

Assessment of MBBS medical students using (OSCE), comparing the modified borderline group method (MBGM), the modified Angoff method (MAM) and the holistic method of 50%

Mohamed Ahmed A/Gadir Elimam Ounsa¹; MD, Elsadig Yousif Mohamed²; MD, Elmuntasir Taha³; MD

- ^{1.} Department of obstetrics and gynaecology, Faculty of Medicine, The National Ribat University, Khartoum, Sudan.
- ^{2.} Department of community medicine, College of Medicine, Majmaah University, Saudi Arabia
- ^{3.} Department of paediatrics, faculty of medicine, The National Ribat University, Khartoum, Sudan.

ABSTRACT

Background: The final clinical examination in obstetrics and gynaecology is usually done by the traditional long obstetric case in the majority of medical schools worldwide. In the past few decades a move towards the (OSCE) has taken place. The objective of this study was to assess the performance of the (MBBS) medical students in the (OSCE) final obstetrics and gynaecology clinical examination, comparing three different standard setting methods: the MBGM, the MAM) and the holistic method of (50%).

Methods: This was a descriptive study conducted in the Faculty of Medicine in the National Ribat University, Sudan. One hundred and three students who constituted all the medical students in the final MBBS clinical examination were enrolled in this study. The students were examined by the (OSCE) in 10 stations. Thirty examiners participated in the examination. The cut score was determined by the (MBGM). The same stations were also scored by another different group of 16 examiners; who acted as judges to determine the cut score by the (MAM). We took the (50%) as a fixed reference cut score for this (OSCE) examination. A comparison was made between the three standard setting methods.

Result: The No. of students who passed the clinical examination by the (MBGM) and (MAM) was the same 85(82.5%) passed the examination by both methods. Results showed that 85 (82.5%) and 94(91.3%) passed by the (MBGM) and the fixed (50%) respectively. Eighty five (82.5%) and 94 (91.3%) of the students passed the exam by (MAM) and the fixed cut off score of (50%) respectively

Conclusion: There is no significant difference in the number of students who passed/failed the clinical examination between (MBGM), (MAM) and the holistic method of (50%). (MAM) can be used in the clinical examination.

Key words: OSCE, standard setting, Modified Borderline Group Method, Modified Angoff Method, the Holistic Method

Introduction:

For many decades, the traditional long obstetric case (LOC) has been the method of medical students' assessment in the clinical examination in most medical schools worldwide. Unfortunately, the conventional clinical and practical examinations are beset with several problems and have been shown to have serious limitations in terms of validity and reliability [1, 2], though, has the advantage of assessing the patient in a holistic approach [3]. In the past few decades the objective structured clinical examination (OSCE) became very popular and many medical schools

started to adopt it as an alternative to the long obstetric case. It is believed to answer the limitation of the long obstetric case [2, 4, 5, and 6].

In the (OSCE), which was first introduced by Hardeen in 1975 [6], the students move round a number of stations, which are structured to test different clinical skills appropriate to them. These skills are patients' examination, history taking and communication besides technical and procedural skills as long as interpretation of laboratory investigations in the light of a patient's problem. At these stations the candidate may meet a patient, real or simulated, a manikin, or less commonly patient data. A set amount of time, usually five or six minutes, is spent at each station; occasionally more complex double time stations may be included. Usually, the same basic circuit is followed by all candidates, although data stations may be changed in order to maintain confidentiality. Old, children and severely ill patients may also be change to prevent fatigue [6,5,1,7,8].An advantage of the (OSCE) over assessments that use real patients is that the patients are standardized across examinees and thus the patient problems are essentially equivalent and examinee responses and scores are comparable [9]. The (OSCE) is a valuable assessment tool to provide outcome measures on the competency performance and evaluate program effectiveness [10].

In order for the (OSCE) to be a suitable alternative to the traditional examination, the criteria of validity, reliability, objectivity and practicability or feasibility are to be ensured [7]. However despite its advantages, the (OSCE) has disadvantages of being exhausting, labour, time and resource intensive, stressful and a strong anxiety-producing exam [11, 12, 13, and 14].The objective of this study was to assess the performance of the MBBS medical students in the (OSCE)final obstetrics and gynaecology clinical examination in the National Ribat University, comparing three different standard setting methods: the modified borderline group method (MBGM), the modified Angoff method (MAM) and the holistic method of (50%).

