## Original Article

# Comparing Objective Structured Clinical Examinations and Traditional Clinical Examinations in the Summative Evaluation of Final-Year Medical Students

Balantine Ugochukwu N Eze, Anthony Jude Edeh, Anthony Ikemefuna Ugochukwu

Department of Surgery, Enugu State University of Science and Technology, Enugu, Enugu State, Nigeria

**ABSTRACT**

**Background:** Medical schools have traditionally assessed medical students using long and short cases. Objective structured clinical examination (OSCE) has been found to be more reliable. **Aim:** To compare OSCE and traditional method of assessment in the summative assessment of final-year medical students. **Methodology:** This was a retrospective cross-sectional study conducted at Enugu State University of Science and Technology College of Medicine. The Department of Internal Medicine organized clinical examinations consisting of long and short cases. The Department of Surgery organized an OSCE consisting of two parts (picture OSCE and clinical OSCE). Students' scores in both internal medicine and surgery were collated and subjected to analysis with SPSS version 23 (IBM; SPSS, Chicago, IL, USA). Pearson's correlation was used to assess the correlations, paired $t$-test was used to compare the mean scores, and Cronbach's alpha was used to assess the reliability. $P < 0.05$ was considered statistically significant. **Results:** Out of the 73 candidates, 41 were female and 32 were male giving a female: male ratio of 1.3:1. Using paired sample t test, there were significant differences between the mean score in long case (mean = 52.86, standard deviation [SD] = 4.315) and mean score in clinical OSCE (mean = 58.356, SD = 7.906), $t$ (72) = −7.181, $P$ = 0.000; mean score in short case (mean = 52.86, SD = 4.097) and mean score in picture OSCE (mean = 48.580, SD = 8.992, $t$ (72) =4.558, $P$ = 0.000; no significant difference between the mean total score in internal medicine clinicals (mean = 105.712, SD = 6.680) and mean total score in surgery clinicals (mean = 106.915, SD = 15.846), $t$ (72) = −0.788, $P$ = 0.433. The Cronbach's alpha for traditional examination and OSCE was 0.437 and 0.863, respectively. **Conclusion:** OSCE gives a similar mean score to traditional method, but OSCE is more reliable.

**KEYWORDS:** *Medical students, objective structured clinical examination, summative assessment, traditional clinical examinations*

## INTRODUCTION

Assessment of clinical skills has a key role in medical education, and the selection of suitable methods of assessment is highly relevant.[1-3] The need to select suitable methods of assessing clinical competence has been a matter of perpetual concern for clinical teachers, course directors, and medical educators.[4-7] The assessment of clinical competence is fundamental to ensure that graduate doctors and other health-care professionals can perform the required duties in patients'

care.[8] Learning will not be thorough if the method of assessment is improper, as students or trainees will not apply themselves fully.[9]

Colleges of medicine have traditionally assessed medical students' clinical competence using long cases, short

*Address for correspondence:* Dr. Balantine Ugochukwu N Eze,
Department of Surgery, Enugu State University of Science and
Technology, Enugu, Enugu State, Nigeria.
E-mail: buneze@rocketmail.com

**Access this article online**

**Quick Response Code:**

**Website:** www.nigerianjsurg.com

**DOI:** 10.4103/njs.NJS_19_20

cases, and *vivas*. Long and short cases involve direct observation of students and examining them in history taking, physical examination, ability to make diagnosis, order and interpret relevant investigations, proffer possible treatment, and manage possible complications. The long and short cases are based on the limited number of patient cases that the students' can encounter and usually have an unstructured format and can be subjective.[5] Judgment based on a limited number of case (s) can be harmful or beneficial to the examinee as the case (s) can be difficult or simple and research has shown that multiple cases are needed for a reliable assessment of skills or ability.[9]

