

Contextualization and validation of the interprofessional collaborator assessment rubric (ICAR) through simulation: Pilot investigation

Fatemeh Keshmiri¹, Sari Ponzer², AmirAli Sohrabpour³, Shervin Farahmand⁴
Farhad Shahi⁵, Shahram Bagheri-Hariri⁶, Kamran Soltani-Arabshahi⁷
Mandana Shirazi^{*8}, Italo Masiello⁹

Received: 30 September 2015

Accepted: 27 January 2016

Published: 1 August 2016

Abstract

Background: Simulation can be used for educating, evaluating and assessing psychometric properties of an instrument. The aim of this study was to contextualize and assess the validity and reliability of the Interprofessional Collaborative Assessment tool (ICAR) in an Iranian context using simulation.

Methods: In this descriptive study, contextualization of the ICAR was assessed through several steps. Firstly, validity assessment was approved through expert panels and Delphi rounds. Secondly, reliability assessment was done by arranging a simulation video and assessing reproducibility, test-retest (ICC), internal consistency (Cronbach's Alpha) and inter-rater reliability (Kappa). The participants included 26 experts, 27 students and 6 staff of the Standardized Simulation Office of Teheran University of Medical Sciences.

Results: Contextualization and validity of the ICAR were approved in an Iranian context. The reliability of the tool was computed to be 0.71 according to Cronbach's Alpha. The test-retest was calculated to be 0.76.

Conclusion: The Iranian ICAR can be a useful tool for evaluating interprofessional collaborative competencies. The development of the instrument through a simulation scenario has been a positive prospect for researchers.

Keywords: Interprofessional Collaboration, Interprofessional Education, Assessment, Validation.

Cite this article as: Keshmiri F, Ponzer S, Sohrabpour AA, Farahmand Sh, Shahi F, Bagheri-Hariri Sh, Soltani-Arabshahi K, Shirazi M, Masiello I. Contextualization and validation of the interprofessional collaborator assessment rubric (ICAR) through simulation: Pilot investigation. *Med J Islam Repub Iran* 2016 (1 August). Vol. 30:403.

Introduction

Learning to work interprofessionally is fundamental to patient safety and a patient-centered approach (1). According to the World Health Organization's definition, interprofessional education (IPE) is the process by which a group of students or

health care providers of different professions learn with, from and about each other (1).

In IPE programs, a competency-based approach is often used to assess the achievements and quality of interprofessional work (2-3). The components of a

¹. PhD Candidate in Medical Education, Faculty of Medicine, Tehran University of Medical Sciences, Tehran, Iran. keshmiri_f@razi.tums.ac.ir

². Professor of Department of Clinical Science and Education, Södersjukhuset, Karolinska Institutet, Stockholm, Sweden. sari.ponzer@ki.se

³. Assistant Professor, Department of Internal Medicine, Tehran University of Medical Sciences, Tehran, Iran. aasohrabpour@tums.ac.ir

⁴. Associate Professor, Department of Emergency Medicine, Imam Khomeini Hospital, Tehran University of Medical Sciences, Tehran, Iran. shfarahmand@tums.ac.ir

⁵. Assistant Professor, Department of Internal Medicine, Oncology Department, Tehran University of Medical Sciences, Tehran, Iran. fshahi@yahoo.com

⁶. Assistant Professor, Department of Emergency Medicine, Imam Khomeini Hospital, Tehran University of Medical Sciences, Tehran, Iran. hariri.shahram@gmail.com

⁷. Professor, Department of Internal Medicine, Department of Medical Education, Iran University of Medical Sciences, Tehran, Iran. soltarab34@gmail.com

⁸. **(Corresponding author)** Post-doctoral Researcher, Department of Clinical Science and Education, Södersjukhuset, Karolinska Institute, Stockholm, Sweden, & Associate Professor, Educational Development Center, Department of Medical Education, Tehran University of Medical Sciences, Tehran, Iran. mandana.shirazi@ki.se

⁹. Associate Professor, Department of Clinical Science and Education, Södersjukhuset, Karolinska Institute, Stockholm, Sweden. italo.Masiello@ki.se

competency-based assessment instrument should include objectives of assessment and its guidelines (4). Some instruments, designed to assess interprofessional competencies, have been developed and validated (5-8). In his study, Oates states that Interprofessional Collaborator Assessment Rubric (ICAR) can be used to assess knowledge/skills and behavioral acquisition (8). The aim of this study was to validate the ICAR (6,9-10) as an interprofessional performance assessment tool in an Iranian context.

