

Validity and reliability of the Persian version of WATCH questionnaire in assessing the clinical learning environment.

Hajar Nejatdarabi¹, Mitra Amini^{2*}, Jalal Yarahmadi³, Parisa Bikineh⁴, Marzieh Kaveh⁴, Hanieh Gholampoor⁴

¹Graduate Student of Medical Education, Shiraz University of Medical Sciences, Shiraz, Iran

²Clinical Education Research Center, Shiraz University of Medical Sciences, Shiraz, Iran

³Assistant professor of Payam Noor University, Shiraz, Iran

⁴Shiraz Medical School, Shiraz University of Medical Sciences, Shiraz, Iran

Abstract

Introduction: In order to attain effective clinical education, it is necessary to assess the current training status continuously. Considering the necessity of evaluating the status of clinical education based on a valid and reliable questionnaire, this study was conducted to assess the validity and reliability of the Persian version of WATCH questionnaire in Shiraz Medical School.

Methods: This was a descriptive, analytical, cross-sectional study conducted on medical students in Shiraz Medical School, southern Iran to assess the validity and reliability of the WATCH questionnaire. The WATCH questionnaire included 15 questions. This questionnaire addresses three areas of teaching and learning, teaching skills, and learner orientation. Initially, the English version was translated into Persian. This translation was then back translated by an interpreter into English and compared to the original English version. Cronbach's alpha method was used to examine the internal consistency. Then, to evaluate the validity, content validity and construct validity were measured. Content validity was determined by medical education experts and construct validity was measured using confirmatory and exploratory factor analysis.

Results: The reliability in teaching skills was 0.97, in teaching and learning was 0.91, and in the learner centered domain was 0.86. By examining the convergence validity of the dimensions of the questionnaire, we observed that all three dimensions of the questionnaire had acceptable convergence validity. In other words, this questionnaire has a coherent validity.

Conclusion: The results of the present study showed that the Persian version of the WATCH questionnaire was a highly reliable and valid instrument for assessing the clinical environment for trainees and medical students. This questionnaire evaluates clinical education from different angles well.

Keywords: Evaluation, Clinical, Environment.

Accepted on September 20, 2018

Introduction

Clinical education can facilitate learning in the clinical setting. Clinical education provides an opportunity for the student to turn his/her theoretical knowledge into the various mental, psychological, and motor skills that are essential for patient care. Clinical education is a huge part of practical training in medical students and affiliated disciplines. With these tutorials, learners can acquire skills and prepare themselves to independently take over the responsibility of taking care of patients in the future. Identifying current issues in clinical education and modifying it, will lead to the success of an educational program and, ultimately, the improvement of the quality of medical services. This will enable medical education

practitioners to get an accurate picture of current conditions in the clinical environment and get a good perspective [1].

The goal of the medical education program is to graduate physicians who play effective roles in the health and quality of lives of people in the society. Moreover, human resources form the basis of health systems [2,3]. In order to attain effective clinical education, it is necessary to assess the current training status continuously. Its strengths and weaknesses should be identified and the quality of education should be investigated and measured according to the defined standards, visually and operationally [4].

The challenge of how to scientifically assess clinical teaching has been the topic of debate between educators and researchers

who are worried about the quality of clinical teaching in medical schools [5].

Many assessment instruments have been designed to deliver feedback to clinical teachers, department heads, and program directors. Some of them measure some aspects of clinical education and only focus on testing psychometric properties, while others measure all aspects [6-8].

One of the tools that are designed to measure the clinical education environment is the Warwick assessment instrument for clinical teaching (WATCH) questionnaire. This new tool is an acceptable, reliable and trustworthy assessment tool of clinical education in the UK [9].

Although many instruments have been created for assessing clinical teaching in Western countries, to the best of our knowledge, most of these instruments have not been validated for the Persian culture. Considering the necessity of evaluating the status of clinical education based on a valid and reliable questionnaire, this study was conducted to assess the validity and reliability of the Persian version of WATCH in Shiraz Medical School.

Method

This was a descriptive, analytical, cross-sectional study conducted on medical students in Shiraz Medical School, southern Iran to assess the validity and reliability of the WATCH questionnaire. The location of the study was educational hospitals affiliated to Shiraz medical school (Nemazee, Chamran, Rajai and Shahid Faghihi). Undergraduate medical education in Shiraz Medical School consists of 2.5 years basic sciences, 1 year pathophysiology of diseases and 3 years clinical rotation.

