



Validity and Reliability of Combination of Observed Practice with Objective Structured Clinical Examination (OSCE) in Midwifery Final Exam

**Mahbobeh Faramarzi¹, Hajar Salmalian^{2*}, Hajar Pasha²
and Zahra Banihosseini²**

¹*Fatemeh Zahra Infertility and Reproductive Health Research Center, Babol University of Medical Sciences, Babol, Iran.*

²*Fatemeh Zahra Infertility and Reproductive Health Research Center, Department of Midwifery, Babol University of Medical Sciences, Babol, Iran.*

Authors' contributions

This work was carried out in collaboration between all authors. Author MF designed the study and wrote the protocol. Author wrote the first draft of the manuscript. Authors HS and HP managed the analyses of the study. Author ZB managed the literature searches. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/ARRB/2015/12775

Editor(s):

(1) George Perry, Dean and Professor of Biology, University of Texas at San Antonio, USA.

Reviewers:

(1) Patricia Katowa-Mukwato, Department of Medical Education Development and Department of Nursing Sciences, University of Zambia, Zambia.

(2) Anonymous, Pusan National University, Korea.

Complete Peer review History: <http://www.sciencedomain.org/review-history.php?iid=667&id=32&aid=6094>

Original Research Article

Received 17th July 2014
Accepted 14th August 2014
Published 13th September 2014

ABSTRACT

Objective: The assessment of midwifery competency in final examination is the one of the most important tests in midwifery educational course of students in bachelor's degree. As objective structured clinical evaluation (OSCE) has become popular and is now part of the midwifery competency assessment, the current study was performed to determine the reliability and validity of the combination of observed practice with OSCE in the final examination of midwifery students.

Methods: A modified final examination model, combination of OSCE and observed practice, was planned in Midwifery Department of Babol University of Medical Sciences. Twenty-seven midwifery students who had passed all midwifery subjects participated in the study in May 2012. The OSCE examination consisted of 8 stations over 6 minutes each. Observed practice consisted of the

*Corresponding author: Email: salmalian2005@yahoo.com;

observation of midwifery students' competency in clinical settings; labor, delivery, and post-partum. **Results:** The correlation coefficient between grade point average (GPA) with OSCE scores, observed practice scores, and modified exam model were 0.70($P<0.001$), 0.61($P<0.01$), 0.66 ($P<0.01$), respectively.

The correlation coefficient between the scores of two observers was significant for all stations and units. Cronbachs' alpha ranged from 0.35 to 0.94. The correlation coefficient between the scores of two observers for modified exam model was 0.59 ($P=0.002$). The exam showed a high internal consistency and Cronbach' alpha coefficient for all midwifery OSCE exams was 0.79.

Conclusion: The combination of observed practice with OSCE examination is a reliable method in assessing midwifery students' competence in final exam. Also, the modified final examination model is an acceptable alternative for traditional final examination.

Keywords: Validity; reliability; observed practice, objective structured clinical examination, final examination; midwifery competency.

1. INTRODUCTION

Examination is an important part of a teaching-learning process which allows instructors to evaluate their students during and at the end of an educational course. Clinical evaluation as an important pillar of midwifery education must incorporate measuring competencies in midwifery. Therefore, it is necessary to use validated evaluation for the assessment of different aspects of midwifery functions. In Iran, to graduate as a midwife, all students fulfilling a bachelor's degree in must midwifery pass the final examination. This assessment of midwifery competency is one the most important tests in midwifery education. A variety of assessment strategies, including clinical interview, observed practice, written examination papers, and objective structured clinical examination (OSCE) are used to assess midwifery students' clinical competence. In Iran, the most common form of final exam is the traditional oral clinical examination (TOCE) in four fields; labor and delivery care, prenatal care, gynecological care, maternal and child health care.

The assessment of midwifery students using the TOCE has been criticized by being highly subjective, unstructured, and biased [1]. In TOCE, often the student is questioned only regarding his final conclusion. The ability to examine a patient and arriving at that conclusion is not observed by the examiners. Also, there is no significant feedback to the candidate [2,3]. Also in TOCE, patients soon become uncooperative and fairness in the judgment becomes least to say, difficult. Not only the patients are disturbed by the repeated examinations, students are also under repeated stress, and the examination is resource intensive [3,4]. The need for another form of examination

that reduces the problems of TOCE in final examination has been proposed in a recent research [5].