Educationists long realized the need for a valid and reliable assessment tool in skill-based subjects such as medicine, surgery, dentistry, pharmacy, nursing, midwifery, physiotherapy, and even police education.[5] Objective structured clinical examination (OSCE) is an objective method of assessing clinical knowledge, professional judgment, interpersonal and professional communication, and problem-solving skills. In this method, a number of stations, each of which contains specific clinical scenario, are used.[10] The examinees are observed and their performance is assessed using structured checklists.[11]

OSCE has become a common method to assess clinical and procedural skills in undergraduate medical education since its introduction by Harden *et al.* in 1975.[12] The OSCE has been found to be a feasible approach to the assessment of clinical competence for use in different cultural and geographical contexts.[13] OSCE has been hailed as the "gold standard" for clinical assessments of medical and other health-care students.[10,14] Standardized patients are used in OSCEs to ensure that each student encounters identically portrayed scenarios.[13] Students have been said to view OSCE as a valid, realistic, and fair assessment method with high levels of satisfaction and positive assessment experience.[10,15]

Despite OSCE being considered worldwide as "a gold standard" for the evaluation of medical students,[10,14] several medical schools in developing countries (including our own medical school) have not yet adapted the use of OSCE.[9] This may be due to variation in viewpoints. Chu *et al.*[16] identified two viewpoints as regards the use of OSCE versus traditional methods in the evaluation of students in the medical sciences. Viewpoint 1 supports the traditional use of live patients, arguing that other assessment models have not yet been demonstrated to be viable alternatives to the actual treatment of patients in the clinical licensure process. This viewpoint also contends that the use of live patients and inherent variances in live patient used

represents the realities of day-to-day practice. Viewpoint 2 argues that the use of live patients in licensure examinations needs to be discontinued considering those examinations' ethical dilemmas of exposing patients to potential harm, as well as their lack of reliability and validity and limited scope.

In the Enugu State University of Science and Technology (ESUT) College of Medicine, Enugu, OSCE was first used in the summative assessment of final-year medical student by the Department of Surgery in June, 2019, while the Department of Internal Medicine still maintained the traditional method of assessment. The aim of this study is to assess and compare OSCE and traditional method of assessment in the summative assessment of final-year medical students.

## METHODOLOGY

This was a retrospective descriptive cross-sectional study of summative assessment of final-year medical students in Surgery and Internal Medicine at ESUT, College of Medicine. Both departments had a written examination consisting of essay questions and multiple choice questions. In addition, the Department of Internal Medicine also organized clinical examinations consisting of long and short cases. This traditional clinical examination basically tested a narrow range of clinical skills under the observation of, normally, two examiners in a given clinical case in the form of patient histories, demonstration of physical examinations, assessment of a narrow range of technical skills, and orals.

The Department of surgery organized an OSCE consisting of two parts. The A part (picture OSCE, replacing traditional short case) in which fifty questions were given to students from slide shows, which contains not only simple questions and case scenarios but also instruments, images of disease conditions and diagnostic images such as X-rays and computed tomography scans for interpretation. The students had 2 min to answer the question (s) in each slide. In this picture OSCE, the slide automatically transits to the next slide projection after 2 min. The B part (clinical OSCE, replacing traditional long case) consists of twenty stations and three rest stations. A ring of a bell and stopwatch are used to control entrance into and exit from the stations. Each student spent 4 minutes at a station. At the end of 4 minutes, the student, on the sound of the bell, leaves the station and moves to the next one. Each student performs the same tasks and was marked and assessed according to the same assessment criteria on the examiner's checklist. The criterion referenced system[9] was adopted.

The students' scores in the traditional examination, OSCE, and final total score in both internal medicine

and surgery were collated and subjected to analysis with SPSS version 23 (IBM; SPSS, Chicago, IL, USA). Pearson's correlation was used to assess the correlations, paired *t*-test was used to compare the mean scores, and Cronbach's alpha was used to assess the reliability. $P < 0.05$ was considered statistically significant.