The properties of the ICAR can be favorable to examiners, instructors and learners because they can increase the quality of instruction, and their functionality as a guideline with respect to the learning expectations and fostering critical reflection on performance (6). However, to assess interprofessional competencies, the researcher has to assess teamwork skills and communication in interaction with other professional members. This remains to be difficult since those assessed have to show effective collaboration and information exchange (11,12). Hence, several issues can affect the assessment, validity and reliability of a tool. Therefore, the quality assurance of a tool should be meticulously run and adapted to different contexts.

The validation process is usually adapted in a real educational context (6), and it seldom is applied in a simulation setting (13-17). Considering the challenges of assessing the reliability of interprofessional collaboration (10,13-14), the aim of this study was to contextualize and assess the validity of the ICAR in an Iranian context, and estimate the reliability of the ICAR reproducibility, internal consistency and inter-rater reliability (6) through simulation.

Methods

In this descriptive study, the ICAR was contextualized and validated from December 2011 to January 2013 (Fig. 1). The original instrument, which is based on a 4-point Likert scale, includes 31 items within six domains: Communication, Collaboration,

Roles and Responsibilities, Collaborative Patient/Client-Family Centered Approach, Team Functioning and Conflict Resolution/Management domains. Scores indicate how often specific interprofessional competencies occur; i.e., score of 1 indicates minimally, score of 2 indicates developing (occasionally), score of 3 designates competent (frequently) and score of 4 represents mastery (consistently) (6).

Study Area

This study was conducted at Tehran University of Medical Sciences in Iran, a developing country in the Middle East. In Iran, medical education systems are centralized and medical curriculums are developed by Ministry of Health. However, interprofessional education programs are not considered in the curriculum. Based on the authors' Knowledge, interprofessional education and collaboration are not considered in the curriculum of Middle Eastern countries. In Iran, running IPE is not an easy task due to the cultural boundaries to the doctor-center discipline. Running IPE in Iran needs some preparation of context; therefore, this study, which contextualized tools and developed educational materials, may be used as a foundation for establishing new IPE system in Iran. This study was conducted to prepare the infrastructure of education and evaluate interprofessional education in Iran.

Participants

The total number of participants in this study during the contextualization and validation phases were 59; of whom, 26 were experts, 27 students, two simulation educators, two standardized patients and two were film producers. In this study, experts were individuals who had at least 10 years of experience in clinical education, non-technical skills training (such as communication and teamwork) and were familiar with IPE.

The process is described in Figure 1.

