http://medicaleducation-bulletin.ir
Letter to Editor (Pages: 834-837)

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). The OSCE exam is organized objectively in different stations, and examinees are asked to perform specific clinical tasks at each station (1). One limitation of the OSCE exam is its high implementation cost (2). Increasing the number of OSCE stations reduces measurement error and enhances test reliability (3). Conversely, reducing station numbers can help save resources like examiners, patients, or standardized patients (SPs) (4). The sequential OSCE method has been proposed to address cost challenges while maintaining high reliability (5). In this approach, students first participate in a preliminary test with fewer stations. For instance, if the main test typically involves 20 stations, the initial test might include only 10 or fewer stations. By reducing the number of stations, patients, SPs, and examiners, the sequential OSCE method significantly lowers examination costs. Students who do not pass the initial test will participate in a supplemental test, with their total performance across both tests determining their final status (3, 4), equivalent to participating in a single comprehensive exam. Students who successfully pass the initial exam are exempted from the subsequent test, with only their first exam performance being considered (4).

In sequential assessment, the results of the first test should accurately predict students' performance in the main test, effectively screening capable learners. Achieving this requires careful consideration of several key points: designing the entire test's blueprint first and then strategically selecting stations for the initial screening test. The screening test must be conceptualized within the comprehensive test's framework and context. Understanding the typical failure rate is crucial, as most similar examinations expect only a small number of students to fail. The screening test should be meticulously designed to minimize both false negatives and false positives (6). False positives occur when students pass the screening exam but are ultimately rejected in the total exam, indicating potential decision-making errors and low test specificity. Since OSCE assesses clinical competence, incorrectly accepting an incompetent student poses a serious threat to patient safety. False negatives represent students who initially fail the test but demonstrate acceptable performance across both tests and are ultimately passed. The number of negative cases is critical: a small number suggests an economical second test, while a large number undermines the primary goal of resource conservation. Ideally, both false positives and false negatives should be minimized, with a particular emphasis on maintaining low false positive rates 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): 834-37. 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 through sequential assessment suggests that student rejection and acceptance rates in the screening test can be determined by calculating the passing rate using standard methods like the borderline regression method. For enhanced certainty and reduced false positives, the acceptance rate can be declared two standard errors of measurement (SEM) higher than the obtained number, providing significant confidence in the selected students' capabilities (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 researchers concluded that the screening test could effectively predict student performance in the total test, while noting that the number of screening test stations and passing score could influence result accuracy, an aspect not fully explored in their initial study (9). In a subsequent investigation, they demonstrated that a screening test comprising one-third of the full examination length could achieve good accuracy, with the cutoff set slightly above the mean pass levels to optimize sensitivity and specificity (10). A further study by the same researchers in 1995 revealed that one-third of the main test stations with high item-total correlation are suitable for initial testing, and interestingly, the station's item difficulty indicator does not significantly impact the screening process (11).

The previous studies aimed to validate sequential OSCE, but methodologically, they were not conducted consecutively and relied on retrospective data analysis. Smee et al. (2003) conducted a comprehensive study on the OSCE of the Canadian Medical System Organization, involving numerous students across multiple centers and implementing a real OSCE screening. The results demonstrated that sequential assessment was attractive to test organizers, offering potential cost reduction, addressing question bank challenges, and mitigating coordination issues with multiple examiners, standardized patients (SPs), and patients. However, alternative perspectives highlighted significant implementation challenges. Despite careful planning, paper correction was slower than anticipated and contained notable errors. Furthermore, students and faculty members, lacking clear understanding of the sequential assessment's purpose and process, propagated misconceptions post-test, resulting in widespread dissatisfaction and potential disruption. This experience underscored the critical importance of comprehensive participant education and justification before implementing sequential testing (12). Cookson et al. (2011) provided another perspective by describing their experience with sequential evaluation at York Hall School of Medicine. In their 2009 final exam involving 127 students, the test comprised 12 OSCE stations and eight real patient encounters (OSLER). Students demonstrating satisfactory performance in the first six stations and four patient exposures were exempted from completing the entire test, with approximately one-third of students required to participate fully. The initial test segment effectively predicted overall student performance, and the approach generated approximately £30,000 in cost savings. However, the researchers cautioned that cost reduction was not exclusively attributable to sequential measurement but likely resulted from combining OSCE and OSLER methodologies (13).