## RESULTS

Out of a total of 82 candidates who sat for the examinations, 7 candidates sat for only internal medicine and 2 candidates sat for only surgery and were excluded from the analysis. Out of the 73 candidates sat for both internal medicine and surgery, 41 were female and 32 were male giving a female: male ratio of 1.3:1. There was a positive correlation between students' score in clinical OSCE (replacement for long case by surgery department) and long case, $r = 0.525$ ($P = 0.000$) and a significant positive correlation between students' score in picture OSCE (replacement for short case by surgery department) and short case, $r = 0.450$ ($P = 0.003$). There was also a positive correlation between scores in internal medicine clinicals (long case + short case) and scores in surgery clinicals (picture OSCE + clinical OSCE), $r = 0.593$ ($P = 0.000$). Using paired sample *t*-test to compare the mean scores, there was a significant difference between the mean scores in long case (mean = 52.86, standard deviation [SD] = 4.315) and scores in clinical OSCE (mean = 58.356, SD = 7.906), t (72) = −7.181, $P = 0.000$. There was also a significant difference between the mean scores in short case (mean = 52.86, SD = 4.097) and picture OSCE (mean = 48.580, SD = 8.992, t (72) = 4.558, $P = 0.000$. There was no significant difference between the mean total scores in internal medicine (long case + short case) clinicals (mean = 105.712, SD = 6.680) and surgery (clinical OSCE + picture OSCE) clinicals (mean = 106.915, SD = 15.846), t (72) = −0.788, $P = 0.433$. The Cronbach's alpha for traditional examination (comparing the individual student's scores in long case versus short case) and OSCE (comparing the individual student's scores in clinical OSCE versus picture OSCE) was 0.437 and 0.863, respectively.

## DISCUSSION

One of the key issues plaguing the undergraduate medical education is the ineffective methods of assessments.[17] Utilizing different patients, different procedures with different difficulty levels introduce some elements of subjectivity in the assessment process[18] and this is the major setback of the traditional method of assessment. Haider *et al.*[19] found that the element of pass by chance and bias was less in OSCE compared to the traditional ways of assessing medical students. OSCE is

also a flexible assessment tool as Spanke *et al.*,[1] while investigating the fairness and objectivity of OSCE using a multiple scenario approach (where all students had to manage the same chief complaint at a station but its underlying scenarios being randomly changed during students' rotation) found that improving rater training is more important to ensure objectivity and fairness of OSCE than providing the same scenario to all students.

Significant positive correlations that were found between long case and clinical OSCE ($P = 0.000$), short case and picture OSCE ($P = 0.000$), and total score in internal medicine clinical and total score in surgery clinical ($P = 0.000$) suggest that both the traditional and OSCE methods of assessment show a high level of agreement when used as summative assessment tools for medical students.

The mean total score in internal medicine clinical (traditional method) and surgery clinical (OSCE) was not significantly different ($P = 0.433$), and this implies that OSCE does not necessarily produce a significantly higher pass rate as erroneously believed by those that oppose OSCE. Eldarir *et al.*,[4] in evaluating clinical students' performance in an OSCE versus traditional examination, found a significantly higher mean score in OSCE ($P < 0.000$). Soni *et al.*[20] also found a higher mean score in OSCE than in traditional method of clinical skills assessment. On the other hand, Siddaram and Anil[21] in a comparative analysis between OSCE and traditional examination as a formative evaluation tool among MSc nursing students found that traditional examination gave a significantly higher mean score contrary to the finding in the current study and other studies above. This may be due to Hawthorne effect as their student participants already knew which group they belonged to and the purpose of the study. Furthermore, OSCE was held 1 month before the traditional clinical examination.