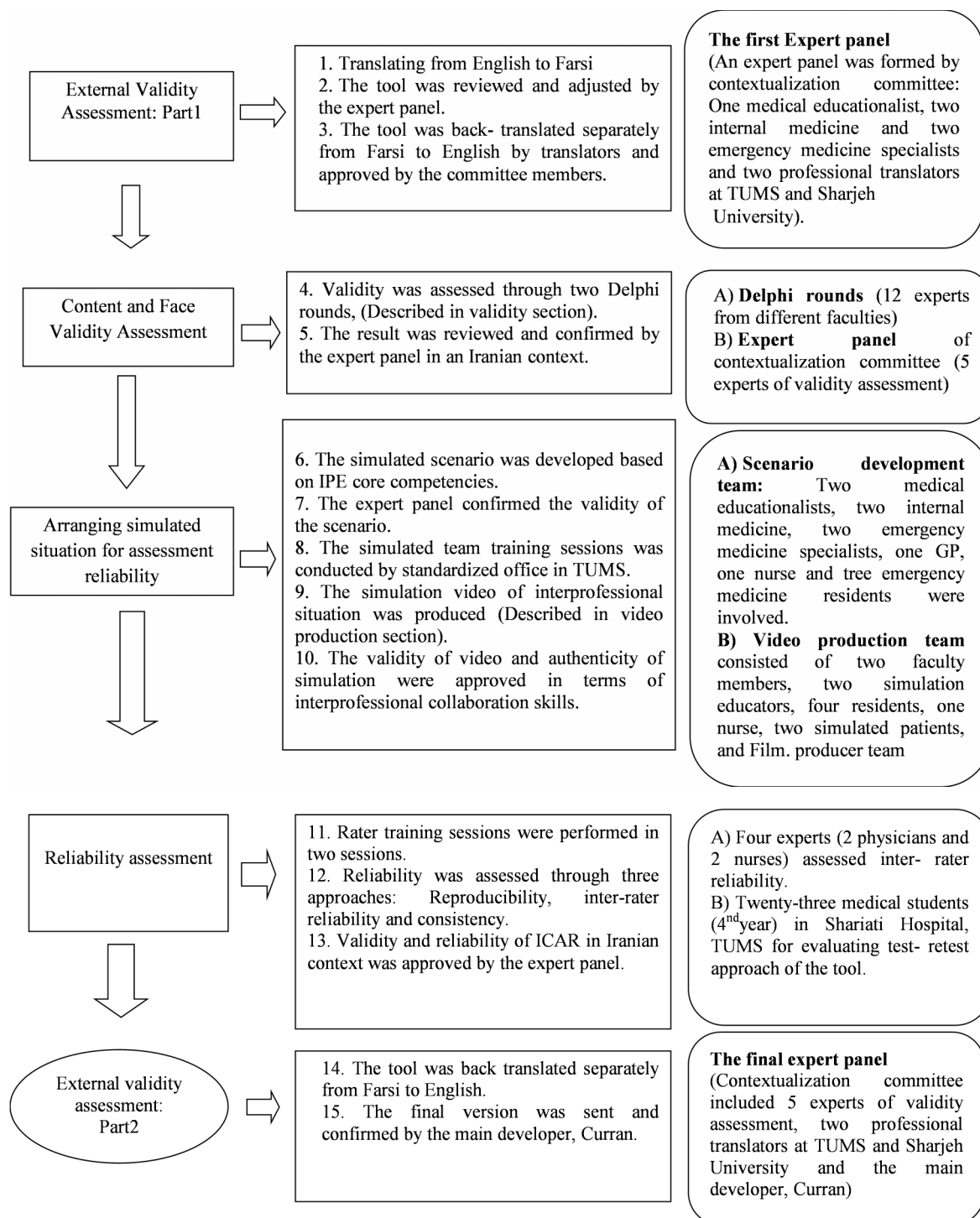


Fig. 1. Contextualization and Validation of ICAR in Iranian Context: Process

Study Design

Contextualization of the instrument: The process was based on the “Toolkit on Translating and Adapting Instruments” (18). The instrument was adapted to the Iranian context through the following process:

Validity Assessment of the ICAR, Part 1: English to Farsi: The expert panel confirmed the external validity of the tool following the translation from English to Farsi and through applying the guidelines of the Human Sciences Research Institute (HSRI) (18). Content and face validity were as-

sessed through two Delphi rounds. In the first round, the translated tool was sent to the experts (response rate was 80%). The results of the first round were gathered after seven days. Then, the experts' opinions were analyzed through the content analysis approach. Two items were suggested to be added in the domains of collaboration and challenge management, which were related to problem solving and providing feedback.

The suggested items were approved in the second Delphi round. The agreement between experts was higher than 90%, and no further suggestions were made. Hence, in the final version, the total number of items was 33 within six domains. The total maximum score was 99, and the results were reviewed and confirmed by an expert panel in the Iranian context.

