use modern psychometric methods (e.g., exploratory factor analysis, Rasch modelling or item response theory) to identify potentially redundant items and to explore whether a shorter version of the PV-PLTQ could be developed without compromising its content coverage.

We selected two to three months as the interval between the two assessment phases. Although the results of a published systematic review indicated that in the first months after transplant, the HRQOL increased after LTx (19), some components of HRQOL may be deteriorated in the short time after LTx due to surgical or psychological complications or some immunosuppressive drug side-effects. It occurs mainly in the first months after transplant and can influence the test-retest reliability. In order to resolve this confounding factor, GRC was used to track participants' changes during two to three months. The ICC values for each domain demonstrated varying levels of reliability. The Total Score domain showed excellent reliability with an ICC of 0.87 [0.76, 0.93]. The Healthcare domain exhibited moderate to good reliability, with an ICC of 0.73 [0.55, 0.85]. The Emotional domain also demonstrated good to excellent reliability with an ICC of 0.82 [0.69, 0.9]. The Medication domain displayed good reliability, with an ICC of 0.80 [0.65, 0.89]. The Physical domain had moderate to good reliability, with an ICC of 0.78 [0.61, 0.87]. Finally, the Cost domain showed moderate reliability, with an ICC of 0.60 [0.36, 0.77]. Overall, most domains exhibited good to excellent reliability, with the exception of the Cost domain, which had moderate reliability. This level of reliability means should not be overinterpreted and only changes exceeding the MDC (3.15 score) represent true changes beyond measurement error. In clinical settings, physicians and researchers should interpret cost domain results with caution, and ideally complement them with additional socioeconomic assessments. Conversely, other domains including physical, emotional, and medication provide more stable measures for monitoring subtle changes in HRQOL. Although Saab et al. (8) did not report the ICC for test-retest reliability, our results are similar to those of other studies translating and validating the PLTQ instrument (9, 10). The changes in the PV-PLTQ questionnaire score were used to evaluate the power of this instrument for responsiveness to change. The result of the ES showed that the PV-PLTQ instrument could be considered responsive to change Persian HRQOL instrument. In total, as expected, PV-PLTQ scores improved in parallel with SF-36 PCS and MCS scores in the longitudinal survey, further supporting convergent validity.

5.2. Convergent Validity

The PV-PLTQ questionnaire and the MCS and PCS showed relatively good, significant convergent validity. Although all the results of the Spearman test showed a statistically significant correlation between the PV-PLTQ instrument domains and MCS and PCS, the correlation between healthcare and cost domains was less than those of others. This may be due to the nature of the SF-36 questionnaire content. The SF-36 does not have questions about patients' financial conditions. Moreover, there are some questions about patients' current health status, but there are no questions about individuals' concerns about healthcare in SF-36 as a generic tool for the assessment of HRQOL. Therefore, the lower correlations do not indicate poor validity but rather highlight that the cost domain measures a distinct dimension of health-related QOL not covered by the SF-36. This underscores the added value of using disease-specific instruments such as the PV-PLTQ, which provide a more comprehensive and culturally sensitive assessment of patients' post-transplant experiences.

No floor or ceiling effects were present in the scores of total PV-PLTQ. It means that the PV-PLTQ questionnaire can generally distinguish LTx recipients with the lowest and highest possible score (16). The Healthcare domain showed a ceiling effect as 37% of participants reported the highest scale in the items of this domain. The items of this domain ask LTx recipients about their "problem following instructions for taking your transplant medications", "trouble about needing to make special arrangements because of frequent doctor visits", "concern about needing to have multiple blood draws", "concern about developing complications from taking your medications incorrectly or forgetting to take them", and "bother by having long waits for doctor appointments". Most of the patients who showed a ceiling effect in this domain (94.3%) had

transplantation more than 12 months ago; therefore, it seems that there is no problem in the items related to the healthcare domain among patients with more than 12 months since their LTx. In line with the original PLTQ, where the Financial domain had the lowest internal consistency among domains, the Cost domain of the PV-PLTQ showed only moderate test–retest reliability (ICC = 0.60) despite good internal consistency (α = 0.82). This pattern likely reflects genuine short-term variability in patients' transplant-related financial burden and the influence of contextual factors such as insurance coverage and timing of medical expenses, rather than a pure measurement problem.

