Research in Health & Medical Sciences *Review Article*

2024 April; 3(2)

Factorial structure, validity, and reliability of the persian version of the professional quality of life questionnaire for nurses

Mahbobeh Abdollahi[1], Somaye Barzanouni[2], Maryam Tatari[1*]

[1] Department of Epidemiology and Biostatistics, School of Health, Torbat Heydariyeh University of Medical Sciences, Torbat Heydariyeh, Iran
[2] Vice Chancellery of Education and Research, Torbat Heydariyeh University of Medical Sciences, Torbat Heydariyeh, Iran

***Corresponding Author Email: maryamtatari2022@gmail.com**
Received: 2025/10; Revised: 2026/2; Accepted: 2026/2

Abstract

Considering the importance of professional quality of life (PROQOL) among nurses in Iran, using updated and revised tools to measure it is essential. Since this tool has been standardized in other countries, adapting it to Iran's culture and context is necessary. This study aimed to determine the psychometric properties of the Persian version of the PROQOL questionnaire for nurses. The study population included nurses working in the educational and medical centers of Torbat Heydariyeh University of Medical Sciences. Three hundred nurses were selected using convenience sampling according to the inclusion criteria. Data were collected using a demographic questionnaire and the 30-item PROQOL questionnaire. Confirmatory factor analysis was performed using Lisrel version 8.8 to examine the factor structure, and internal reliability was assessed with SPSS version 20. Face and content validity were first confirmed using qualitative methods. Due to incomplete questionnaires, data from 267 participants were analyzed. Confirmatory factor analysis showed that after removing three items (2, 15, and 29), the model fit indices improved. Internal reliability was confirmed as Cronbach's alpha for all dimensions was greater than 0.7. The Persian version of the questionnaire, with 27 items, has acceptable construct validity and good internal reliability in the first and third dimensions.

Keywords: Validity, Reliability, Professional Quality of Life Questionnaire

Introduction

Professional quality of life (PROQOL) is an important concept linked to individuals' personality traits and work environment and is significant for health systems and healthcare providers. Employees with higher PROQOL demonstrate stronger organizational identity, greater job satisfaction and performance, and lower intention to leave their job. Studies have shown that programs designed to enhance PROQOL reduce complaints, absenteeism, and disciplinary actions, while increasing positive attitudes and participation in suggestion systems.

PROQOL consists of two dimensions: compassion satisfaction and compassion fatigue. Compassion satisfaction refers to the individual's satisfaction with their ability to perform their work correctly and effectively. It reflects an individual's attitude toward their job and the positive feelings they have about it. This dimension is one of the most important factors affecting organizational productivity and is influenced by personality, organizational, social, and cultural factors.

Compassion fatigue is another key dimension of PROQOL and was first defined as an adverse psychological outcome experienced by nurses. It is a significant variable that directly impacts nurses' PROQOL and is related to their compassion satisfaction. Compassion fatigue consists of two components: secondary traumatic stress and burnout.

Secondary traumatic stress is a negative feeling that arises from witnessing fear and occupational incidents experienced by others. It does not result from direct exposure to incidents or accidents. Secondary traumatic stress can occur suddenly and without warning, and its symptoms may include feelings of helplessness, confusion, and lack of support from others. It is considered an occupational hazard for healthcare workers, especially nurses, as it can affect their ability to provide quality care to patients.

Burnout, the other component of compassion fatigue, consists of three elements: exhaustion, cynicism, and reduced self-efficacy. Its signs often include weakness, hopelessness, emotional fatigue, social withdrawal, irritability, failure, cynicism, and indifference.

Nursing, as the largest professional workforce at the frontline of healthcare delivery, involves multiple and complex roles and responsibilities. Due to this complexity, nursing is a professional activity that requires a high sense of responsibility, attention, and vigilance.

In Iran, several studies have investigated the concept of professional quality of life, including those by Nasraji et al. (2006), Shabani Nejad et al. (2012), and Zakarian et al. (2014). All these studies used self-constructed questionnaires or the Darren Van Laar questionnaire, but none employed the PROQOL questionnaire. The PROQOL version 5 was developed by B. Hudnall Stamm in 2009. This questionnaire includes two subscales: compassion satisfaction and compassion fatigue, with the latter divided into secondary traumatic stress and burnout. Each subscale is independent, and its scores cannot be combined into a single overall score. The questionnaire has 30 items rated on a 5-point Likert scale (1 = never to 5 = always), with 10 items per subscale, and subscale scores are calculated by summing the relevant items. In a study by Pashib et al. (2016), this questionnaire was used but not fully validated.

