An Assessment of the Correlation between Tests of Clinical Competence and Tests of Cognitive Knowledge amongst Nigerian Resident Doctors in Surgery

Abstract

Background: Assessment of clinical competence involves the assessment of cognition and assessment of clinical performance (behaviour in practice). The limitations of the traditional long case examination (LCE) in the assessment of clinical performance led to its replacement with the objective structured clinical examination (OSCE) by many institutions. **Aims:** To determine and compare the abilities of the OSCE and LCE to predict candidates' performance in the tests of cognitive knowledge in the fellowship examination of the National Postgraduate Medical College of Nigeria in the Faculty of Surgery. **Materials and Methods:** The results of the OSCE, LCE, written papers, picture tests (PTs), vivas, and the total clinical score (TCS) of surgical residents who took part in the fellowship examination over six consecutive examination periods were compared by using the Pearson's correlation coefficient. A P-value less than .01 was considered as significant. **Results:** The OSCE had a weak but statistically significant positive correlation (.175) with the LCE. Both the OSCE and LCE had similar correlations with the total written papers (TWP) and PTs. The viva had a higher correlation with the OSCE than the LCE. The TCS when compared with either the OSCE or LCE alone had a higher correlation with most of the tests of cognitive knowledge. **Conclusion:** Neither the OSCE nor the LCE showed any superiority over the other in terms of the ability to predict performance in the tests of cognition. The TCS appears superior to either the OSCE or the LCE as a predictor of the candidates' overall knowledge of surgery. Therefore, both the OSCE and the LCE should be retained as part of the examination.

Keywords: *Long case, Nigerian National Postgraduate Medical College, objective structured clinical examination, resident doctors, surgery*

Abstrait

Contexte:
L'évaluation de la compétence clinique implique l'évaluation de la cognition et l'évaluation de la performance clinique (comportement dans la pratique). Les limites de l'examen de cas long traditionnel (LCE) dans l'évaluation de la performance clinique ont conduit à son remplacement par l'examen clinique objectif structuré (ECOS) par de nombreuses institutions.

Objectifs:
Déterminer et comparer les capacités de l'ECOS et du LCE à prédire les performances des candidats aux tests de connaissances cognitives lors de l'examen de bourse de la faculté de chirurgie du Nigerian National Postgraduate Medical College.

Méthodes:
Les résultats de l'ECOS, du LCE, des documents écrits, des tests d'image, des vivas et du score clinique total (TCS) des résidents en chirurgie qui ont participé à l'examen de bourse sur six périodes d'examen consécutives ont été comparés en utilisant le coefficient de corrélation de Pearson. Une valeur de p inférieure à 0,01 était considérée comme significative.

Résultats:
L'ECOS avait une corrélation positive faible mais statistiquement significative (0,175) avec le LCE. L'ECOS et le LCE avaient des corrélations similaires avec le nombre total d'épreuves écrites (TWP) et les tests d'images. La soutenance avait une corrélation plus élevée avec l'ECOS que le LCE. Le score clinique total (TCS), comparé à l'ECOS ou au LCE seul, présentait une corrélation plus élevée avec la plupart des tests de connaissances cognitives.

Abdulrazzaq O. Lawal[1],
Abdul-Hakeem O. Abiola[2],
Muhammad Y. M. Habeebu[3],
Rufus W. Ojewola[1],
Kehinde H. Tijani[1]*

[1]Department of Surgery, College of Medicine, University of Lagos, Lagos, Nigeria, [2]Department of Community Health, College of Medicine, University of Lagos, Lagos, Nigeria, [3]Department of Radiation Biology, Radiotherapy and Radiography, College of Medicine, University of Lagos, Lagos, Nigeria

Address for correspondence:
Prof. Kehinde Habeeb Tijani, Department of Surgery, College of Medicine, University of Lagos, Lagos, Nigeria.
E-mail: Habeeb_tijani@yahoo.com

How to cite this article: Lawal AO, Abiola AHO, Habeebu MY, Ojewola RW, Tijani KH. An assessment of the correlation between tests of clinical competence and tests of cognitive knowledge amongst Nigerian resident doctors in surgery. J West Afr Coll Surg 2020;10:12-6.