The study conducted by Saab et al., which first developed the questionnaire, established its validity and reliability as a robust tool for LTx recipients (8). Similarly, Molski's investigation into the cultural validation of this questionnaire for the Brazilian population corroborated these findings, demonstrating that the questionnaire maintains its validity and reliability even when applied in a different cultural context (10). This alignment between our findings and those of Saab et al. and Molski underscores the consistency and generalizability of the questionnaire across diverse settings. Such convergence in findings not only reinforces the credibility of the questionnaire but also highlights its potential utility in cross-cultural research (8, 12).

Despite the strengths of our study, including a systematic translation and cross-cultural adaptation process, robust psychometric evaluation, and a relatively large sample size, several limitations should be acknowledged. First, participants were recruited from a single national referral liver transplant centre; thus, our findings may not fully represent LTx populations in other regions or healthcare settings. This limitation may restrict the generalisability of our findings to other settings and regions. Future multi-centre studies across different geographic and sociocultural contexts are needed to confirm the generalisability of the PV-PLTQ. Second, the cross-sectional design and the relatively modest sample size, particularly for the test–retest subsample, may have reduced the precision of some reliability estimates. Third, the 2 - 3-month interval between the first and second assessment may have introduced recall bias, even though we used the Global Rating of Change (GRC) and restricted test–retest analyses to patients who reported "no change" in their overall status. Fourth, the PV-PLTQ relies on self-reported HRQOL, so the data are inherently influenced by patients' subjective perceptions, mood, and expectations. Moreover, because the questionnaires were completed in a face-to-face clinical setting, the presence of research staff and the clinical environment may have promoted social desirability and other response biases, potentially leading some participants to over-report favourable HRQOL. These issues should be considered when interpreting our results and planning future validation studies. Fifth, although COSMIN recommends assessment of structural validity using factor analysis, we did not perform exploratory or confirmatory factor analysis in this study. Our sample size, while adequate for the planned analyses, was at the lower end for a stable factor analysis of a 32-item, eight-domain instrument, and our primary aim was to preserve the original PLTQ structure rather than to redevelop it. As a result, construct validity was examined through internal consistency, test–retest reliability, hypothesis-driven correlations with the SF-36, and subgroup analyses, but not through EFA or CFA.

Based on these limitations, several specific areas for future research can be identified. First, multicentre studies in different regions of Iran and in more heterogeneous clinical settings are needed to further validate the PV-PLTQ and to examine its performance in earlier post-transplant phases, in newly transplanted patients, and in those with more complex comorbidities. Second, head-to-head comparisons of the PV-PLTQ with other disease-specific instruments such as the Chronic Liver Disease Questionnaire and with generic tools (e.g., SF-36) in newly transplanted and long-term recipients would help clarify the added value and complementary roles of these measures. Fifth, longitudinal studies should investigate responsiveness to clinical change and determine indices such as minimal important change and minimal important difference for the PV-PLTQ, in order to facilitate its use in intervention trials and routine follow-up. Finally, future psychometric research using factor analysis and modern methods (e.g., Rasch or item response theory) may confirm the structural validity of the PV-PLTQ and explore whether a shorter form can be developed without compromising content coverage, which could further support the design of tailored interventions to improve post-transplant well-being.

In terms of clinical implications, the PV-PLTQ can serve as a valuable tool for clinicians and researchers by offering a sensitive, LTx-specific assessment of HRQOL. Routine use of the questionnaire in outpatient follow-up may help identify individuals with low domain-specific scores—such as emotional well-being or financial concerns—who might otherwise be overlooked. Early detection of potential issues can prompt timely referral to mental health services, social support programs, or financial counseling, potentially improving overall outcomes for liver transplant recipients. Additionally, the PV-PLTQ could be used in future clinical trials to evaluate the efficacy of interventions (e.g., new pharmacological regimens or rehabilitation programs) aimed at enhancing QOL after LTx. By recognizing the limitations of the current study, exploring avenues for future research, and outlining practical clinical applications, we believe that the PV-PLTQ instrument provides a robust foundation for ongoing efforts to optimize post-transplant care and improve patient-centered outcomes.