Research method:

The current study was descriptive study design conducted in the Faculty of Medicine in the National Ribat University, Khartoum, Sudan. All the medical students in the final MBBS obstetrics and gynaecology clinical examination were enrolled in this study. One hundred and three students were examined in 10 interactive stations representing 10 different skills in both obstetrics and gynaecology, each station was given 7 minutes. Thirty examiners, with at least 10 years' experience, participated in the assessment. In the border line group method, the global rating in the check list is categorized in four categories: clear pass, borderline (satisfactory), border line (unsatisfactory), and clear fail. The examiner gave the student marks according to his/her real performance. The examiner depending on his/ her experience was asked to judge clinically and then rate whether student irrespective of his /her real score, is a clear pass, clear fail, or borderline (satisfactory/unsatisfactory). The real scores of the border line students were averaged in each station and the sum of the averages of all stations was taken as the cut off score of the examination. In the modified Angoff method,the same stations were also scored by a different group of 16 examiners; obstetricians/gynaecologists of long standing experience in the profession who acted as judges referees. Each judge was asked to define the criteria of the border line (minimally competent) student and then determined the probability that this student could answer the questions in each station correctly. The judge's estimates were averaged in each station and the sum of the averages of all stations was taken as the cut off score of the examination. We took also the fixed (50%) as cut score for this (OSCE) examination.

SPSS version 20 was employed for data analysis. Analysis was done by simple frequency. In comparison between two groups, the Chi Square was used to test significance and p-value <0.05 was considered significant. The ethical approval was obtained from the ethical committee of the National Ribat University. The data was kept confidential and used only for the purpose of this study.

Results:

Table (1) shows the performance of the students when the cut off score was determined by the (MBGM) and (MAM). Equal number of students 85(82.5%) passed the examination by both standard setting methods. The difference was not significant. Table (2) shows the performance when the cut off score was determined by (MBGM) and the fixed (50%). 85(82.5%) passed by the (MBGM) and 94(91.3%) passed by the fixed (50%). There was no statistically significant difference between the two methods. The results showed that 85 (82.5%) and 94 (91.3%) of the students passed the exam by (MAM) and the fixed cut off score of (50%) respectively. The difference was not significant as shown in table (3).

Table (1)
The performance of the students: comparison between the (MBGM) and (MAM)

Parameter	Results		Total	p
	Passed	Failed		
OSCE(MBGM)	85(82.5%)	18(17.5%)	103(50%)	1.000
OSCE(MAM)	85(82.5 %)	18(17.5 %)	103(50%)	

Table (2)
The performance of the students: comparison between the (MBGM) and fixed cut off score of (50%) CS(50%).

Parameter	Results		Total	p
	Passed	Failed		
(MBGM)	85(82.5%)	18(17.5%)	103(50%)	1.000
CS (50%)	94(91.3%)	9(8.7%)	103(50%)	

Table (3)
The performance of the students; comparison between the (MAM) and the fixed cut off score of(50%)(CS= (50%).

Parameter	Results		Total	p
	Passed	Failed		
(MAM)	85(82.5%)	18(17.5%)	103(50%)	0.065
CS (50%)	94(91.3%)	9(8.7%)	103(50%)	

Discussion:

Methods for student assessment in medical education have changed over the past few decades. A movement from a standard of pen-and-paper tests of knowledge and facts toward a more complex system of evaluation has occurred. Today medical students are tested on knowledge, attitudes, and skills across multiple settings and methods, which are often triangulated to reach summative decisions. Current educational and assessment strategies include problem-based learning, computer simulations, faculty global ratings and checklists, standardized patients, and team-based learning [7].

Student assessment is an integral part of educational programs. Since it drives students' learning and highlights significant goals and objectives of the course, teachers and administrative pay careful attention to its different parts. However, standard setting is an area in the field of assessment which is not dealt with so frequently [15]. Standard-setting consists of establishing performance standards, which are statements describing the knowledge and skills that students must attain if they are to perform at a specified performance level (basic, proficient, advanced, and so forth), and it consists of setting cuts scores [16]. A standard, also known as the minimum pass level, separates the competent students from those who are not [15, 17].