Since PROQOL is highly important in nursing communities, including among Iranian nurses, measuring it with revised and updated tools provides more accurate and current results for nursing managers. Because this tool has been standardized and psychometrically tested in different countries, it is appropriate to adapt it to the culture and context of Iran. This will allow for the assessment of Iranian nurses' PROQOL through the determination of its factor structure and reliability.

Method

This methodological study was conducted to determine the psychometric properties of the Persian version of the PROQOL tool for nurses. The study population included nurses working in educational and medical centers of Torbat Heydariyeh University of Medical Sciences.

The sample size was determined based on the recommendation of 5–10 participants per item for psychometric evaluation. Therefore, for the 30-item PROQOL, 300 nurses were selected, corresponding to 10 participants per item. Inclusion criteria were: holding an associate, bachelor's, or master's degree in nursing or higher; employment in hospital wards or outpatient clinics; at least one year of work experience in these settings; and willingness to complete the questionnaire. Exclusion criteria included withdrawal of consent during the study or having a specific medical condition that could affect PROQOL.

Data were collected using a demographic questionnaire and the 30-item PROQOL for nurses. The questionnaire consists of two subscales: compassion satisfaction and compassion fatigue, with the latter further divided into secondary traumatic stress and burnout. Each subscale is independent, and subscale scores cannot be combined to provide a total score. Items are rated on a 5-point Likert scale (1 = never to 5 = always), with 10 items per subscale, and subscale scores are obtained by summing the items within each subscale.

Following official permission, the first step in this study was the translation of the tool. The standard forward-backward translation method was used. The original English version of the 30-item questionnaire was first translated into Persian by a bilingual translator. Then, a second translator, independent from the first, back-translated the Persian version into English. Both versions were reviewed by a bilingual expert for consistency and accuracy. Finally, the primary researcher, with input from a panel of experts, finalized the approved Persian version of the questionnaire.

Validity of the questionnaire was assessed in three stages: face validity, content validity, and construct validity. Validity refers to the extent to which an instrument accurately measures the concept it intends to assess. Face validity, considered a type of content validity, was evaluated qualitatively through expert opinion. It addressed aspects such as apparent credibility, logical flow, appropriateness, clarity, brevity, and comprehensiveness of items from the perspective of the target group (nurses). In qualitative face validity, items were reviewed by a panel of 10 experts in terms of difficulty, relevance, coherence, and ambiguity. Quantitative face validity was assessed using the item impact method, where items were rated on a Likert scale from 1 (not important) to 5 (very important). Items with an impact score greater than 1.5 based on a sample of 20 participants were considered suitable for further analysis.

Content validity examined whether the instrument adequately covers all important aspects of the concept. Quantitative content validity was assessed using the Content Validity Ratio (CVR) and Content Validity Index (CVI). CVR was calculated based on expert judgment regarding the necessity of each item. Experts rated items as "essential," "useful but not essential," or "not necessary," and the CVR was calculated using Lawshe's formula. Items with CVR values exceeding the Lawshe table threshold were considered statistically necessary ($P < 0.05$). CVI assessed whether each item was appropriately designed to measure the construct, based on criteria of simplicity, relevance, and clarity. Items with CVI below 0.70 were removed, those between 0.70 and 0.79 were revised, and those above 0.79 were retained. Content validity assessment involved 6 to 15 experts, and the CVI was analyzed using the modified kappa coefficient according to standard kappa analysis.

Construct validity examines the extent to which a tool measures abstract, latent, or unobservable concepts, as constructs are theoretical and broader than specific behaviors. In essence, construct validity evaluates the adequacy of an instrument in measuring the intended constructs. In this study, confirmatory factor analysis is used to assess construct validity. Factor analysis is one of the most reliable methods for determining construct validity, especially for tools that measure psychological characteristics. It helps instrument developers confirm whether the designed tool accurately measures the intended constructs by examining the fundamental factors it aims to capture. Most instrument developers use this method to verify whether the proposed factors actually exist.