The study population consisted of all medical students (n=767) of Shiraz Medical School with at least 6 months of clinical education experience in clinical settings comprised of 291 year-seven, 240 year-six, and 236 year-five students. Considering the different number of students in different medical courses, random sampling with proportional volume was used. 260 medical students (94 year-seven, 86 year-six, and 80 year-five students) were selected.

The WATCH questionnaire included 15 questions. Each question, using Likert criterion, was checked with a score of 1 to 5. The best score was 5 points and the worst was rated 1. So the lowest score of this questionnaire was 15 and the highest score was 75. This questionnaire addresses three areas of teaching and learning, teaching skills, and learner orientation.

Questions 5, 7, 9, 11, 12, 13, and 14 were related to the field of teaching and learning, questions 1, 6, and 8 were related to the field of teaching skills and questions 2, 3, 4, 10, and 15 were related to the learner-centered field. The score of each of these fields comes from the sum of questions related to the same field. The lowest score was in training and learning is 7 and the highest score is 35. In the area of teaching skills, the lowest score is 3 and the highest score is 15. In the learner's area, the lowest score is 5, and the highest score is 25. The validity and reliability of this questionnaire were studied by Haider et al. in England and a Cronbach's alpha of 0.88 was reported. Structural validity also showed that 69.77% of the total variance is expressed by the subject learning, learning skills, and orientation of the students, indicating good validity and reliability [9].

Initially, the English version was translated into Persian. This translation was then back translated by an interpreter into English and compared to the original English version. After the implementation of different stages of translation, the final Persian version with 15 questions in three areas was prepared and confirmed by five faculty members of the medical education department. In addition to the 15 questions in this questionnaire, demographic data such as age, sex, and marital status were added to the questionnaire. Cronbach's alpha method was used to examine the internal consistency. Then, to evaluate the validity, content validity and construct validity were measured. Content validity was determined by medical education experts and construct validity was measured using confirmatory and exploratory factor analysis. Data were analysed using descriptive statistics such as mean and standard deviation and analytical tests such as t-test, analysis of variance, and exploratory factor analysis using SPSS software. Confirmatory factor analysis was also examined by Amos software. For assuring confidentiality, the students' names were not asked. They were also assured that the information obtained from the questionnaires would generally be expressed and the researchers will not mention their names and views, and students have no obligation to complete the questionnaires.

Results

260 medical students with a mean \pm SD age of 24.46 ± 5.4 y were questioned, of which 150 (57.7%) were women and 110 (42.3%) were men and 185 (71%) were single and 75 (29%) were married. Table 1 shows the mean, standard deviation, and the lowest and highest scores for each area as well as the total score for clinical education.

Table 1. Descriptive statistics for all areas and the whole questionnaire and assessment of clinical education status.

Field	Field name	Item	Mean \pm SD	Minimum score	Maximum score
1	Teaching skills	1, 6, 8	12.05 \pm 2.22	10	15
2	Teaching and learning	5, 7, 9, 11, 12, 13, 14	18.93 \pm 2.23	9	35
3	Learner-centered	2, 3, 4, 10, 15	20.6 \pm 5.93	16	25

As shown, it can be said that Shiraz Medical School is at an acceptable level in terms of clinical education. There were no statistically significant difference between sex and marital status of medical students and questionnaire scores.

Reliability of the WATCH questionnaire

The reliability of the questionnaire was calculated for all questions and each question based on Cronbach's alpha and Interclass Correlation Index (ICC). In the WATCH questionnaire, 40 people were allowed to retry. The ICC in teaching skills was 0.97, in teaching and learning was 0.91, and in the learner centered domain was 0.86.

Validity of the watch questionnaire

Several methods have been used to assess the validity of the Persian version of the WATCH questionnaire.

Table 2. The results of convergence and divergence of the watch questionnaire.

Questionnaire's dimension	The number of questions	The convergence range	The number of Correlations above 0.4	Convergence validity percent	Divergence validity range	Divergence validity percent
Teaching skills	3	0.88-0.96	3	100	0.2-0.38	100
Teaching and learning	7	0.77-0.84	7	100	0.18-0.34	100
Learner-centered	5	0.79-0.89	5	100	0.13-0.37	100

Pearson correlation coefficient was used to check the tool internal consistency. According to this table, all correlation coefficients between the dimensions of the WATCH questionnaire with $P < 0.001$ were meaningful. With respect to the correlation between domains, the highest correlation was found between the areas of teaching skills and learner-centered skills, which were 0.75. Two areas of learner-centered learning and teaching and learning, the lowest level of correlation were equal to 0.50.