5.3. Conclusions

We cross-culturally adapted the PV-PLTQ instrument. This tool seems to be reliable, valid, and responsive to change. Therefore, the use of the PV-PLTQ questionnaire can be recommended in research and clinical settings to evaluate the HRQOL of Persian-speaking LTx recipients.

Acknowledgements

The researchers sincerely thank all of the liver transplantation team members in Imam Khomeini Hospital Complex for their help and support.

Supplementary Material

Supplementary material(s) is available here [To read supplementary materials, please refer to the journal website and open PDF/HTML].

Footnotes

AI Use Disclosure: The authors declare that no generative AI tools were used in the creation of this article.

Authors' Contribution: M. J., N. A., A. N., M. N., A. J., and B. A. M. made substantial contributions to the conception and design of the study. M. J., N. A., and M. N. contributed to the acquisition of the data. M. J. and M. S. analyzed the data. B. A. M. and A. N. were responsible for data safety monitoring during the study. M. J. and N. A. interpreted the data and drafted the manuscript. All authors had access to the data, revised the manuscript, and approved it.

Conflict of Interests Statement: The authors declare no conflict of interest.

Data Availability: The dataset presented in the study is available on request from the corresponding author during submission or after publication.

Ethical Approval: This study was approved by the Ethics Committee of Tehran University of Medical Sciences (approval ID: IR.TUMS.IKHC.REC.1397.068).

Funding/Support: This research has been supported by the Tehran University of Medical Sciences (grant number: 97-01-205-37867). The funder of the study had no role in study design, data collection, analysis, interpretation, or writing the manuscript. The corresponding author had full access to all the data in the study and was responsible for publishing the results.

Informed Consent: Written informed consent was obtained from the participants.

References

1. Malek-Hosseini SA, Jafarian A, Nikeghbalian S, Poustchi H, Lankarani KB, Nasiri Toosi M, et al. Liver Transplantation Status in Iran: A Multi-center Report on the Main Transplant Indicators and Survival Rates. *Arch Iran Med*. 2018;**21**(7):275-82. [PubMed ID: 30041524].
2. Stine JG, Stukenborg GJ, Wang J, Adkins A, Niccum B, Zimmet A, et al. Liver transplant candidates have impaired quality of life across health domains as assessed by computerized testing. *Ann Hepatol*. 2020;**19**(1):62-8. [PubMed ID: 31558420]. [PubMed Central ID: PMC7252261]. https://doi.org/10.1016/j.aohep.2019.06.018.
3. Dueland S, Line PD, Hagness M, Foss A, Andersen MH. Long-term quality of life after liver transplantation for non-resectable colorectal metastases confined to the liver. *BJS Open*. 2019;**3**(2):180-5. [PubMed ID: 30957065]. [PubMed Central ID: PMC6433324]. https://doi.org/10.1002/bjs5.50116.
4. Dabrowska-Bender M, Kozaczuk A, Paczek L, Milkiewicz P, Sloniewski R, Staniszewska A. Patient Quality of Life After Liver Transplantation in Terms of Emotional Problems and the Impact of Sociodemographic Factors. *Transplant Proc*. 2018;**50**(7):2031-8.

[PubMed ID: 30177104]. https://doi.org/10.1016/j.transproceed.2018.03.113.

