http://medicaleducation-bulletin.ir
Letter to Editor (Pages: 833-836)

A Reflection on Sequential OSCE in Medical Education

*Zahra Abdolreza Gharehbagh[1, 2]

[1]Department of Medical Education, Tehran University of Medical Sciences, Tehran, Iran. [2]Faculty Member, Department of Pediatric Nursing, Faculty of Nursing and Midwifery, Tehran Medical Sciences, Islamic Azad University, Tehran, Iran.

Dear Editor-in-chief

Practical exams are frequently used in the training and evaluation of medical students. One such exam is the objective structured clinical examination (OSCE). OSCE exam is organized objectively in different stations, and the examinees are asked to perform certain clinical tasks in each station (1). One of the limitations of the OSCE exam is the high cost of its implementation (2). Increasing the number of OSCE stations reduces the measurement error and increases the reliability of the test (3). However, reducing the number of stations can save the number of examiners, patients, or SPs (4). The sequential OSCE method has been proposed to deal with the cost problem in the OSCE exam, which saves resources while claiming high reliability (5). In sequential OSCE, all students first participate in a test with a limited number of stations. For example, if there are supposed to be 20 stations in the main test, the preliminary test only contains 10 stations or less. Reducing the number of stations, patients, SPs, and examiners lowers the cost of conducting this exam. A student who fails this test will participate in the supplemental test, and their total performance in the two tests will be the criterion for deciding their status (3, 4), as if they participated in the main exam once. However, the students who pass the initial exam are exempted from the next exam, and only the first exam is considered for them (4). In sequential assessment, the results of the first test should accurately predict students' performance in the main test, that is, it should be able to screen capable learners correctly. The correct and reliable performance of the sieve test requires observing several points: designing the stations of the whole test first and then selecting what to include in the first test from among them. In other words, the screening test should be considered in the context and text of the blueprint of the whole test. It should also be noted how many students usually fail the test. In most similar exams, a small number of students are expected to fail. Another important point is that the screening test should be designed so that false negatives and false positives are low (6). False positive means students who have passed the screening exam but are ultimately rejected in the total exam. If the false positive is low, it indicates fewer mistakes in decision-making and high specificity of the screening test. OSCE is a test to confirm the clinical competence of learners, and if an incompetent student is declared accepted, it can be a serious threat. False negative means students who failed the test, but their performance was acceptable in two tests, and they were finally declared passed. If the number of negative cases (i.e., students who failed the first test) is small, it means that a second (supplementary) test is to be held for a small number of students and is economical. If there are many negative cases, the second test must be repeated for a large number of students, meaning the main goal (i.e., saving money) is not achieved. Ideally, false positives and false negatives should be low, but low false positives are more important in the screening test (6, 7).

Key Words: Cost, Objective Structured Clinical Examination, Sequential, Medical Education.

*Please cite this article as: Abdolreza Gharehbagh Z. A Reflection on Sequential OSCE in Medical Education. Med Edu Bull 2023; 4(4): 833-36. DOI: **10.22034/MEB.2024.434280.1089**

Corresponding Author:

Zahra Abdolreza Gharehbagh, Department of Medical Education, Tehran University of Medical Sciences, Tehran, Iran.
Email: Zgharebagh@iautmu.ac.ir
Received date: Jul. 25, 2023; Accepted date: Dec.12, 2023

A review article about implementing OSCE by sequential assessment states that the rejection and acceptance rates of students in the screening test can be determined by calculating the passing rate using a standard determination method such as the borderline regression method. Then, for higher certainty and lower false positives, the acceptance rate can be declared as two standard errors of measurement (SEM) higher than the obtained number. In this case, it can be said with great confidence that this group of students is really capable and worthy of acceptance (5, 8). The first empirical study on sequential OSCE was conducted retrospectively by Colliver et al. (1992) at the Southern Illinois University School of Medicine. The authors concluded that the screening test can predict the performance of students in the total test well. However, they mentioned that the number of screening test stations and the passing score could affect the accuracy of the results of the screening test, which was not addressed in this study (9). These researchers investigated the above two cases in another study (1992). The results showed that good accuracy can be attained with a screening test that is only one-third the length of the full examination, and the cutoff for this screen should be set slightly above the mean of the pass levels to maximize sensitivity and specificity (10). The remaining issue was the characteristics of the stations used in screening. For this purpose, another study was conducted by the same researchers (1995), and the results showed that one-third of the main test stations, where the item-total correlation index is high, are suitable as initial tests. Therefore, the item difficulty indicator of the station has no effect (11).