The performance of the students when the cut off score was determined by the (MBGM) and the (MAM) was determined. The results showed an equal number of students passed the examination when the cut off score was decided by either method. In comparison, Kaufman et al reported that the Angoff and border line methods were shown to provide reasonable and defensible approaches to standard setting and are practical to be used by non-psychometricians in medical schools, Boursicot KA et al found the border line more consistent in determining the pass score than the Angoff [18, 19]. It is reported that (MBGM) was the more credible and defensible method of standard setting, and appeared well suited to a small-scale OSCE [20, 21]. Gunilla and Peter reported in their study

comparing a version of Angoff method (extended Angoff) which is a two-step procedure and the border line group method that both methods provided reasonable and trustworthy approaches to standard setting [22]. Carlson et al reported that the border line group method offered a slightly more reliable cut score when compared to the standard set by the Angoff method, but was more challenging to implement and both methods demonstrated sufficient inter-rater reliability which is statistically significant [23]. The (MBGM) is an examinee referenced criterion and (MAM) is a test or an item referenced criterion and both are criterion (absolute) referenced types of standard setting methods that are most appropriate for tests of competence, where the purpose is to establish that the examinees know enough for a particular purpose. These include final or exit examinations [17].

The performance of the students when the cut off score was determined by the (MBGM) and the fixed percentage of (50%) was assessed. Though, the number of students who passed by the fixed percentage was more, but the difference was not statistically significant. In the fixed percentage standard setting method, the cut off score of the examination is determined by judges prior to the examination [17]. or, may be already determined by the regulatory body. Traditionally, most medical schools used this sort of standard setting in their assessment and the cut off score applied is usually (50%). This method is easy to apply and best suited to situations where there is a desire to identify a certain number of the best (or worst) examinees [17]. However, it has serious disadvantages; first, the standard for the test will vary from one group to another depending on the ability of the group and second, some candidates will pass and others will fail regardless of how correctly they performed [5].

The results here with the (MBGM) are the same as the performance of the students when the cut off score was determined by the (MAM) and the fixed percentage of (50%). In contrast to our findings, Kaufman et al when comparing between the fixed percentage (holistic) method and the relative, the Angoff and the border line methods, found a highly inconsistent results with a wide divergence between it and the relative method while the results of the Angoff and borderline were comparable [18]. However, our results showed that the difference between the holistic (fixed) method and the modified Angoff method was not significant. Our findings showed that the Angoff method could be used in the clinical examination as it is used in the written examination. Others supported this finding [24]. The modified Angoff method, which is a variation of Angoff method, is the most widely studied standard-setting method [16]. It may be a feasible way to set a standard passing score with less time consumed and more independent rather than group work by instructors [25]. Being reported by others, the Angoff was found to be credible and have good inter-rater reliability and moderate test-retest reliability [26, 4, and 27].

In this study we involved sixteen examiners to act as judges. The number is usually less than that. We think this increased number of ratters will maintain the variance among the rating low and hence we will obtain more accurate cut off score.

Today, there is a plethora of standard setting methods [8, 15, 26, 28], but none is a gold standard and hence, no absolute cut off score that could be generated from any one [8, 29], so, any method chosen may not matter unless it fits for the purpose of the assessment. However, when a certain standard setting method is chosen for a particular examination, the passing score has to be reliable and defensible [4].

Conclusion:

There is no significant difference in the number of students who passed/failed the clinical examination between (MBGM), (MAM) and the holistic method of (50%). (MAM) can be used in the clinical examination.

References:

1. Erfanian F, Khadivzadeh T. Evaluation of midwifery students' competency in providing intrauterine device services using objective structured clinical examination. *IJNMR/Summer 2011; Vol. 16 (3) 191-196 ANY PAGES?*
2. Newble D. Techniques for measuring clinical competence: objective structured clinical examinations. *Medical Education 2004; Volume 38 (2):199-203*
3. Ponnampertuma GG, Karunathilake IM, McAleer S, Davis MH. The long case and its modifications: a literature review. *Med Educ. 2009 Oct; 43(10):936-41.*
4. Morrison H, McNally H, Wylie C, McFaul P and Thompson W. The passing score in the Objective Structured Clinical Examination. *Medical Education 1996; Volume 30, Issue 5:345-348*
5. Pell G and Roberts T. E. Setting standards for student assessment. *International Journal of Research & Method in Education 2006; Vol. 29 (1):91-103*
6. Harden RM, Stevenson M, Downie W W, Wilson G.M. Assessment of clinical competence using objective structured Examination. *British Medical Journal 1975, 1, 447-451*