The use of the OSCE would circumvent the disadvantages of the TOCE. The OSCE is, however, costly and time-consuming particularly if used for the assessment of large numbers of students. In a study in Iran, they compared the student' satisfaction in OSCE and traditional final examination. This study showed that the traditional final examination is not suitable tool in the evaluation of clinical competence in midwifery [5]. Also, Faramarzi et al. [6] reported that OSCE exam was superior to traditional oral exam in the reduction of test anxiety in the final examination of Midwifery students.

OSCE was first introduced by Harden¹ in 1975 [7]. It is designed to test clinical skill performance and competence in skills. An OSCE usually comprises a circuit of short (the usual is 5–10 minutes although some use up to 15 minute) stations, in which each candidate is examined on a one-to-one basis with one or two impartial examiner(s) and either real or simulated patients. Each station has a different examiner, as opposed to the traditional method of clinical examinations where a candidate would be assigned to an examiner for the entire examination. Candidates rotate through the stations, completing all the stations on their circuit. In this way, all candidates take the same stations. It is considered to be an improvement over traditional examination methods because the stations can be standardized enabling fairer peer comparison and complex procedures can be assessed without endangering patient's health [7-9].

The concept of validity of a measurement tool is considered to be the degree to which the tool measures what it claims to measure [10]. Reliability is the extent to which a measurement gives results that are consistent [11]. Previous studies showed that OSCE is strongly recommended as a valid and reliable means of evaluating health professional students' clinical skills. Nickbakht et al. reported good validity and reliability for OSCE in evaluating the clinical competence of audiology students [12]. Smith et al. reported that OSCE is a valuable strategy for enhancing the assessment of students' clinical competence, and for embracing diversity within midwifery education and training [13]. Moattari et al. [14] introduced OSCE as a reliable and valid means of evaluating nursing students' clinical skills.

Despite the general acceptance of OSCE method, there is a debate over the value of OSCE testing compared to more traditional methods. A review of reliability and validity research does not clearly show superiority of OSCE testing [15]. Also, a lot of research emphasized that for a comprehensive assessment of clinical competence, other methods of assessments should be used in conjunction with the OSCE [16-18]. Therefore in this study, modified final examination, combination of OSCE and observed practice, was planned.

The use of OSCE has progressively increased. There have been numerous studies using OSCE to validate education situations such as medical licensing, resident and student education, and technical skills [19-21]. In the midwifery literature, few studies have reported the use of OSCE in midwifery skills. Because the OSCE has great scope to adequately test students' competency and has not been conducted in combination with observe a practice in Iranian midwifery final examination, this study was conducted to assess the validity and reliability OSCE in modified final test of midwifery students.

2. MATERIALS AND METHODS

2.1 Study Design and Participants

This study is an observational study to determine the validity and reliability of OSCE examination for midwifery students in final examination. Twenty-seven female midwifery students who had passed all midwifery subjects participated in the study in May 2012. Six months before the

final exam, an examination coordinating committee was established with members who are committed to the evaluative and educational process to develop and OSCE implement. The members in this committee were the teaching staff members of Nursing and Midwifery schools of Babol University of Medical Sciences. The committee members participated in workshops to train using the OSCE assessment model. In series of 5 meeting, the staff agreed about the planning modified final examination, the stations, and the checklists.

2.2 Data Collection

The modified final examination consisted of 2 parts. The first part was observed practice-observation of midwifery students' competency in clinical settings. This clinical exam differed from traditional final examination. Observed practice consisted of three separate exercises in which midwives were observed by two faculties of Midwifery Department as they conduct a focused or complete history, perform a focused or complete physical examination, or present a patient's case orally [22]. We designed assessment forms customized to the goals of our exam. There was a separate form for interviewing, physical examination, and oral case presentation. Each of the three assessment forms (interviewing, physical examination and oral presentation) was composed of a checklist of core skills and a global rating scale to assess the overall observed performance.