Conclusion:

Ni l'OSCE ni le LCE n'ont montré de supériorité sur l'autre en termes de capacité à prédire la performance dans les tests de cognition. Le TCS semble supérieur à l'ECOS ou au LCE en tant que prédicteur des connaissances globales des candidats en chirurgie. Par conséquent, l'ECOS et le LCE devraient être retenus dans le cadre de l'examen.

Mots clés: *Examen clinique objectif structuré, cas long, médecins résidents, chirurgie, Collège National Nigérian de Médecine Postdoctorale*

Introduction

Assessment of clinical competence involves the assessment of cognition and assessment of performance (behaviour in practice).[1] Presently, there appears to be no single test that can adequately assess all aspects of clinical competence. Assessing cognition deals with knowledge and its application, whereas the assessment of behaviour deals with the assessment, under controlled conditions, of competence in practice. The traditional written papers including the oral viva[1,2] assess cognitive knowledge, whereas the clinical examinations assess competence or performance.

Although there is little doubt that the tasks undertaken in a traditional long case examination (LCE) resemble real-life situations, its limitations are well known, especially its reported low reliability.[3,4] The objective structured clinical examination (OSCE) was introduced with the aim of overcoming these limitations. Since it was initiated by Harden, the OSCE has gained worldwide popularity with many institutions and examination bodies preferring it to and abandoning the LCE. Although examination bodies in North America and most of Europe have completely abandoned the LCE, there appears to be some reluctance in the West African subregion to do the same.[5] Studies on OSCE in our environment have been sparse and have mainly looked at the perception of students towards the examination.[6,7] Although different studies worldwide have tried to assess the correlation between the tests of clinical performance (OSCE or LCE) and other tests of clinical competence, most of these studies have had one major limitation—inability to compare the OSCE and LCE in the same set of candidates at the same examination sitting—because most training institutions and examination bodies simply abandoned the LCE and replaced it with the OSCE.[7-11] The Faculty of Surgery of the National Postgraduate Medical College of Nigeria was one of the earliest faculties in the college to fully introduce the OSCE into its Part I (intermediate) fellowship examinations in 2010. The faculty however retained the LCE while the OSCE replaced the short cases and one of the oral viva examinations, thereby giving the authors a unique opportunity to assess the OSCE and LCE in the same candidates in the same examination.

The objective of this study was to assess the ability of the candidates' scores in the tests of clinical performance to predict their performance in the tests of cognitive knowledge by determining the correlation between the clinical examination (OSCE and LCE) scores and the scores of the different tests of cognitive knowledge in surgical residents who sat for the Part 1 examinations in surgery.

Materials and Methods

This was a retrospective study that involved surgical resident doctors who fully participated in the Part I examination of the National Postgraduate Medical College of Niger between November 2013 and May 2016 (six examinations). At each examination, there were 20 OSCE stations consisting of two history stations, three clinical examination stations, and at least one performance station where candidates were tasked with demonstrating a specific skill on a mannequin. Each history, clinical examination, and performance station was manned by a minimum of two examiners using a checklist to score the candidates, and was followed by another station where the candidates were required to answer open-ended questions related to the previous clinical activity. Each OSCE station lasted 5 min. The other OSCE stations included questions on surgical instruments and patient management problems. For the LCE, each candidate was assessed by a minimum of three examiners. Each candidate was given 30 min to clerk a real-life patient and 15 min to be assessed by the examiners. The non-clinical examination (tests of cognitive knowledge) included one paper of essays and multiple-choice questions (MCQs) on principles of surgery, a second paper of essays and MCQs on operative surgery, a 20-min oral viva, and a 30-min picture test (PT). The oral viva was divided into equal segments that consisted of operative surgery and pathology (10 min) and principles of surgery (10 min). Every candidate was expected to answer questions on at least two topics in each segment. In order to reduce bias on the choice of the topics, each candidate was required to pick the topic via a ballot system. In the PTs, the candidates were showed different pictures such as clinical lesions and radiographs followed by related open-ended questions. The PT spanned 30 min and involved 20 different pictures with the candidate having 90 s to review a picture and answer the accompanying questions.