In factor analysis, correlated variables are summarized into new variables called factors. By breaking down complex and numerous variables into factors, researchers can identify meaningful patterns. Factor analysis can be exploratory or confirmatory. In exploratory factor analysis, some underlying factors are responsible for shared variance among variables. Based on this approach, the questionnaire was administered to 300 nurses in this study. Exploratory factor analysis was then conducted to identify the number of factors underlying the items, which were measured on a 5-point Likert scale.

The suitability of factor analysis was assessed using the Kaiser-Meyer-Olkin (KMO) measure and Bartlett's test of sphericity. A KMO value of 0.6 or higher indicates an adequate sample size for factor analysis. When factor extraction is justified, exploratory factor analysis using principal component analysis with Varimax rotation is applied. Rotation helps achieve clearer results, and Varimax is recommended for its efficiency. All data from the total sample were included in the analysis. Items with factor loadings of 0.3 or higher were considered to belong to a factor. The number of saturated factors was determined using three main criteria: eigenvalues, the proportion of variance explained by each factor, and the scree plot.

For confirmatory factor analysis, AMOS version 20 was used. Model fit was assessed using fit indices and factor loadings. Fit indices included the goodness-of-fit index (GFI), adjusted goodness-of-fit index (AGFI), root mean square error of approximation (RMSEA), and comparative fit index (CFI). Values of CFI, GFI, and AGFI greater than 0.9, and RMSEA less than 0.1 indicate an acceptable model fit.

Reliability of the instrument was evaluated through two approaches: (a) internal consistency using Cronbach's alpha and (b) external consistency or stability using the test-retest method. For stability, the questionnaire was administered to 30 participants with a two-week interval, and the intraclass correlation coefficient (ICC) was calculated to assess reliability. Internal consistency was assessed using Cronbach's alpha.

Results

In this study, 300 participants were initially recruited, but due to sample loss, data from 267 questionnaires were ultimately analyzed. The mean age of participants was 34.84 ± 6.89 years. Most participants (64.8%, n = 173) were female. The mean number of work shifts was 34.71 ± 20, the mean overtime hours was 93.88 ± 45.07, and the mean work experience was 10.58 ± 6.65 years. Other demographic information is presented in Table 1.

The questionnaire was first translated into Persian using the forward-backward method. Face and content validity of the translated version were assessed qualitatively. After confirming face and content validity, the suitability of the three-factor model was examined. Sampling adequacy was evaluated using the Kaiser-Meyer-Olkin (KMO) test. The KMO statistic was 0.89, exceeding the minimum threshold of 0.8, indicating adequate sampling. Bartlett's test of sphericity was used to assess the justification for factor extraction. The Bartlett test statistic was 1167.55, which was significant at the 95% confidence level ($P < 0.001$), confirming that factor extraction was appropriate.

Initially, the three-factor model with all 30 items was fitted to the data (Model 1). Confirmatory factor analysis indicated that items 2 and 29 had standardized factor loadings below 0.4 and t-values less than 2. Additionally, the RMSEA was greater than 0.1, and the NFI and CFI were below 0.9, indicating poor model fit (Table 2). In Model 2, items 2 and 29 were removed, and confirmatory factor analysis was conducted with 28 items. Fit indices for the 28-item model are reported in Table 2, showing improvement in model fit after removing these two items. Standardized factor loadings, t-values, and multiple correlations are presented in Table 3.

Results indicated that in the second model, item 15 had a standardized factor loading below 0.4 and a t-value less than 2, suggesting that item 15 was not significant in the Persian version. Internal reliability of the tool was assessed using Cronbach's alpha. Cronbach's alpha for the first, second, and third dimensions was 0.89, 0.51, and 0.81, respectively, all above