The second part was the OSCE examination. The OSCE examination consisted of 8 stations over 6 minutes each. One minute was given between stations to facilitate change and reading of instructions. On the door of the station room was an instruction sheet of expected of the station, which consisted of three parts: the goal of the station, the intended task, and the examiners' way of assessment. After going through four stations, the examinees took a short break. All students completed the circuit around a period of 48 minutes.

The aim of each station was to test a particular clinical competence. Students' skills in the management of gynecological problems were assessed in stations 1, 2, and 4. In stations of 3 and 6 students' skills in maternal and child health care was evaluated. A standardized patient in station 3 was selected among the expert midwives employing in the family planning unit of health center. She was trained for her

participation in the study. The management of abnormal labor was assessed in station 5. Students' skills in labor management and prenatal care were evaluated with two cases and laboratory tests. The tasks that involved the 8 stations are listed in Table 1.

Two teaching staff members observed the students' performance in the stations and were registered on a checklist sheet. There was no relationship between two observers or examiners. All the necessary instruments and instructions for student's performed requested skills and their performances were assessed by criterion reference for each station. The criterion based scoring was used with each checklist item. The observer filled the check lists based on related rating scale, each checklist item was scored as 0 (omitted step), 1 (incorrect or inadequate), and 2(correct or adequate). All aspects of this protocol were approved by the

Medical Ethics Committee of Babol University of Medical Sciences.

2.3 Statistical Analysis

To assess the reliability and validity of the OSCE, the statistical analyses of Cronbach alpha and Pearson's correlation coefficient were used. Concurrent validity is measured by comparing one testing method to another that aimed to measure the same construct.

For the construct validity of the test, the correlation coefficient between the total mark of midwifery OSCE and grade point average was calculated. For internal consistency of the test, the correlation coefficient between the mean marks of each station with total mark all stations (total mark OSCE) was calculated. Correlation coefficient between marks of two observers in each station was used in the reliability of OSCE exam.

Table 1. Midwifery objective structured clinical examination stations in students' final examination

Stations	Assessed skill
1	Bimanual examination of uterus: put the fingers the full length the vagina, palpate the uterus for size, shape, location, position, consistency, and mobility (a mannequin with 8 weeks uterus size)
2	Obtaining a specimen for Papanicolaou: Insert the appropriate size speculum, visualize the cervix and fix the speculum. Insert a cervical brush and rotate it, spread the cells on the labeled slide and fix it (a mannequin with equipments for the work).
3	Evaluation of growth child: Drawing height growth and diagnosing height growth disorder and present methods for solving (a growth height chart cure of child with 8 months age)
4	Evaluation and management of cervicitis: Diagnosis of endocervicit discharge, erosion in the cervix, description of adequate drug for treatment (A slide with cervicitis)
5	Management of abnormal delivery: Diagnosis of shoulder dystocia and choice of adequate maneuvers for management (A slide with shoulder dystocia and treatments)
6	Ability maternal education for complementary nutrition of child: conduct of interview with the time frame, adequate advances for complementary nutrition (a simulated mother with 6 month old child).
7	Interpretation of screening test in pregnancy with abnormal fetus: Diagnosis of abnormal markers in abnormal fetus and making the decision for referring (report a laboratory test with diagnosis of Down syndrome)
8	Management of high risk pregnancy: Diagnosis of preeclampsia, severity, prevention of eclampsia with administration of drug (a report with laboratory test of a woman diagnosed with severe preeclampsia)

3. RESULTS

Twenty-seven female midwifery students with a mean age of 23.81 years (SD 0.73) participated in final examination.

To assess construct validity of modified exam model, observed practice OSCE, and examinees' total mark of modified exam was compared with grade point average at graduation stage. Correlation coefficient between grade point average (at graduation stage) with OSCE scores, observed practice scores, and modified exam model were 0.70($P<0.001$), 0.61($P<0.01$), 0.66 ($P<0.01$), respectively.

The correlation coefficient between the scores of two observers was significant for all stations and units (Table 2). Cronbachs' alpha ranged from 0.35 to 0.94. The correlation coefficient between the scores of two observers for modified exam model was 0.59 ($P=0.002$).