The LCE, written papers, and oral viva were marked over 20 marks, each using a close marking system with the implications that the highest possible mark was 12, whereas the lowest was 8. The maximum marks for the OSCE and PT were 100 each.

After a formal application was made to the college registrar, the authors were granted access to the hard copies of the spreadsheets containing the marks of all the candidates who sat for the examinations. Relevant data were collected from the sheets.

The SPSS version 20 software was used in data analysis. The correlation between the OSCE, LCE, and the other different forms of the examination were compared using the Pearson's correlation coefficient. A *P*-value less than .01 was considered as significant.

Results

A total of 420 candidates sat for the examination (55 candidates in 2013, 147 in 2014, 144 in 2015, and 74 in 2016), but two candidates who had some of their marks missing were excluded from the final analysis. Table 1 shows the correlation coefficients between OSCE, LCE, and the other forms of assessment. The OSCE and LCE had a correlation of .175 (P < .001) between them [Figure 1]. When compared with the LCE, the OSCE had a higher correlation with the written papers on surgical principles, whereas the LCE had a higher correlation with the operative surgery. Both the OSCE and LCE had similar correlations with the total written papers (TWP) and the PTs. The correlation between the oral viva and the OSCE was three times its correlation with the LCE. The total clinical score (TCS)—the sum of the marks in the LCE and OSCE—when compared with either the OSCE or LCE alone had a higher correlation with the principles paper, the TWP, and the PTs. The OSCE had a higher correlation than the TCS with the oral viva.

Discussion

The OSCE had a weak but statistically significant positive correlation (.175) with the LCE. Other studies have also found a weak (<.5) correlation between the two forms of clinical examination. Tijani *et al.* found a weak correlation of .37 between performance in the final year OSCE and the in-course assessment LCE in 612 undergraduate students in surgery.[12] Bakhsh *et al.* also found a correlation of .40 in 904 undergraduates in internal medicine.[13]

The reasons for the relatively lower correlation found in our study are unclear. The fact that these other studies were on undergraduate students, whereas the present study was on surgical residents with at least 3 years of postgraduate training may be contributory.

Moreover, in each of those two studies, even though the same set of students were involved, the interval between the LCE and the OSCE was between 12 and 24 months.[12,13] The implication of this is that although the physical identity of each

	POS	OPS	TWP	PT	OV	LCE	OSCE
LCE							
r	*.194	*.211	*.242	*.183	*.137	-	*.175
P-value	<.001	<.001	<.001	<.001	.005	-	<.001
OSCE							
r	*.306	.087	.227	*.193	*.392	*.175	
P-value	<.001	.075	<.001	<.001	<.001	<.001	
TCS							
r	*.330	*.280	*.299	*.233	*.314		
P-value	<.001	<.001	<.001	<.001	<.001		

Table 1: Pearson's correlation coefficients between the different forms of examination

POS: principles of surgery, OPS: operating surgery, TWP: total written papers, PT: picture test, OV: oral viva, LCE: long case examination, OSCE: objective structured clinical examination, TCS: total clinical score, *r*: Pearson's correlation coefficient. *Statistically significant
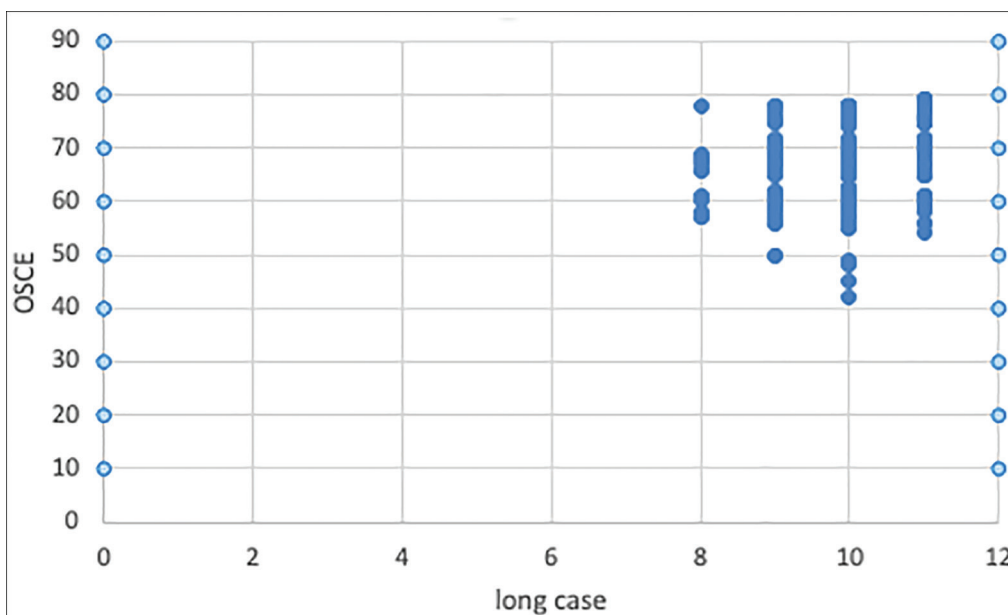