In a study utilizing graduate OSCE data from England, Pell et al. (2013) explored the theoretical foundations of sequential testing, examining its impacts from candidate and institutional perspectives while assessing potential cost savings and reliability improvements for borderline students. The research revealed that sequential testing in OSCEs enhances reliability for borderline students by increasing the number of observations, thereby bringing "observed" student marks closer to "true" marks. However, the researchers emphasized that achieving full reliability benefits requires maintaining high station-level assessment quality. The proposed system demonstrated financial advantages, strong validity, and was well-received by students and stakeholders (14).

Currie et al. (2016) further investigated the impact of station numbers and screening test quorums using OSCE data from the University of Aberdeen Medical School across two consecutive years. Their findings showed consistently high sensitivity across 6-14 stations, with specificity generally increasing as the number of screening stations expanded. The study broadened existing evidence supporting sequential OSCEs' capacity to deliver reliable sensitivity and specificity across various screening station configurations. While the highest specificity emerged in 11, 13, and 14-station tests, the researchers determined these were impractical for implementation. Instead, they recommended an

eight-station test, which demonstrated a sensitivity of 88-89% and specificity of 83-86%, as the most suitable approach (15).

Mortaz Hejri et al. (2016) conducted a comprehensive study to develop an optimal screening test model for pre-internship OSCE, analyzing psychometric characteristics using classical test and item-response theories. The researchers designed multiple hypothetical screening tests with varying station numbers, psychometric characteristics, and acceptance rates, comparing their results against the main test. Their findings concluded that sequential OSCE represents an efficient examination method. Critically, they determined that an effective screening OSCE with minimal error probability requires carefully selecting stations with high discrimination or total correlation values and implementing a relatively stringent cut-off score (16). Homer et al. (2018) investigated the diagnostic accuracy of a sequential testing model by comparing screening test (S1) and second test (S2) pass/fail decisions while evaluating failing students' longitudinal performance. Their research revealed a modest yet significant improvement in diagnostic accuracy, with approximately 2-4% of students potentially misclassified under traditional models. Notably, after a re-sitting year, weaker students demonstrated substantial ranking improvements, advancing 20 to 30 percentile points (17).

A subsequent 2019 study explored student perspectives on sequential OSCE by surveying final-year students at a Scottish university. The findings revealed high stress levels, with 98.1% feeling anxious about the sequential OSCE, 89.7% fearing failure if requiring a second day, and 78.5% perceiving the second day as equivalent to a re-sit. Conversely, 61.7% appreciated the reduced number of exam days. Open comments further highlighted increased stress, anxiety, and frustration associated with the sequential OSCE approach (18). These studies collectively demonstrate the complex landscape of sequential OSCE implementation, revealing both potential advantages in terms of diagnostic accuracy and cost-effectiveness, as well as significant psychological challenges for students undergoing this assessment method.

CONCLUSION

In general, implementing a sequential OSCE can help reduce costs while maintaining the test's reliability and validity. However, overlooking specific implementation considerations can potentially compromise test quality or negate potential cost-effectiveness. To mitigate these risks, it is crucial to develop a comprehensive hypothetical sequential test model and meticulously compare its specifications and features with the existing examination format to determine the potential effectiveness and utility of a sequential OSCE approach.

Key challenges in designing a sequential OSCE include strategically determining the number of screening test stations, selecting appropriate stations for the initial screening, establishing precise passing criteria for the screening test, and addressing complex policy, regulatory, and logistical considerations. These critical factors require careful planning and systematic evaluation to ensure the sequential OSCE achieves its intended goals of resource optimization without sacrificing assessment integrity and educational standards.

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.