The scores of modified model were calculated through the sum of scores of total mark of 8 stations of OSCE with a total mark of 3 clinical units of observed practice. Internal consistency of the modified model was calculated through correlation coefficient between the mark of each station with total mark of 8 stations, and the mark of each unit with total mark of 3 clinical units. The exam showed a high internal consistency and Cronbach' alpha coefficient for all the midwifery OSCE exams was 0.79. Table 2 shows the internal consistency for each station and clinical unit.

4. DISCUSSION

This was the first midwifery modified final examination, the combination of OSCE and observed practice in Iran, and it yielded remarkable reliability and validity. Reliability for each station and unit was significant. Internal consistency of all the stations and clinical units was acceptable. The lowest score belonged to stations 7 and 8. There were a high degree of controversy regarding the design of the questions and the checklists of these stations which seem to explain the low score with lack of consensus among the designers. Low alpha Cronbach in these stations reflected a diversity of objective for two cases. Also, the lowest internal consistency belonged to postpartum unit.

High correlation coefficient between OSCE scores and grade point average of students (0.70) showed that our study had a good face

and content validity. Also, there was a high correlation coefficient between modified exam model and grade point average of students (0.66). Like this results, some research showed that for a comprehensive assessment of clinical competence, observed practice should be used in conjunction with the OSCE [16-18]. Similar to our study, Auewarakul et al. [23] found OSCEs to be one of the evaluation methods with the most valid evidence and concluded that "There is clearly sufficient evidence of validity to support the utilization of the OSCE format". In contrast, a study reported that performance on a multiple choice writ-ten test was a better predictor of clinical performance of family physicians than the use of unannounced standardized patients [24].

Although the use of OSCEs is acknowledged as effective assessment tool and is seen as the gold standard for evaluating clinical performance [19], no study has previously reported the reliability and validity of OSCE in Iranian midwifery students' final examination. Also, few studies showed the implication of OSCE in the field of midwifery in Iran. Aghajani et al. reported that The OSCE may be a powerful assessment in the evaluation of clinical competence in midwifery [5]. Erfanian and Khadivzadeh in the evaluation of midwifery students' competency in providing intrauterine device services using OSCE reported that OSCE is a valid evaluation method which provides valuable information which cannot be obtained by more traditional assessment modalities [25]. A study suggested that midwifery students perceive OSCE as a valid means of assessment and increase confidence in performing clinical skills [8]. Smith et.al explored the validity and reliability of OSCE for assessing midwifery students' competence in lactation and infant feeding practices. They reported that OSCE, when viewed alongside other forms of assessment, and with the relevance to the topic under examination, may be considered a valuable strategy for enhancing the assessment of students' clinical competence, and for embracing diversity within midwifery education and training [13].

How did our study obtained a reasonably high content validity and reliability test? The high content and concurrent validity in the use of OSCEs is obtainable with the application of a "test blueprint" (a framework for content areas of the test) [26]. Well-designed checklists and performance will clearly influence the validity [27]. High reliability of OSCE is influenced by large numbers of stations,

Table 2. Reliability and internal consistency of modified final examination in midwifery education

Parts of examination		Correlation between two observers**		Correlation of station or unit with total mark***	
		Cronbach's	P- value	Cronbach's	P-value
OSCE stations	1	0.66	<0.001	0.77	<0.001
	2	0.45	0.02	0.52	0.003
	3	0.89	<0.001	0.61	<0.001
	4	0.94	<0.001	0.54	0.02
	5	0.76	<0.001	0.51	0.03
	6	0.82	<0.001	0.53	0.04
	7	0.35	0.04	0.34	0.03
	8	0.57	0.003	0.31	0.04
Observed* Practice units	1	0.75	<0.001	0.90	<0.001
	2	0.67	0.01	0.83	<0.001
	3	0.62	0.001	0.82	<0.001
Modified model		0.59	0.002	-	-

* Observed practice units: 1 labor, 2 delivery, 3 post-partum **Reliability of OSCE exam was assessed with correlation coefficient between the mark of two observers in each station, unit, and total marks *** Internal consistency of OSCE exam was assessed with correlation coefficient between the mark of each station with the total mark of 8 stations. The internal consistency of observed practice exam was assessed with correlation coefficient between each unit with total mark of units

good standardization of patients, test length 1-2 hours, controlling confider variables including, student fatigue, personal bias, and high anxiety about the testing method [28-31]. Some reasons that contributed in the high content validity in this study were: establishing the coordinating committee with expert faculty midwives, long hours of meetings and discussion for planning the study and stations, developing appropriate checklists according to midwifery references, and implication of related workshops for OSCE training. Also, the reasons that increased reliability of test included, relatively appropriate number of stations, duration of test 1- hours, separation between two observers, and the preparation of students for this assessment model.