Table 1. Demographic Characteristics of the Study Participants

Variable	Number (%)
Workplace	
9 Dey Hospital	117 (43/8)
Social Security Hospital	70 (26/2)
Other	80 (30)
Job Position	
Nurse	205 (78/2)
Head Nurse	26 (9/9)
Supervisor	5 (2)
Other	26 (9/9)
Work Shift	
Fixed Morning	28 (10/6)
Fixed Afternoon	3 (1/1)
Fixed Night	1 (0/4)
Rotating	232 (87/9)
Employment Status	
Permanent	181 (68/8)
Contractual	58 (22/1)
Plan-based/Intern	24 (9/1)
Marital Status	
Single	31 (11/7)
Married	234 (88/3)
Health Status	
Healthy	252 (95/8)
Having a disease	11 (4/2)

Table 2. Comparison of the Results from Confirmatory Factor Analysis

	CFI	NFI	RMSEA	X^2/df
Model1	0/88	0/86	0/12	3/66
Model2	0/91	0/90	0/10	3/51

Table 3. Standard Factor Loadings, T-Statistics, and Squared Multiple Correlations in Model 2

Dimension	Question	Standard Factor Loading	**T**	**Squared Multiple Correlation**
First	3	0.67	11/92	0/45
	6	0.60	10/36	0/36
	12	.66	11/78	0/44
	16	0.70	16/57	0/49
	18	0.78	14/77	0/61
	20	0.69	12/30	0/47
	22	0.47	7/68	0/22
	24	0.74	13/73	0/55
	27	0.69	12/31	0/47
	30	0.66	11/51	0/43
Second	1	0.45	7/42	0/20
	4	0.35	5/50	0/12
	8	0.71	-12/78	0/50
	10	0.66	-11/62	0/43
	15	0.20	3/22	0/04
	17	0.39	6/37	0/15
	19	0.66	-11/73	0/44

	21	0.61	-10/50	0/37
	26	0.55	-9/38	0/30
Third	5	0.35	5/55	0/12
	7	0.36	5/73	0/13
	9	0.66	11/46	0/43
	11	0.70	12/54	0/49
	13	0.82	15/72	0/67
	14	0.47	7/81	0/22
	23	0.53	8/78	0/28
	25	0.68	12/13	0/46
	28	0.54	7/31	0/20

the acceptable threshold of 0.7, indicating good internal consistency.

Discussion

This study aimed to translate and evaluate the psychometric properties of the questionnaire in the Iranian population. Findings showed that the Persian version of the questionnaire, after forward-backward translation and validation, demonstrated acceptable validity and reliability, although three items (2, 15, and 29) needed to be adjusted or removed during confirmatory factor analysis.

The demographic characteristics of participants, with a mean age of 34.84 ± 6.89 years, indicate a relatively young population. The predominance of female participants (64.8%) may reflect greater accessibility of this group to the researchers or the gender composition of the target population. The mean work experience of 10.58 ± 6.65 years suggests moderate professional experience among participants. These findings differ from the study by Refaat et al. (2025), which examined the psychometric properties of the Carmen-Q questionnaire in Iranian workers with a mean age of 38.59 years and a predominance of males (66.2%), indicating demographic differences likely due to variations in the target populations.

The translation and content validity process in this study was performed using the forward-backward method, along with qualitative assessment of face and content validity. This approach aligns with international standards, such as the CTAQ checklist developed by Alazeh et al. (2024), where forward-backward translation is considered a key step in the cultural adaptation of questionnaires. Shoushtari-Moghadam et al. (2024), in the Persian translation and validation of the Retirement Resources Questionnaire, applied the WHO protocol, which includes more comprehensive steps such as expert panels and pilot testing. Both approaches are valid, but using combined methods can enhance the credibility of the translation process.

Sampling adequacy was confirmed by a KMO statistic of 0.89 and a significant Bartlett's test ($P < 0.001$). These values indicate that the sample size of 267 was sufficient for factor analysis. Similarly, Shoushtari-Moghadam et al. (2024), with 335 participants, reported acceptable KMO values for both exploratory and confirmatory factor analyses. These results support the adequacy of sampling in the present study.

Confirmatory factor analysis showed that the initial three-factor model with 30 items did not have an acceptable fit. The NFI and CFI values were below 0.9, and RMSEA was greater than 0.1, indicating that the original factor structure was not supported in the Iranian population. After removing items 2 and 29, fit indices improved, but item 15 remained non-significant due to a standardized factor loading below 0.4 and a t-value less than 2.