5. Durand F. How to improve long-term outcome after liver transplantation? *Liver Int*. 2018;**38**(1):134-8. [PubMed ID: 29427483]. https://doi.org/10.1111/liv.13651.
6. Rodrigue JR, Nelson DR, Reed AI, Hanto DW, Curry MP. Is Model for End-Stage Liver Disease score associated with quality of life after liver transplantation? *Prog Transplant*. 2011;**21**(3):207-14. [PubMed ID: 21977881]. https://doi.org/10.1177/152692481102100305.
7. Dabrowska-Bender M, Michalowicz B, Paczek L. Assessment of the Quality of Life in Patients After Liver Transplantation as an Important Part of Treatment Results. *Transplant Proc*. 2016;**48**(5):1697-702. [PubMed ID: 27496474]. https://doi.org/10.1016/j.transproceed.2015.12.139.
8. Saab S, Ng V, Landaverde C, Lee SJ, Comulada WS, Arevalo J, et al. Development of a disease-specific questionnaire to measure health-related quality of life in liver transplant recipients. *Liver Transpl*. 2011;**17**(5):567-79. [PubMed ID: 21506245]. https://doi.org/10.1002/lt.22267.
9. Xiao P, Yujian N, Hongjuan S, Xiaohong L, Teng MS, Wenxin SZ. The reliability and validity of the Chinese version of Post Liver Transplant Quality of Life Questionnaire. *J Nurs Sci*. 2015;**30**(2):28-31.
10. Molski C, Mattiello R, Sarria EE, Saab S, Medeiros R, Brandao A. Cultural validation of the post-Liver transplant quality of life (pLTQ) questionnaire for the Brazilian population. *Ann Hepatol*. 2016;**15**(3):377-85. [PubMed ID: 27049491]. https://doi.org/10.5604/16652681.1198810.
11. Kim WR. Quality of life instruments for liver transplant recipients: too many choices? *Liver Transpl*. 2000;**6**(6):784-5. [PubMed ID: 11084069]. https://doi.org/10.1053/jlts.2000.19026.
12. Montazeri A, Goshtasebi A, Vahdaninia M, Gandek B. The Short Form Health Survey (SF-36): translation and validation study of the Iranian version. *Qual Life Res*. 2005;**14**(3):875-82. [PubMed ID: 16022079]. https://doi.org/10.1007/s11136-004-1014-5.
13. Sousa VD, Rojjanasrirat W. Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: a clear and user-friendly guideline. *J Eval Clin Pract*. 2011;**17**(2):268-74. [PubMed ID: 20874835]. https://doi.org/10.1111/j.1365-2753.2010.01434.x.
14. Dontje ML, Dall PM, Skelton DA, Gill JMR, Chastin SFM, Seniors U. Reliability, minimal detectable change and responsiveness to change: Indicators to select the best method to measure sedentary behaviour in older adults in different study designs. *PLoS One*. 2018;**13**(4). e0195424. [PubMed ID: 29649234]. [PubMed Central ID: PMC5896945]. https://doi.org/10.1371/journal.pone.0195424.
15. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med*. 2016;**15**(2):155-63. [PubMed ID: 27330520]. [PubMed Central ID: PMC4913118]. https://doi.org/10.1016/j.jcm.2016.02.012.
16. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;**60**(1):34-42. [PubMed ID: 17161752]. https://doi.org/10.1016/j.jclinepi.2006.03.012.
17. Eekhout I, de Vet HC, Twisk JW, Brand JP, de Boer MR, Heymans MW. Missing data in a multi-item instrument were best handled by multiple imputation at the item score level. *J Clin Epidemiol*. 2014;**67**(3):335-42. [PubMed ID: 24291505]. https://doi.org/10.1016/j.jclinepi.2013.09.009.
18. Nunnally JC. *Psychometric Theory 3E*. New York, USA: Tata McGraw-Hill Education; 1994.
19. Fritz CO, Morris PE, Richler JJ. Effect size estimates: current use, calculations, and interpretation. *J Exp Psychol Gen*. 2012;**141**(1):2-18. [PubMed ID: 21823805]. https://doi.org/10.1037/a0024338.
20. Streiner DL, Norman GR, Cairney J. *Health Measurement Scales: A practical guide to their development and use*. New York, USA: Oxford University Press; 2024. https://doi.org/10.1093/med/9780192869487.001.0001.