This study found OSCE to be a more reliable assessment tool ($\alpha = 0.863$) than traditional method of assessment ($\alpha = 0.437$). The reliability of the traditional method found in this study is unacceptable ($\alpha < 0.5$). This supports the fact that OSCE is a fairer assessment tool.[7,13,22] The low reliability of the traditional clinical examination meant a significant variation between individual student's score in long and short cases. The student's scores is not consistently low or high as the case might be. This is not surprising as the traditional clinical examination has been shown to be unreliable.[23] Bias in any examination is a threat to its validity, marks awarded being dependent not only on a candidates' performance but also on other factors which generally are construct irrelevant.[24] In any clinical examinations (where a

candidate's performance is subjectively observed by an examiner), there is a potential risk that an examiners' judgments will depend in part upon their personality, attitudes, or predispositions in general (resulting most obviously in the examiner being a hawk or a dove).[25]

The benefits of the OSCE go well beyond its role as an assessment tool. These include the positive impact that the examination has on student's learning, encouraging students to focus their attention on the development of clinical skills as well as on the acquisition of relevant knowledge.[14] On the one hand, students have reported difficulties with time management and stress control during the OSCE examination.[6,7] While some faculty members have acknowledged the accuracy of the OSCE, some have criticized its limitations for assessing the integrated approach to patients and complained that the examination was remarkably time and effort-consuming.[26] The same study concluded that OSCEs test the students' knowledge and skills in a compartmentalized fashion, rather than looking at the patient as a whole.[26] These are misconceptions.[14]

The most frequently cited argument against the use of the OSCE relates to cost. The OSCE may be viewed as resource intensive with the associated expense militating against its adoption in practice.[14] Researchers have shown that it is possible to organize a good OSCE with very limited resources.[27,28] Limitation of cost did not prevent the Department of Surgery from organizing a successful OSCE.

Although OSCE is not perfect, it is a better proven assessment tool than the traditional system. It is considered to be a much greater improvement over traditional examination methods because the stations can be standardized, enabling fairer peer comparison, and complex procedures can also be assessed without endangering patients' health.[9] According to Vagholkar,[29] there is a need for adequate sensitization toward the OSCE pattern of examination as there is excessive resistance from most faculty members, especially those who have not gone through the OSCE evaluation themselves.

## CONCLUSION

OSCE, as shown in this study, gives a similar mean score to traditional method when used as a summative assessment tool for final-year medical students. OSCE has also been found to be a more reliable method of assessment. We recommend that all medical schools (including our own) should fully adopt OSCE as the standard assessment tool, especially in summative assessment, as it is a reliable and fair tool for assessment devoid of abuses of examiners or patient bias.

**Conflicts of interest**

There are no conflicts of interest.

## REFERENCES

1. Spanke J, Raus C, Haase A, Angelow A, Ludwig F, Weckmann G, *et al*. Fairness and objectivity of a multiple scenario objective structured clinical examination. GMS J Med Educ 2019;36:Doc26.
2. Fowell SL, Bligh JG. Recent developments in assessing medical students. Postgrad Med J 1998;74:18-24.
3. Newble DI. Assessing clinical competence at the undergraduate level. Med Educ 1992;26:504-11.
4. Eldarir SA, Nagwa A, Hamid A. Objective structured clinical evaluation (OSCE) versus traditional clinical students' achievement at maternity nursing: A comparative approach. IOSR JDMS 2013;4:63-8.
5. Khan M, Noor SM, Siraj MU. Students' perceptions of OSCE in dentistry: A study from Khyber College of Dentistry, Pakistan. Adv Health Prof Educ 2015;1:30-6.
6. Shahzad A, Saeed MH, Paiker S. Dental students' concerns regarding OSPE and OSCE: A qualitative feedback for process improvement. BDJ Open 2017;3:17009.
7. Troncon LE. Clinical skills assessment: Limitations to the introduction of an "OSCE" (Objective Structured Clinical Examination) in a traditional Brazilian medical school. Sao Paulo Med J 2004;122:12-7.
8. Brookes D. Objective structured clinical examination assessment. Nurs Times 2007;104:30-1.
9. Onwudiegwu U. OSCE: Design, development and deployment. J West Afr Coll Surg 2018;8:1-22.
10. Beitia G, Beltrán I, Ortega A, Pérez-Mediavilla A, Ramírez MJ. An objective structured clinical examination to assess fifth year pharmacy internship performance. Farma J 2019;4:118.
11. Harden RM, Gleeson FA. Assessment of clinical competence using an objective structured clinical examination (OSCE). Med Educ 1979;13:41-54.
12. Harden RM, Stevenson M, Downie WW, Wilson GM. Assessment of clinical competence using objective structured examination. Br Med J 1975;1:447-51.
13. Patrício MF, Julião M, Fareleira F, Carneiro AV. Is the OSCE a feasible tool to assess competencies in undergraduate medical education? Med Teach 2013;35:503-14.
14. Harden RM. Misconceptions and the OSCE. Med Teach 2015;37:608-10.
15. Jaiswal P, Mehta RK. Medical students' perception regarding objective structured clinical examination in Medical College, Chitwan. J Chitwan Med Coll 2019;9:52-60.
16. Chu TM, Makhoul NM, Silva DR, Gonzales TS, Letra A, Mays KA. Should live patient licensing examinations in dentistry be discontinued? Two viewpoints: Viewpoint 1: Alternative assessment models are not yet viable replacements for live patients in clinical licensure exams and viewpoint 2: Ethical and patient care concerns about live patient exams require full acceptance of justifiable alternatives. J Dent Educ 2018;82:246-51.
17. Javaeed A. The Crisis of Health Professions Education in Pakistan. MedEdPublish 2019. p. 8.
18. Omu FE. Attitudes of nursing faculty members and graduates towards the objective structured clinical examination (OSCE).