These findings are comparable to the study by Refaat et al. (2025), who, after translating the Carmen-Q questionnaire, made minor modifications to four items and ultimately reported a four-factor structure for the Persian version, explaining 47.14% of the total variance. Differences in the number of factors and removed items may reflect cultural and linguistic differences between study populations. The inadequate fit of the initial

model highlights the need for cultural adaptation beyond simple translation. As emphasized in the CTAQ checklist, post-translation and cultural adaptation assessment is essential to ensure construct validity in a new context.

Regarding fit indices, although the initial model in this study showed RMSEA greater than 0.1 and NFI and CFI below 0.9, different sources propose varying thresholds for these indices. According to guidelines provided in the R software package, RMSEA values below 0.08 (Avang, 2012) to 0.05 indicate good fit, and CFI values above 0.90 to 0.96 are considered acceptable. In the study by Shoushtari-Moghadam et al. (2024), RMSEA was 0.05, CFI was not reported, IFI was 0.88, and TLI was 0.87, which is close to the acceptable range. The improvement in fit indices after removing problematic items in the present study is consistent with the logic of model modification in confirmatory factor analysis.

Internal reliability of the tool, assessed using Cronbach's alpha, was 0.89, 0.51, and 0.81 for the three dimensions, respectively. Reliability was acceptable in the first and third dimensions (above 0.7) but poor in the second dimension (0.51), indicating internal inconsistency among items in the second subscale or a low number of items in that subscale. Refaat et al. (2025) reported Cronbach's alpha above 0.8 for all subscales in their questionnaire. While Cronbach's alpha is widely used, it is influenced by the number of items, and increasing the number of items can artificially inflate alpha even if inter-item correlations are weak. Madadizadeh and Baharinia (2025) emphasized that in congeneric models, where factor loadings differ across items, McDonald's omega provides a more accurate estimate of reliability than Cronbach's alpha. Therefore, future studies are recommended to report both Cronbach's alpha and omega coefficients.

Overall, the Persian version of the questionnaire, after removing three non-significant items, demonstrates relatively good psychometric properties. However, the low reliability in the second dimension suggests the need to review its content and potentially revise or rewrite items in future research.

Conclusion

Based on the findings of this study, the Persian version of the questionnaire, with 27 items (after removing items 2, 29, and 15), demonstrates acceptable construct validity and good internal reliability in the first and third dimensions. However, the low reliability of the second dimension (0.51) indicates the need for revision and modification of the items within this subscale. The forward-backward translation process and the confirmation of face and content validity were appropriate steps in the validation of the instrument.

Considering the importance of assessing this construct in the Iranian population, the use of this questionnaire is recommended with attention to the stated limitations and after potential revision of the second dimension. Future studies are advised to reassess reliability using the test-retest method as well as McDonald's omega coefficient. Also, examining the factor structure of the instrument in different populations and with larger sample sizes may further support its construct validity.

Limitations

The main limitation of this study was the low reliability in one of the questionnaire dimensions, which requires caution in interpreting and generalizing findings related to that subscale. Also, sample attrition (from 300 to 267 participants) and the predominance of female participants in the sample may affect the generalizability of the findings.

Acknowledgments

This study was approved by the Ethics Committee of Torbat Heydariyeh University of Medical Sciences with the ethical code IR.THUMS.REC.1396.1.

Conflict of Interest

The authors declare that there is no conflict of interest in this study.

References

1-Stamm BH. The concise ProQOL manual. Pocatello, ID: proqol. org; 2010.

2-Hooper C, Craig J, Janvrin DR, Wetsel MA, Reimels E. Compassion satisfaction, burnout, and compassion fatigue among emergency nurses compared with nurses in other selected inpatient specialties. Journal of emergency nursing. 2010;36(5):420-7. https://doi.org/10.1016/j.jen.2009.11.027

3-Najjar N, Davis LW, Beck-Coon K, Carney Doebbeling C. Compassion fatigue: A review of the research to date and relevance to cancer-care providers. Journal of Health Psychology. 2009;14(2):267-77. https://doi.org/10.1177/1359105308100211

4-Figley CR. Compassion fatigue: Psychotherapists' chronic lack of self care. Journal of clinical psychology. 2002;58(11):1433-41. https://doi.org/10.1002/jclp.10090