Exploratory factor analysis

The KMO index for data was 0.92 and the Bartlett score was 1872.074 and its p-value was < 0.000 . The results indicate that the data is proper for factor analysis and samples are sufficient. To extract the number and quality of factors, using the above table, the scree plot and the main component matrix were used.

Scree plot

Scree plot diagram (Figure 1) is plotted based on the specific values of each agent in the preliminary solution and helps to identify the optimal factors. Generally, agents are selected that fall with a high slope. The factors on the low slope generally have less involvement in problem solving. As shown, the first three factors fell with a relatively high slope, and the rest of the factors are almost on the same line. Therefore, it can be concluded intuitively that these data are categorized in three areas. The Main components or agents extracted from the

Factual validity of WATCH questionnaire

To examine the factual validity, the Persian translation of the questionnaire was examined by five faculty members of the medical education group and, with the suggestion of these professors, little changes were made.

Convergence validity and divergence validity between dimensions of WATCH questionnaire

Correlation validity (correlation between dimensions of the questionnaire and questions of the same dimension) was measured by Pearson's correlation coefficient. By examining the convergence validity of the dimensions of the questionnaire, we observed that all three dimensions of the questionnaire had acceptable convergence validity. In other words, this questionnaire has a coherent validity (Table 2).

WATCH Questionnaire using the Varimax rotation are shown in Table 3.

Table 3. Main components or agents extracted from the WATCH questionnaire using the Varimax rotation.

Questions	Teaching skills area	Teaching and learning area	Learner-centered area
1 Having effective communication with students/trainees (by listening, clearly answering and reasonably explaining)	0.74	0.209	
2 Creating a passion for the trainee relative to education and learning	0.23	0.073	0.76
3 Promote or increase active involvement of trainees during learning	0.31	0.16	0.83
4 Avoid criticism and discrimination by professors and respect for justice and fairness among trainees	0.13	0.38	0.64
5 Provide consistent, explicit and constructive feedback to the instructors	0.27	0.77	0.19

6	The occurrence of appropriate professional and ethical conduct	0.890	0.22	0.28
7	Matching education with the needs of learners according to the goals of learning outcomes	0.02	0.69	0.15
8	Stimulating feedback, problem solving, self-directing, and independent learning skills	0.683	0.003	0.27
9	Having clinical competence, accurate analysis, diagnostic, cognitive and therapeutic skills	0.09	0.82	0.22
10	Having a role model for trainees	0.38	0.143	0.86
11	Sufficient clinical and medical knowledge	0.27	0.73	0.023
12	Up-to-date information and awareness of the latest scientific achievements	0.21	0.66	0.2
13	The ability to learn in a variety of situations (near the bed, the operating room, section) and involving patients in training (if needed)	0.13	0.85	0.09
14	Teaching concepts and creating well-organized skills	0.38	0.79	0.016
15	Demonstrates professional and ethical conduct	0.2	0.188	0.70
16	Eigen values	7.8	3.45	2.3
17	Cumulative predicted variance	62.54	68.31	73.85

3 factors have a special value greater than one. Therefore, exploratory factor analysis summarized the 15 questions of this questionnaire in 3 domains. These three areas account for 85.73% of the total variance.

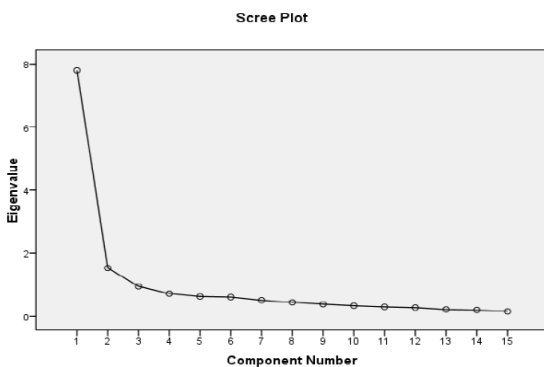


Figure 1. Scree plot diagram.

Confirmatory factor analysis

The researcher used a confirmatory factor analysis model to confirm the discovered model of exploratory factor analysis (Figure 2).