Simulated Scenario Development and Video Production: The reliability of the translated instrument was confirmed through a simulation scenario:

1. The simulated scenario was developed based on the objectives of the ICAR and the core competencies for interprofessional collaborative practice (3,6,19). The scenario was about a case of multiple traumas in an emergency ward with emphasis on interprofessional collaborative competency in the health care team. The scenario was developed during 18 sessions (36 hours).

2. The expert panel members confirmed the validity of the scenario to increase the scenario reality.

3. The Standardized Simulation Office at TUMS conducted the simulated team training sessions to produce a video. The simulation training sessions were conducted and all the simulated team members exercised their role-play. Simulation educators watched their final exercise, and confirmed quality of their role-play.

4. The simulation situation of the emergency ward was conducted at TUMS skill lab.

5. The simulated scenario was recorded and a video was produced. The duration of the video was 30 minutes, and its production took 12 hours of teamwork.

6. The validity of the video and authenticity of the simulation were approved in terms of interprofessional collaboration skills.

Reliability Assessment of the ICAR: In this study, the reliability of the tool was assessed by performing reproducibility (test-retest), internal consistency and inter-rater reliability analyses. To do so, two rater-training sessions were provided. In the first session, raters gained familiarity with the reliability assessment process, the ICAR and its guideline. Then, the raters watched the video produced in the earlier phase and completed the Iranians' ICAR. Afterwards, they received feedback from the main investigator of the research (M SH).

The test-retest approach was done by the raters filling out the ICAR with one-week interval between the first and second occasions (20). Internal consistency was calculated per domain through Cronbach's Alpha. The agreement between the 23 raters and experts' score was assessed through correlation (21).

Inter-rater reliability was approved by assessing the correlation between four expert raters after watching the video and applying non-parametric analysis tests (Kappa coefficient) (22,23).

External Validity Assessment, Part 2: Farsi to English: After demonstrating the validity and reliability of the ICAR in an Iranian context, the tool was back translated separately from Farsi to English. The translated version was sent to the main developer, Professor Curran, who confirmed the consistency of the final ICAR. The reproducibility of the ICAR was assessed by ICC analysis, the most appropriate and commonly utilized test for assessing parametric data (21,24,25). In this study, the test-retest approach was measured by the ICC, internal consistency was measured by Cronbach's Alpha and inter-rater reliability was assessed by Kappa coefficient. The agreement between the raters and experts' scores was assessed using descriptive statistical analysis. SPSS Version 16 was used

Table 1. Psychometric Properties of the Inter-professional Collaborator Assessment Rubric (ICAR)

ICAR domains	Test- Retest: ICC	Internal Consistency: Cronbach's Alpha	Agreement between raters' scores (23) and experts' score (%)
Communication	0.73	0.65	75.12
Roles and responsibilities	0.80	0.71	72.65
Collaboration	0.75	0.74	84.34
Collaborative patient centered approach	0.75	0.71	77.17
Team function	0.83	0.75	74.78
Conflict management	0.75	0.75	67.80
ICAR Domains	0.76	0.71	75.31

for data analysis.

Ethical Considerations

The Ethical Committee of Teheran University of Medical Sciences (ID: 91-01-30-17052) approved this study. Prof. Curran gave permission to use the ICAR instrument. Inform consent was taken for all the raters who participated in this study.

Results

Demographic Characteristics of the Study Participants

The mean±SD age of the expert panel members was 43.5±6.5yrs., and their mean±SD work experience was 10±8.5yrs. In this study, 50% (n=6) of the participants of Delphi rounds were men. The participants' mean±SD age was 41.5±8.5 yrs., and their mean±SD of work experience was 12±2.5yrs. The mean (Mean±(SD)) age of the video production committee was 37.5±9.6yrs., and their mean±SD of work experience was 10±8.5yrs. The mean±SD age of the four raters in inter-rater reliability assessment was 33.2±6.9yrs. The major-

ity of the raters in reproducibility assessment were female (52%) and their mean±SD age was 28.4±2yrs.