Because of several limitations, generalizations from our results should be made with caution. Small sample size was the first limitation. We recommend further research with a larger sample. The second limitation was the number of 8 stations. Final examination is a comprehensive examination in which more stations can increase the validity of the exam. Further researches with a larger number of stations are recommended. Also, this study combined OSCE with observed practice. Further research can explore whether other new assessment methods are reliable method in assessing midwifery competency in the final exam. Thus, although this study

advocates OSCE method as a reliable assessment for final examination more research is needed about the best uses of the OSCE method and how to maximize reliability and validity in midwifery education.

Hence, this study is not advocate replacing the clinical examination on real patients with OSCE. We are rather advocating a combination of the OSCE and observed practice in clinical settings (labor and delivery room) in midwifery students' final exam.

5. CONCLUSION

The combination of observed practice with OSCE examination has been proven as a reliable method for assessing midwifery's competence in final exam. Although it is nearly impossible to have assessment method that will satisfy all criteria of a good test, modified final examination, combination OSCE with observed practice, is acceptable alternative for traditional final examination.

ACKNOWLEDGEMENTS

The authors would like to thank Babol University of Medical Sciences for all their support, the midwifery students for their participation in this study, and to Dr. Evangeline Foronda for the English editing.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Ali E. Use of Video-Projected Structured Clinical Examination (ViPSCE) instead of the traditional oral (Viva) examination in the assessment of final year medical students. *Educ Health (Abingdon)*. 2004;17(1):17-26.
2. Smith J, Plotkin M, Vinh TQ, Thanh CN, Huu NV, Thu Ha NT, et al. Introduction of OSCE for clinical assessment in obstetrics and gynaecology in Hue, Vietnam. *South East Asian Journal of Medical Education*. 2008;2(1):25-30.
3. Bakhsh TM, Sibiany AM, Al Mashat FM, Meccawy AA, Al Thubaity FK. Comparison of students' performance in the traditional oral clinical examination and the objective structured clinical examination. *Saudi Med J*. 2009;30(4):555-7.
4. Tabatabai K, Yazdi M, Ebrahimzadeh S, Bahreini Toosi V, Bahreini Toosi K. Medical students viewpoints about the evaluation methods at internship stage Mashad University. *MUMS*. 2002;45(76):97-103.
5. Aghajani Delavar M, Salmalian H, Faramarzi M, Pasha H, bakhtiari A, Nickpour M, Mohsenzadeh LF. Using the Objective structured clinical examinations in undergraduate midwifery students. *Journal of Medicine and Life*; 2012. In press.
6. Faramarzi M, Pasha H*, Bakhtiari A, Salmalian H, Agajani Delavar M, Nasiri Amiri F, Nikpour M. Test anxiety in objective structured clinical examinations (OSCEs) compared with traditional assessment methods in undergraduate midwifery students. *Health*. 2013;5:2204-2209.
7. Harden RM, Stevenson M, Downie WW, Wilson GM. Assessment of clinical competence using objective structured examinations. *British Medical Journal*. 1975;1:447-451.
8. Barry M, Noonan M, Bradshaw C, Murphy ST. An exploration of student midwives' experiences of the objective structured clinical examination assessment process. *Nurse Educ Today*; 2011. In press.
9. Ross M, Carroll G, Knight J, Chamberlain M, Fothergill-Bourbonnais F, Inton J. Using the OSCE to measure clinical skills performance in nursing. *J Adv Nurse*. 1988;13:45-56.
10. Brains Willnat, Manheim Rich. *Empirical Political Analysis* 8th edition. Boston, MA: Longman p. 2011;105.
11. Kramer GP, Bernstein DA, Phares V. *Introduction to clinical psychology*. 7th ed. Upper Saddle River, NJ: Pearson Prentice Hall, 2009. Print.
12. Nickbakht M, Amiri M, Latifi SM. Study of the reliability and validity of objective structured clinical examination (OSCE) in the assessment of clinical skills of audiology students. *Glob J Health Sci*. 2013;5(3):64-8. DOI: 10.5539/gjhs.v5n3p64.
13. Smith V, Muldoon K, Biesty L. The Objective Structured Clinical Examination (OSCE) as a strategy for assessing clinical competence in midwifery education in Ireland: A critical review. *Nurse Educ Pract*. 2012;12:242-247.
14. Moattari M, Abdollah-zargar S, Mousavinasab M, Zare N, Beygi Marvdast P. Reliability and validity of OSCE in evaluating clinical skills of nursing students. *Journal of Medical Education Summer*. 2009;13(3):79-85.
15. Turner JL, Dankoski ME. Objective structured clinical exams: A critical review. *Fam Med*. 2008;40(8):574-8.
16. Barman A. Critiques on the objective structured clinical examination. *Ann Acad Med Singapore*. 2005;34:478-482.
17. Carraccio C, Englander R. The objective structured clinical examination: A step in direction of competency-based evaluation. *Arcb Pediatr Adolesc Med*. 2000;154:736-741.
18. Konje JC, Abrans KR, Taylor DJ. How discriminatory is the Objective Structured Clinical Examination (OSCE) in the assessment of clinical competence of medical students. *J Obstet Gynaecol*. 2001;21:223-227.
19. Sloan DA, Donnelly MB, Johnson SB, et al. Use of an Objective Structured Clinical Examination (OSCE) to measure improvement in clinical competence during the surgical internship. *Surgery*. 1993;114:343-351.