Figure 1: Scatter plot of the OSCE and LCE scores

student would remain unchanged, it would be reasonable to assume that the level of knowledge and clinical competence of the individual student would have significantly changed during this period. Wallenstein *et al.* also found a strong correlation (.6) between the performance in the OSCE of new intake residents and their average performance in other tests of clinical competence 18 months later.[14] It is also not clear whether the system of marking in the LCE affected the strength of correlation with the OSCE. With close marking in LCE, the minimum score was 8 and the maximum 12, which meant that only five different marks were possible. As a result, most of the points on the scatter diagram overlapped as shown in Figure 1. It may be conjectured that if the LCE scoring were to be opened up, a clearer picture of the strength of correlation between the two could be obtained.

Both the OSCE and LCE had positive but weak correlations with all the other components of the examination. These correlations were however statistically significant. Kirton and Kravitz also found an average correlation of .6 between the OSCE and the written papers in undergraduate students, whereas Al Rushood and Al-Eisa also found a correlation of .44 in undergraduate students in paediatrics.[15,16] Komasawa *et al.* however found no statistically significant relationship between the OSCE and TWP.[17] In the present study, the OSCE had a correlation of about .23 with the TWP, which was statistically significant. The OSCE (when compared with the LCE) had a stronger correlation with the written papers on surgical principles and the oral viva, whereas the LCE had a better correlation with the operative surgery component of the written papers. Both the OSCE and LCE had similar correlations with the TWP and PTs. Our study was therefore unable to demonstrate any superiority of the OSCE over the LCE or vice versa in terms of their correlation with the candidates' performance in the non-clinical forms of assessment. These findings were similar to the findings of Johnson and Reynard who did not find any consistent correlation between the OSCE and other forms of (non-clinical) examination in their study on undergraduate students in emergency medicine.[18] However, they were in contrast with a study by Tijani *et al.* that reported a consistent superiority of the OSCE over the LCE in terms of its correlation with other forms of non-clinical assessment in the undergraduate examinations in surgery.[12]

The literature is sparse on studies comparing OSCE and LCE in the same set of candidates, as almost all undergraduate and postgraduate examination bodies that adopted the OSCE simply abandoned the LCE at the same time. Indeed LCE can be said to have a face validity and authenticity as it is difficult to contest the logical assumption that the clinical method rehearsed in the course of a long case is the way that a good clinician should practise. It is also most likely true that an experienced examiner could by active questioning during an LCE differentiate between a good and a bad candidate. The decision to replace the LCE with OSCE has however been critically challenged.[19] Wass *et al.* in the UK reported a study of 214 candidates, who all had two LCEs alongside a 20-station

OCSE.[20] Their study reported reliability of .84 and .72 for LCE and OSCE. They concluded that given the same amount of time, the LCEs are in terms of reliability no worse and no better than OSCEs in assessing clinical competence. In the present study, the TCS (a combination of the scores of the LC and OSCE) when compared with LCE or OSCE alone had a stronger positive correlation with all the other components of the Part 1 examinations except the oral viva, which had a slightly higher correlation with the OSCE. The TCS therefore appears to be a better predictor of the candidates' performance in the non-clinical tests, thereby making it a good predictor of the candidates' overall knowledge of surgery.