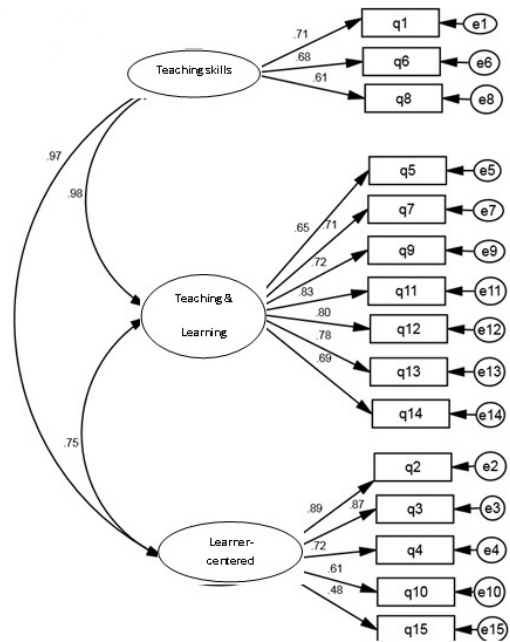


Figure 2. Confirmatory factor model.

Discussion

This study evaluated the reliability, and content and construct validity of the Persian version of WATCH questionnaire using a sample of medical students. The WATCH questionnaire is one of the specialized tools for evaluating different aspects of the clinical learning environment.

According to the pervious article about WATCH questionnaire [9], the results of this survey revealed a more positive than negative clinical education in Shiraz Medical School. In all areas and the overall score of the questionnaire, the grades were above the average indicating that Shiraz Medical School is at an acceptable level in terms of clinical education.

The results of a study about validating PHEEM questionnaire in Tehran medical school were similar to our results. It revealed a more positive than negative environment with room for improvement [10].

In order to control the quality of the questionnaires, a wide range of different issues of the questionnaire should be taken into account. One of the most important aspects of the questionnaire is the reliability of the questionnaire. Reliability is often defined in such a way that the results can be repeated in frequent measurements or length of time. The most common way to test reliability is to use Cronbach's alpha coefficient. Cronbach's alpha is often used to confirm the internal consistency of a measurement scale. In this study, the Cronbach's alpha coefficient for the whole questionnaire was 0.85. Also, this coefficient for teaching skills was 0.83, for teaching and learning field was 0.88 and for the learner-

centered learning area was 0.81, all of which were of a desirable level.

In a study conducted by Haider et al. to validate this questionnaire, it was concluded that the Cronbach's alpha coefficient for the field of teaching skills was 0.887, the field of teaching and learning was 0.883, and the learner-centered field of 0.84, which were all highly evaluated and consistent with our research. Moreover, Cronbach's Alpha coefficient for all questions of the WATCH questionnaire was 0.97, which confirmed the high reliability of this questionnaire [9].

To our knowledge, the translation, and psychometry and validation of this questionnaire were done in other language (besides English) for the first time. Therefore, no other study was found to compare the results.

Different methods were used to assess the validity of this questionnaire. In this study, first, the content validity of the questionnaire was evaluated. Validity and integrity of a test are usually determined by people specializing in the subject. As a result, in order to assess content validity of the Persian translation of the questionnaire, five faculty members of the medical education group were asked to check it. With the suggestion of these professors, the changes were made and then the questionnaire was approved by the professors. Pearson correlation coefficient was used to check the internal consistency of the instrument.

All correlation coefficients between the dimensions of the questionnaire were positive and meaningful. In correlation between domains, the highest correlation was found between the areas of teaching skills and learner-centered skills, which were 0.75. The two areas of learner-centered learning and learning and teaching had the lowest level of correlation, which was 0.50.

The results of exploratory factor analysis showed that 3 components had a special value greater than 1, which means that these 15 questions can be summarized in 3 domains. Based on the screen plot, it can be concluded that in the first three factors, the slope of the graph is sharp, but then the slope of the graph is uniform and constant. The three selective factors were able to explain 73.85% of the total variance (changes of the observations). Therefore, the complexity of analysis with 15 variables can be reduced by using the first three factors, while we lose only 26% of the information in the data.

In Haider et al.'s study, using the exploratory factor analysis, the 15 questions of the questionnaire were summarized in three factors, which determined only 72.69% of the total variance [9], which is consistent with our study. In this research questions 1, 6, and 8 were placed in the field of teaching skills, questions 5, 7, 9, 11, 12, 13, and 14 in the field of teaching and learning and questions 2, 3, 4, 10, and 15 in the learner-centered field which was completely compatible with the research about WATCH questionnaire [9].

Parvizi et al. studied the validity and reliability of the Persian version of the ACLEEM questionnaire and showed that it was a multidimensional, valid, and reliable instrument for assessing

the environment of outpatient clinics of medical residents [11]. Other clinical education assessment tools articles introduced valid and reliable instruments in this field [6-13].