The above studies aimed to confirm the use of sequential OSCE, but from a methodological point of view, they were not conducted consecutively, and the data analysis was done retrospectively. Smee et al. (2003) conducted a study on the OSCE of the Canadian Medical System Organization with the participation of a large number of students simultaneously in several centers. They decided to hold a real OSCE screening. The results of this study showed that sequential assessment was attractive for the test organizers, could reduce costs and solve the question bank problem to some extent, and could alleviate the coordination problem with a large number of examiners, SPs, and patients. However, there were other views that did not agree with this method. Despite the planned arrangements, correcting the papers did not proceed at the desired speed and contained evident errors. Also, students and faculties, who could not find justification regarding the purpose and process of successive assessments, spread false notions after the test, leading to dissatisfaction and disruption. This issue shows the necessity of adequate justification and training of people before holding sequential tests (12). Cookson et al. (2011) also describe the experience of holding an OSCE using the sequential evaluation method. In this study, the 2009 final exam of York Hall School of Medicine, consisting of 12 OSCE stations and eight real patient encounters (the OSLER), was conducted on 127 students using a sequential measurement method. The students whose performance in the first six stations and four real exposures were evaluated as completely satisfactory were exempted from continuing the test. The rest of the students (about a third) had to participate in the entire test, and their performance in the entire test was the criterion for deciding their status. The first part of the test predicted the overall performance of the students well, and the cost savings was £30,000. However, the researchers noted that the cost reduction was not solely due to sequential measurement and was likely linked to combining OSCE and OSLER as well (13).

In a study that used data from a graduate OSCE in England, Pell et al. (2013) aimed to outline the theoretical case for sequential testing, explore the impacts from both candidate and institutional perspectives, and estimate cost savings and improvements to reliability for borderline students. The results showed that sequential testing in OSCEs increased reliability for borderline students as the increased number of observations implied that “observed” student marks were closer to “true” marks. However, the station-level quality of the assessment should be sufficiently high to achieve the full benefits of reliability. The introduction of such a system has financial benefits and good validity inferences and has proved acceptable to students and other stakeholders (14). Currie et al. (2016) published an article where they examined the effect of the number of different stations and quorums of the screening test on the results. They used OSCE data from the Medical School of the University of Aberdeen in two consecutive years. The data demonstrated that in two separate years, the sensitivity across 6-14 stations was consistently high irrespective of the number of screening stations, while the specificity generally increased (with progressive narrowing of the 95% confidence intervals) as the number of screening OSCE stations increased. Their data expanded the body of evidence that sequential OSCEs can confer a reasonably high sensitivity and specificity for different numbers of screening day OSCE stations. The highest specificity was seen in the 11, 13, and 14-station tests, but the researchers concluded that this number of stations is not efficient in terms of implementation, and the eight-station test with a sensitivity of 88-89%, and a specificity of 83-86% is a suitable choice (15).

Mortaz Hejri et al. (2016) conducted a study to plan the optimal model of the screening test using the data of the pre-internship OSCE, considering three basic factors. The psychometric characteristics of pre-internship OSCE stations of Tehran University of Medical Sciences were determined based on the classical test and the item-response theories. Several screening tests with different numbers of stations, psychometric characteristics, and acceptance rates were designed hypothetically, and their results were compared with the main test. The researchers concluded that sequential OSCE can be an efficient and suitable method for conducting OSCE. Also, an efficient screening OSCE with the lowest probability of an error rate requires the careful selection of stations with high values of discrimination or total correlation and the use of a relatively stringent cut-off score (16). Homer et al. (2018) compared a ‘screening’ test (S1) ‘decisions’ to second ‘test’ (S2) overall pass/fail decisions to assess diagnostic accuracy in a sequential model and evaluated the longitudinal performance of failing students using changes in percentile ranks over a full, repeated year. The results showed a small but important improvement in diagnostic accuracy under a sequential model (of the order 2-4% of students misclassified under a traditional model). Also, after a re-sitting year, weaker students’ rankings relative to their peers improved by 20 to 30 percentile points (17). Another study (2019) aimed to examine the acceptability of the sequential OSCE by identifying students’ views. Final-year students at a Scottish university completed a questionnaire after day 1 of a sequential OSCE. Most respondents strongly agreed/agreed that they felt stressed about the sequential OSCE (98.1%), would feel like a failure if taking Day 2 (89.7 %), and that Day 2 seemed the same as a re-sit (78.5%). However, 61.7% agreed that fewer exam days were a positive aspect of the sequential OSCE. Open comments indicated feelings of increased stress, anxiety, and frustration associated with the sequential OSCE (18).