The study had some limitations. First, it was retrospective. Second, it did not take into consideration variables like previous experience with the examination because this may have an effect on the performance of the candidates as some of them were repeating the examination.

In conclusion, none of the methods of assessing clinical competence (OSCE and LCE) was able to independently show any consistent superiority over the other in terms of correlation with the candidates' performance in the tests of cognition. It is therefore recommended that the TCS should be retained as part of the summative assessment of resident doctors in surgery.

Conflicts of interest

There are no conflicts of interest.

References

1. Miller GE. The assessment of clinical skills/competence/performance. Acad Med 1990;65:S63-7.
2. Al-Wardy NM. Assessment methods in undergraduate medical education. Sultan Qaboos Univ Med J 2010;10:203-9.
3. Smee S. Skill based assessment. BMJ 2003;326:703-6.
4. Newble D. Techniques for measuring clinical competence: Objective structured clinical examinations. Med Educ 2004;38:199-203.
5. Dong T, Saguil A, Artino AR Jr, Gilliland WR, Waechter DM, Lopreaito J, *et al*. Relationship between OSCE scores and other typical medical school performance indicators: A 5-year cohort study. Mil Med 2012;177:44-6.
6. Ameh N, Abdul MA, Adesiyun GA, Avidime S. Objective structured clinical examination vs traditional clinical examination: An evaluation of students' perception and preference in a Nigerian medical school. Niger Med J 2014;55:310-3.
7. Nasir AA, Yusuf AS, Abdur-Rahman LO, Babalola OM, Adeyeye AA, Popoola AA, *et al*. Medical students' perception of objective structured clinical examination: A feedback for process improvement. J Surg Educ 2014;71:701-6.

8. Simon SR, Volkan K, Hamann C, Duffey C, Fletcher SW. The relationship between second-year medical students' OSCE scores and USMLE step 1 scores. Med Teach 2002;24:535-9.

9. Bang JB, Choi KK. Correlation between clinical clerkship achievement and objective structured clinical examination (OSCE) scores of graduating dental students on conservative dentistry. Restor Dent Endod 2013;38:79-84.

10. Eftekhar H, Labaf A, Anvari P, Jamali A, Sheybaee-Moghaddam F. Association of the pre-internship objective structured clinical examination in final year medical students with comprehensive written examinations. Med Educ Online 2012;17. doi:10.3402/meo.v17i0.15958.

11. Adeyemi SD, Omo-Dare P, Rao CR. A comparative study of the traditional long case with the objective structured clinical examination in Lagos, Nigeria. Med Educ 1984;18:106-9.

12. Tijani KH, Giwa SO, Abiola AO, Adesanya AA, Nwawolo CC, Hassan JO. A comparison of the objective structured clinical examination and the traditional oral clinical examination in a Nigerian university. J West Afr Coll Surg 2017;7:59-72.

13. 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;30:555-7.

14. Wallenstein J, Heron S, Santen S, Shayne P, Ander D. A core competency-based objective structured clinical examination (OSCE) can predict future resident performance. Acad Emerg Med 2010;17:S67-71.

15. Kirton SB, Kravitz L. Objective structured clinical examinations (OSCEs) compared with traditional assessment methods. Am J Pharm Educ 2011;75:111.

16. Al Rushood M, Al-Eisa A. Factors predicting students' performance in the final pediatrics OSCE. PLoS One 2020;15: e0236484.

17. Komasawa N, Terasaki F, Nakano T, Kawata R. Relationships between objective structured clinical examination, computer-based testing, and clinical clerkship performance in Japanese medical students. PLoS One 2020;15:e0230792.

18. Johnson G, Reynard K. Assessment of an objective structured clinical examination (OSCE) for undergraduate students in accident and emergency medicine. J Accid Emerg Med 1994;11: 223-6.

19. Teoh NC, Bowden FJ. The case for resurrecting the long case. BMJ 2008;336:1250.

20. Wass V, Jones R, Van der Vleuten C. Standardized or real patients to test clinical competence? The long case revisited. Med Educ 2001;35:321-5.