20. Prislín MD, Fitzpatrick CF, Lie D, et al. Use of an objective structured clinical examination in evaluating student performance. *Fam Med.* 1998;30:338–344.
21. Swift SE, Carter JF. Institution and validation of an observed structured assessment of technical skills (OSATS) for obstetrics and gynecology residents and faculty. *Am J Obstet Gynecol.* 2006;195:617–23.
22. Holmboe ES. Faculty and the observation of trainees's clinical skills: Problems and opportunities. *Acad Med.* 2004;79:1622. DOI:10.1097/00001888-200401000-00006, PMID146909924.
23. Auewarakul C, Downing SM, Jaturatamrong U, Praditsuwan R. Sources of validity evidence for an internal medicine student evaluation system: An evaluative study of assessment methods. *Med Educ.* 2005;39:276-83.
24. Ram P, van der Vleuten C, Rethans JJ, Schouten B, Homba S, Grol R. Assessment in general practice: The predictive value of written knowledge tests and a multiple-station examination for actual medical performance in daily practice. *Med Educ.* 1999;33:197-203.
25. Erfanian F, Khadivzadeh T. Evaluation of midwifery students' competency in providing intrauterine device services using objective structured clinical examination. *IJNMR.* 2011;16(3):191-196.
26. Newble DI, Noare J, Elmslie RG. Validity and reliability of a new examination of the clinical competence of medical students. *Med Educ.* 1981;15:46-52.
27. Mazor KM, Ockene JK, Rogers HJ, Carlin MM, Quirk ME. The relationship between checklist scores on a communication OSCE and analogue patients' perceptions of communication. *Adv Health Sci Educ.* 2005;10(1):37-51.
28. Rutala PJ, Witzke DB, Leko EO, Fulginiti JV, Taylor PJ. Student fatigue as a variable affecting performance in an objective standardized clinical examination. *Acad Med.* 1990;65:S53-S54.
29. Tamblyn RM, Klass DJ, Schnabl GK, Kopelow ML. Sources of unreliability and bias in standardized patient rating. *Teach Learn Med.* 1991;3:74-85.
30. Swanson DB, Norcini JJ. Factors influencing reliability of tests using standardized patients. *Teach Learn Med.* 1989;1:158-66.
31. Van Der Vleuten CPM, Swanson DB. Assessment of clinical skills with standardized patients: state of the art. *Teach Learn Med.* 1990;2:58-76.

© 2015 Faramarzi et al.; This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Peer-review history:

The peer review history for this paper can be accessed here:
<http://www.sciencedomain.org/review-history.php?iid=667&id=32&aid=6094>