Validation of the ICAR

The content and face validity of the Iranian ICAR was approved through two Delphi rounds. The validation of scenario and video production was approved through consensus.

Psychometric Properties

The internal consistency according to Cronbach's Alpha was 0.71, and the test-retest reliability was 0.76 (Table 1).

The highest test-retest reliability was in the team function domain (ICC=0.83), and the lowest in the domain of communication (ICC=0.73). The agreement between the raters and the experts' score was calculated to be 67.8 and 84.3 (Table 1). The inter-rater reliability was calculated by kappa coefficient and turned out to be K=0.7 (Table 2).

Discussion

Measuring interprofessional collaboration

Table 2. Inter-rater Reliability between the Raters

Raters	Test
1 vs 2	Kappa=0.7 P<0.001
1 vs 3	Kappa=0.71 P<0.001
1 vs 4	Kappa=0.72 P<0.001
2 vs 3	Kappa=0.72 P<0.001
2 vs 4	Kappa=0.71 P<0.001
3 vs 4	Kappa=0.78 P<0.001
1,2 ,3 and 4	Intra class coefficient (ICC)=0.87 P<0.001

competencies is still a challenge; therefore, applying a valid and reliable instrument remains a significant issue. The findings of this study confirmed the innovative validation process by endorsing the reliability of the ICAR within a simulated setting. In this study, several professionals were engaged in the research to be in line with the assumption of multidisciplinary attribution of IPE.

In our study, the tools' validity was strongly agreed upon by applying the accepted HSRI's guideline (18), Delphi rounds, an expert panel, and receiving feedback from the instrument's main author (6), which was in line with the research procedure of other studies (26,27). We followed the research process by Curran et al. (6) and applied two rounds to compile the instrument in a rigorous fashion just as other researchers have done (28).

We used simulation to assess the psychometric properties (reliability) of the instrument in the interprofessional education domain. The controlled setting allowed the researchers to assess the interprofessional skills by controlling for confounding factors and possible Hawthorne effect that may occur in a real setting. In their study, Hayward et al. (10) assessed the reliability of another interprofessional performance tool in a real context, but despite the extended observation, they were not able to assess all the domains of their instrument. In this study, to avoid the challenges, reproducibility was assessed within a simulated situation. Shirazi et al. (29) conducted a research to evaluate the reliability of standardized patients (SP) and to assess communication skills as opposed to standard raters. Despite the limited number of assessors, which was a limitation, the researchers found an acceptable correlation among all raters (29). In this study, to increase the statistical power and obviate potential biases, the number of raters was increased to 23.

The highest internal consistency was in the domain of team function and challenge management. The lowest internal con-

sistency was computed in the communication domain. This could be interpreted as an inconsistency between items because of different issues such as 'respect', which are more subjective and more dependent on raters' perceptions and experience. However, in accordance with this latest finding, studies from Shyne et al. (30) and Najafi et al. (31) found similar low internal consistency with the Team STEPPS Teamwork Attitudes Questionnaire (T-TAQ) and SDOT, respectively. In contrast with our findings, Hayward et al. study (10) found internal consistency of communication to have the highest domain. This may be due to its use of a multi-source feedback (MSF) process to assess collaboration competencies (10).

The highest agreement between the 23 raters and the experts' scores was in the field of collaboration (84.3%), which could be related to the raters' familiarity with these concepts and gaining experience in the group work. Moreover, the items' construct in this domain consisted of simplicity, clarity and relevance of items for the raters. The lowest accordance was in the domain of conflict management (67.8%). These findings may be due to the possible limited involvement of the 4th-year medical students in health care management issues or lack of training. However, this information could not be confirmed.

The correlation between experts' rating was approved by non-parametric statistical tests. The Kappa coefficient (32) demonstrated propriety agreement among each pair raters and between all raters' (four) scores. In line with our study, findings from Schmitz demonstrated acceptable or highly acceptable coefficients (33). However, in contrast to our findings, Boulet et al. found a low correlation among their raters (0.09 and 0.29), which can be due to the lack of training or an assessment guide in their study (34). In this study, to inhibit the possible low inter-rater reliability among the raters, the research group compiled a guideline for reliability assessment and conducted training sessions.