CONCLUSION

In general, holding sequential OSCE can help save costs while maintaining characteristics such as the reliability and validity of the test. However, ignoring specific considerations in its implementation can cause serious damage to the quality of the test or even make it not cost-effective. For this purpose, it is necessary to plan a hypothetical sequential test before conducting a sequential OSCE and compare the specifications and features of the current test to determine whether holding a consecutive OSCE is effective and useful. Some of the challenges in this regard are determining the number of stations for the sieve test, how to choose stations for the screening test, determining the quorum to pass the screening test, and issues related to policies, regulations, and information.

REFERENCES

1. Ghanbari A, Monfared A, Hoseinzadeh T, Moaddab F, Sedighi A. The Impact of the Nursing Process Education on Critical Thinking of Nursing Students. Research in Education Medical 2017; 9 (2): 25- 33.

2. Hejri SM, Jalili M. Sequential Objective Structured Clinical Examination based on item response theory in Iran. J Educ Eval Health Prof. 2017 Sep 8;14:19.

3. Jalili M, Khabaz Mafinejad M, Gandomkar R, Mortaz Hejri S. Principles and Methods of Student Assessment in Health Professions. Tehran: The Academy of Medical Sciences of Iran, 2019.

4. Mortaz Hejri S, Jalili M. Sequential OSCE based on item response theory. J Educ Eval Health Prof 2017; 14:19.

5. Gormley GJ, McCusker D, Booley MA, McNeice A. The use of real patients in OSCEs: a survey of medical students' predictions and opinions. Med Teach. 2011;33(8):684.

6. Rothman, A.I., Blackmore, D.E., Dauphinee, W.D. and Reznick, R. 'Tests of sequential testing in two years' results of Part 2 of the Medical Council of Canada Qualifying Examination', Academic Medicine, 1997; 72: S22-S24.

7. Muijtjens, A.M., van Vollenhoven, F.H., van Luijk, S.J., van der Vleuten, C.P. 'Sequential testing in the assessment of clinical skills', Academic Medicine, 2000; 75(4):369.

8. Currie GP, Sivasubramaniam S, Cleland J. Sequential testing in a high stakes OSCE: Determining number of screening tests. Med Teach. 2016 Jul;38(7):708-14.

9. Colliver, J.A., Mast, T.A., Vu, N.V. and Barrows, H.S. 'Sequential testing with a performance-based examination using standardized patients', Academic Medicine, 1991;66(9 Suppl): S64-6.

10. Colliver, J.A., Vu, N.V., Barrows, H.S. 'Screening test length for sequential testing with a standardized-patient examination: a receiver operating characteristic (ROC) analysis', Academic Medicine, 1992; 67: 592-95.

11. Colliver, J.A., Markwell, S.J., Travis, T.A., Schrage, J.P., et al. 'Sequential testing with a standardized-patient examination: an ROC analysis of the effects of case-total correlations and difficulty levels of screening test cases', Proceedings of the 6th Ottawa International Conference on Medical Education, 1995; pp. 26-9.

12. Smee SM, Dauphinee WD, Blackmore DE, Rothman AI, Reznick RK, Des Marchais J. A sequenced OSCE for licensure: administrative issues, results and myths. Adv Health Sci Educ Theory Pract. 2003;8(3):223-36.

13. Cookson J, Crossley J, Fagan G, McKendree J, Mohsen A. A final clinical examination using a sequential design to improve cost-effectiveness. Med Educ. 2011 Jul;45(7):741-7.

14. Pell G, Fuller R, Homer M, Roberts T. Advancing the objective structured clinical examination: sequential testing in theory and practice. Med Educ. 2013 Jun;47(6):569-77. doi: 10.1111/medu.12136. PMID: 23662874.

15. Currie GP, Sivasubramaniam S, Cleland J. Sequential testing in a high stakes OSCE: Determining number of screening tests. Med Teach. 2016 Jul;38(7):708-14.

16. Mortaz Hejri S, Yazdani K, Labaf A, Norcini JJ, Jalili M. Introducing a model for optimal design of sequential objective structured clinical examinations. Adv Health Sci Educ Theory Pract. 2016 Dec;21(5):1047-1060. doi: 10.1007/s10459-016-9673-x.

17. Homer, MS, Fuller, R and Pell, G. The benefits of sequential testing: Improved diagnostic accuracy and better outcomes for failing students. Medical Teacher, 2018;40 (3):275-84. ISSN 0142-159X.

18. Duncumb M, Cleland J. Student Perceptions of a Sequential Objective Structured Clinical Examination. Journal of the Royal College of Physicians of Edinburgh, 2019;49(3):245-49.