5-Lee D-J, Singhapakdi A, Sirgy MJ. Further validation of a need-based quality-of-work-life (QWL) measure: Evidence from marketing practitioners. Applied Research in Quality of Life. 2007;2(4):273. https://doi.org/10.1007/s11482-008-9042-x

6-Dahl S-Å, Nesheim T, Olsen KM. Quality of work: Concept and measurement. 2009. https://doi.org/10.2139/ssrn.1489881

7-Mousavi SMH, Dargahi H, Asgari M, Sharifiyan R, Shaham G, Mokhtari Z. Investigation of Staff's Job Satisfaction and Related Factors in TUMS Teaching Hospital. Journal of Payavard Salamat. 2016;10(2):133-44.

8-Joinson C. Coping with compassion fatigue. Nursing. 1992;22(4):116, 8. https://doi.org/10.1097/00152193-199204000-00035

9-Hansen N, Sverke M, Näswall K. Predicting nurse burnout from demands and resources in three acute care hospitals under different forms of ownership: A cross-sectional questionnaire survey. International Journal of Nursing Studies. 2009;46(1):96-107. https://doi.org/10.1016/j.ijnurstu.2008.08.002

10-Beck CT. Secondary traumatic stress in nurses: A systematic review. Archives of psychiatric nursing. 2011;25(1):1-10. https://doi.org/10.1016/j.apnu.2010.05.005

11-Maslach C. Job burnout: New directions in research and intervention. Current directions in psychological science. 2003;12(5):189-92. https://doi.org/10.1111/1467-8721.01258

12-Saraji GN, Dargahi H. Study of quality of work life (QWL). Iranian journal of public health. 2006;35(4):8-14.

13-Shabaninejad H, Arab M, Rashidian A, Zeraati H, Bahrami S. Quality of working life of Family Physicians in Mazandaran. Hakim Research Journal. 2012;15(2):178-84.

14-Zakerian SA, Teymouri G, AhmadNejad I, AbbasiNia M, Rahmani A, Asghari M. Evaluation and effect Dimension of quality of work life on job satisfaction in the automotive industry. Journal of Ergonomics. 2014;1(3):36-46.

15-Pashib M, Abbaspour S, Tadayyon H, KHalafi A. Quality of Professional life among Nurses of hospitals in Torbat Heydariyeh city in 2016. Journal of Tanin Salamat (Health Chimes). 2016;4(1):36-41.

16-Jensen MP. Questionnaire validation: a brief guide for readers of the research literature. The Clinical journal of pain. 2003;19(6):345-52. https://doi.org/10.1097/00002508-200311000-00002

17-Stamm BH. The ProQOL manual. Retrieved July. 2005;16:2007.

18-Sprang G, Clark JJ, Whitt-Woosley A. Compassion fatigue, compassion satisfaction, and burnout: Factors impacting a professional's quality of life. Journal of Loss and Trauma. 2007;12(3):259-80. https://doi.org/10.1080/15325020701238093

19-Lauvrud C, Nonstad K, Palmstierna T. Occurrence of post traumatic stress symptoms and their relationship to professional quality of life (ProQoL) in nursing staff at a forensic psychiatric security unit: a cross-sectional study. Health and quality of life outcomes. 2009;7(1):31. https://doi.org/10.1186/1477-7525-7-31

20-Kutner MH, Nachtsheim C, Neter J, Li W. Applied linear statistical models: McGraw-Hill Irwin; 2005.

21-Bernstein IH, Nunnally J. Psychometric theory. New York: McGraw-Hill Oliva, TA, Oliver, RL, & MacMillan, IC (1992) A catastrophe model for developing service satisfaction strategies Journal of Marketing. 1994;56:83-95. https://doi.org/10.1177/002224299205600306

22-Polit DF, Beck CT, Owen SV. Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. Research in nursing & health. 2007;30(4):459-67. https://doi.org/10.1002/nur.20199

23-Gh A, Heydarali H, Mahmoudi N. Considering practicality, reliability, validity finding of spiritual questioner in students. Psychological researches. 2009;1(3):1-23.

24-Press U. Bollen, KA Structural Equations With Latent Variables. New; 1950. p. 256.