Limitations

The validity assessment of an interprofessional assessment tool in a real setting is a difficult task as there are many confounders. Moreover, In Iran, we do not have interprofessional wards where we could run such studies. Therefore, in this study, we used a video of a simulated setting to assess the ICAR psychometric properties. Although this was a psychometric study, one of the limitations was the reproducibility assessment of the instrument by medical students due to the restriction in the implementation of the interprofessional meetings. Limited generalizability across institutions, countries and even within the setting of an institution (i.e., beyond the trauma team) might be a limitation.

Conclusion

Assessing IPE is a complex task due to the difficult task of assessing team competencies. This study confirmed the validity and reliability of the modified ICAR to evaluate interprofessional collaborative competencies in an Iranian context, which was done using a novel process of using a simulated scenario. IPE in an Iranian context is a new endeavor and we recommend conducting more research on this topic.

Acknowledgements

Sincere thanks go to Professor Curran for his comments and kind contribution and also to the medical students at TUMS who took part in this research.

References

1. WHO: World Health Organization. Framework for action on interprofessional education & collaborative practice. Geneva: World Health Organization. April 2012, Available from: <http://whqlibdoc.who.int> . 2010.
2. Davis M, Harden R. Competency-based assessment: making it a reality. *MedTeach* 2003; 25(6):565-568
3. Canadian Interprofessional Health Collaborative C: A National Interprofessional Competency Framework. May 2010. Available from: www.cihc.ca.
4. Harden M. Outcome-Based Education: the future is today. *MedTeach* 2007;29(7):625-9.
5. Canadian Interprofessional Health Collaborative C: An Inventory of Quantitative Tools Measuring Interprofessional Education and Collaborative Practice Outcomes. May 2012. Available from: www.cihc.ca.accessed
6. Curran V, Hollett A, Casimiro L, McCarthy P, Banfield V, Hall P, et al. Development and validation of the interprofessional collaborator assessment rubric (ICAR). *J Interprof Care* 2011; 25:339-344.
7. Thannhauser J, Russell-Mayhew S, Scott C. Measures of interprofessional education and collaboration. *J Interprof Care* 2010;24(4):336-349.
8. Oates M, Davidson M. A critical appraisal of instruments to measure outcomes of interprofessional education. *Med Educ* 2015;49:386-398.
9. Curran V, Casimiro L, Banfield V, Hall P, Lackie K, Simmons B, et al. Research for Interprofessional Competency-Based Evaluation (RICE). *J Interprof Care* 2009;23(3):297-300.
10. Hayward MF, Curran V, Curtis B, Schulz H, Murphy S. Reliability of the Interprofessional Collaborator Assessment Rubric (ICAR) in Multi Source Feedback (MSF) with post-graduate medical residents. *BMC Med Educ* 2014;14:1-9.
11. Joshi R, Ling FW, Jaeger J. Assessment of a 360-degree instrument to evaluate residents' competency in interpersonal and communication skills. *Acad Med* 2004;79(5):458-463.
12. Musial JL, Rubinfeld IS, Parker AO, Reickert CA, Adams SA, Rao S, et al. Developing a scoring rubric for resident research presentations: a pilot study. *J Surg Res* 2007;142(2):304-307.
13. McEvoy M, Hand W, Furse C, Field LC, Clark CA, Moitra VK, et al. Validity and Reliability Assessment of Detailed Scoring Checklists for Use during Perioperative Emergency Simulation Training. *Simul Healthc* 2014;9(5):295-303.
14. Wright M, Segall N, Hobbs G, Phillips-Bute B, Maynard L, Taekman JM. Standardized Assessment for Evaluation of Team Skills Validity and Feasibility. *Simul Healthc* 2013;8(5):292-303.
15. Rethans JJ, Gorter S, Bokken L, Morrison L. Unannounced standardised patients in real practice: a systematic literature review. *Med Educ* 2007; 41(6):537-549.
16. Cox M, Irby DM, Epstein RM. Assessment in medical education. *N Engl J Med* 2007;356(4):387-396.
17. Gorter S, Rethans JJ, Scherpbier A, De'sire'e H, Houben H, Vleuten C, et al. Developing case specific checklists for standardized-patient—based assessments in internal medicine: A review of the literature. *Acad Med* 2000;75(11):1130-1137.
18. Chávez L, Canino G. Toolkit on Translating and Adapting Instruments. Human Services Research Institute. may 2012, Available from: www.tecathsri.org.