One of the strengths of this study was that clinical education assessment is one of the education research priorities in Eastern Mediterranean Region and Iran [14]. Also, the integration of medical education and health system in Iran provides a good opportunity of using facilities in health care delivery system for improving education [15,16]. Some of the limitations of this study were the lack of access to all medical students due to the intensive training programs, the crowding of clinical education environments, and the crowding of patients. Finally, participants were distributed in different levels of undergraduate course from 5th to 7th years. Therefore, the findings might have been affected by the heterogeneity in their medical expertise.

Conclusion

The results of the present study showed that the Persian version of the WATCH questionnaire was a highly reliable and valid instrument for assessing the clinical environment for trainees and medical students. This questionnaire evaluates clinical education from different angles well. Therefore, the researchers recommend the use of the Persian version of the WATCH questionnaire for evaluating and assessing clinical education for trainees and medical students. The results of clinical education measurement using this tool will help educational managers and planners to monitor teaching and learning process and provide a better and more effective learning environment for the education of medical students.

Acknowledgement

This study was obtained from a thesis for obtaining Master degree in medical education that was done by the first author Hajar Nejatdarabi with proposal number 12134 that was approved by Deputy of Research of Shiraz University of Medical Sciences. The authors thank all the students who participated in this study.

References

1. Ramani S, Leinster S. AMEE Guide no. 34: Teaching in the clinical environment. *Med Teach* 2008; 30: 347-364.
2. Ramani S. Twelve tips to improve bedside teaching. *Med Teach* 2003; 25: 112-115.
3. Hirsh DA, Ogur B, Thibault GE, Cox M. Continuity as an organizing principle for clinical education reform. *N Engl J Med* 2007; 356: 858-866.
4. Fluit CR, Bolhuis S, Grol R, Ham M, Feskens R, Laan R, Wensing M. Evaluation and feedback for effective clinical teaching in postgraduate medical education: Validation of an assessment instrument incorporating the CanMEDS roles. *Med Teach* 2012; 34: 893-901.
5. Amini M, Honardar M. The view of faculties and medical students about evaluation of faculty teaching experiences. *Koomesh* 2008; 9: 171-177 (Persian).

6. Stalmeijer RE, Dolmans DH, Wolfhagen IH, Muijtjens AM, Scherpbier AJ. The development of an instrument for evaluating clinical teachers: Involving stakeholders to determine content validity. *Med Teach* 2008; 30: 272-277.
7. Kikukawa M, Stalmeijer RE, Emura S, Roff S, Scherpbier AJ. An instrument for evaluating clinical teaching in Japan: content validity and cultural sensitivity. *BMC Med Educ* 2014; 14: 17.
8. Copeland HL, Hewson MG. Developing and testing an instrument to measure the effectiveness of clinical teaching in an academic medical center. *Acad Med* 2000; 75: 161-166.
9. Haider SI, Johnson N, Thistlethwaite JE, Fagan G, Bari MF. WATCH: Warwick Assessment instrument for Clinical teaching: development and testing. *Med Teach* 2015; 37: 289-295.
10. Jalili M, Hejri SM, Ghalandari M, Moradi-Lakeh M, Mirzazadeh A, Roff S. Validating modified PHEEM questionnaire for measuring educational environment in academic emergency departments. *Arch Iran Med* 2014; 17: 372-377.
11. Parvizi MM, Amini M, Dehghani MR, Jafari P, Parvizi Z. Psychometric properties of the Persian version of the Ambulatory Care Learning Educational Environment Measure (ACLEEM) questionnaire, Shiraz, Iran. *Adv Med Educ Pract* 2016; 29: 559-566.
12. Liesel CH. Developing and testing an instrument to measure the effectiveness of clinical teaching in an academic medical center. *Acad Med* 2000; 75: 161-166.
13. Bierer SB, Hull AL. Examination of a clinical teaching effectiveness instrument used for summative faculty assessment. *Eval Health Prof* 2007; 30: 339-361.
14. Amini M, Kojuri J, Lotfi F, Karimian Z, Abadi ASH. Research priorities in medical education in the Eastern Mediterranean Region. *EMHJ* 2012; 18: 687-692.
15. Amini M, Kojuri J, Karimian Z, Lotfi F, Moghadami M, Dehghani MR. Talents for future: report of the second national medical science Olympiad in Islamic republic of Iran. *Iran Red Crescent Med J* 2011; 13: 377-381.
16. Salehi A, Harris N, Lotfi F, Hashemi N, Kojuri J, Amini M. Reform in medical and health sciences educational system: a Delphi study of faculty members views at Shiraz University of Medical Sciences. *East Mediter Health J* 2014; 20: 151-161.

***Correspondence to**

Mitra Amini

Clinical Education Research Center

Shiraz University of Medical Sciences

Shiraz

Iran