Open J Nurs 2016;6:353-64.

19. Haider I, Khan A, Imam SM, Ajmal F, Khan M, Ayub M. Perceptions of final professional MBBS Students and Their examiners about objective structured clinical examination (OSCE): A Combined Examiner and Examinee Survey. J Med Sci 2016;24:206-11.

20. Soni R, Rani S, Thokchom S, Sarkar S. A comparative study to assess the opinion and level of satisfaction of the student nurses regarding objective structured clinical examination (OSCE) and Traditional Method of Clinical Skills Assessment related to Antenatal Examination. Int J Nurs Midwif Res 2017;4:9-12.

21. Siddaram S, Anil S. A comparative analysis between objective structured clinical examination (OSCE) and conventional examination (CE) as Formative Evaluation Tool. Int J Nurs Educ 2018;10:102-5.

22. Omu AE, Al-Azemi MK, Omu FE, Al-Harmi J, Diejomaoh MF. Attitudes of academic staff and students towards the objective structured clinical examination (OSCE) in obstetrics and gynaecology. Creat Educ 2016;7:886-97.

23. Wilson GM, Lever R, Harden RM, Robertson JI. Examination of clinical examiners. Lancet 1969;1:37-40.

24. Al-Saegh RM, Scherpbier AJ, Alhibaly HA, Hmood AR, Almkhtar MA. Perception of OSCE examination in Iraqi undergraduate medical students. Karbala J Med 2015;8:2056-69.

25. McManus IC, Elder AT, Dacre J. Investigating possible ethnicity and sex bias in clinical examiners: An analysis of data from the MRCP (UK) PACES and nPACES examinations. BMC Med Educ 2013;13:103.

26. Gupta P, Dewan P, Singh T. Objective structured clinical examination (OSCE) Revisited. Indian Pediatr 2010;47:911-20.

27. Lucchetti G, Ezequiel OS, Lucchetti AL. An OSCE with very limited resources: Is it possible? Med Teach 2017;39:227.

28. Abdelaziz A, Hany M, Atwa H, Talaat W, Hosny S. Development, implementation, and evaluation of an integrated multidisciplinary objective structured clinical examination (OSCE) in primary health care settings within limited resources. Med Teach 2016;38:272-9.

29. Vagholkar K. OSCE as a summative assessment tool for undergraduate students of surgery – Our experience. Indian J Surg 2019;81:412.