19. Interprofessional Education Collaborative (IPEC): Core competencies for interprofessional collaborative practice: Report of an expert panel. Washington, D.C. May 2011. Available from: www.aacn.nche.edu
20. Norman GR, van der Vleuten CPM, Newble DI. Psychometric Method In J. A. Shea and S. Fortnag (eds), International Handbook of Research in Medical Education Series. (Vol. 7): Springer International Handbooks of Education 2002:pp.99.
21. Terwee C, Bota S, Boera M, 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.
22. Keen AJA, Klein S, Alexander D. Assessing the Communication Skills of Doctors in Training: Reliability and Sources of Error. *Adv Health Sci Educ* 2003;8(1):5-16.
23. Xu S, Lorber MF. Interrater agreement statistics with skewed data: evaluation of alternatives to Cohen's kappa. *J Consult Clin Psychol* 2014;826(6):1219-27.
24. Zaki R, Bulgiba A, Nordin N, Ismail NA. A Systematic Review of Statistical Methods Used to Test for Reliability of Medical Instruments Measuring Continuous Variables. *Iran J Basic Med Sci* 2013;16(6):803-807.
25. Al-shair K, Kolsum U, Berry P, Smith J, Caress A, Singh D. Development, dimensions, reliability and validity of the novel Manchester COPD fatigue scale. *Int J Chron Obstruct Pulmon* 2009;64:950-955.
26. Tomasik T. Reliability and validity of the Delphi method in guideline development for family physicians. *Qual Prim Care* 2010;18(5):317-26.
27. Schlegel C, Woermann U, Rethans J. Validity evidence and reliability of a simulated patient feedback instrument. *BMC Med Educ* 2012;12:1-5.
28. Dizon JM, Grimmer-Somers K, Kumar S. The physical therapy profile questionnaire (PTPQ): development, validation and pilot testing. *BMC Res Notes* 2011;19(4):362-368.
29. Shirazi M, Labaf A, Monjazbi F, Jalili M, Mirzazadeh M, Ponzer S, et al. Assessing medical students' communication skills by the use of standardized Patients: emphasizing standardized patients' quality assurance. *Acad Psych* 2014; 38(3):354-360.
30. Shayne P, Gallahue F, Rinnert T. Reliability of a Core Competency Checklist Assessment in the Emergency Department: The Standardized Direct Observation Assessment Tool. *Acad Emerg Med* 2006;13(7):727-732.
31. Najafi M, Keshmiri F, Najafi M, Shirazi M. Assessment of Validity and Reliability of Team STEPPS Teamwork Attitudes Questionnaire (T-TAQ) in Iran. *Payavard Salamat* 2013;7(5):389-398.
32. Rothman AI, Cusimano M. A comparison of physician examiners', standardized patients', and communication experts' ratings of international medical graduates' English proficiency. *Acad Med* 2000;75(12):1206-11.
33. Schmitz CC, Chipman J, Luxenberg M, Beilman GJ. Professionalism and Communication in the Intensive Care Unit Reliability and Validity of a Simulated Family Conference. *Simul Healthc* 2008; 3(4):224-38.
34. Boulet JR, van Zanten M, de Champlain A, Hawkins RE, Peitzman SJ. Checklist Content on a Standardized Patient Assessment: An Ex Post Facto Review. *Adv Health Sci Educ* 2008;13(1):59-69.