7. Turner JL, Dankoski ME. Objective structured clinical Exams: A Critical review. *Fam Med*. 2008; 40(8):574-8.
8. Barman A. Standard setting in student assessment: is a defensible method yet to come?. *Ann Acad Med Singapore* 2008 Nov; 37(11):957-63.
9. Brannick MT, Erol-Korkmaz HT & Prewett M. A systematic review of the reliability of objective structured clinical examination scores. *Medical Education* 2011; 45: 1181–1189.
10. Short MW, Jorgensen JE, Edwards JA, Blankenship RB, Roth BJ. Assessing intern core competencies with an objective structured clinical examination. *J Grad Med Educ*. 2009 Sep; 1(1):30-6.
11. Brand HS1, Schoonheim-Klein M. Is the OSCE more stressful? Examination anxiety and its consequences in different assessment methods in dental education, *Eur J Dent Educ*. 2009 Aug; 13(3):147-53.
12. A Barman. Critiques on the objective structured clinical examination. *Ann Acad Med Singapore* 2005; 34: 478-82
13. Iqbal M, Khizar B, Zaidi Z. Revising an objective structured clinical examination in a resource-limited Pakistani Medical School. *Education for Health* 2009; Volume 22, issue 1: 2009.1-9.
14. Pierre RB, Wierenga A, Barton M, Branday JM, Christie CDC. Student evaluation of an OSCE in paediatrics at the University of the West Indies, Jamaica, *Medical Education* <http://www.biomedcentral.com/1472-6920/4/22>
15. Hejri SM, Jalili M. Standard setting in medical education: fundamental concepts and emerging challenges. *MJIRI* 2014; Vol.28 (34):1-6
16. Brandon PR. Using Test Standard-Setting Methods in Educational Program Evaluation: Addressing the Issue of how good is good enough. *Journal of multidisciplinary evaluation (JMDE: 3)*. <http://evaluation.wmich.edu/jmde/> 1-29.
17. Norcini JJ. Setting standards on educational tests. *Med Educ*. 2003 May; 37(5): 464-9.
18. Kaufman DM, Mann KV, Muijtjens AM, van der Vleuten CP. A comparison of standard-setting procedures for an OSCE in undergraduate medical education. *Acad Med*. 2000 Mar; 75(3):267-71.
19. Boursicot KA, Roberts TE, Pell G. Using borderline methods to compare passing standards for OSCE at graduation across three medical schools, *Med Educ*. 2007 Nov; 41(11):1024-31.
20. Humphrey-Murto S, MacFadyen JC. Standard setting: a comparison of case-author and modified borderline-group methods in a small-scale OSCE. *Acad Med*. 2002 Jul; 77(7):729-32.
21. Wood TJ, Humphrey-Murto SM, Norman GR. Standard setting in a small scale OSCE: a comparison of the Modified Borderline-Group Method and the Borderline Regression Method. *Adv Health Sci Educ Theory Pract*. 2006 May; 11(2):115-22.
22. Näsström G and Nyström P. A comparison of two different methods for setting performance standards for a test with constructed-response items. *Practical Assessment Research & Evaluation* 2008; Vol. 13 (9): 1-12
23. Carlson J, Tomkowiak J, Knott P. Simulation-based examinations in physician assistant education: A comparison of two standard-setting methods. *The Journal of physician assistant education association*. 2010, 21(2):7-14.
24. Livingston SA and Zieky MJ. *Passing scores: A manual for setting standards of performance for education and occupational tests*. 1982. Educational testing service
25. Senthong V, Chindaprasirt J, Sawanyawisuth K, et al. Group versus modified individual standard-setting on multiple-choice questions with the Angoff method for fourth-year medical students in the internal medicine clerkship. *Adv Med Educ Pract*. 2013:195-200.
26. George S, Haque MS and Oyeboode F. Standard setting: Comparison of two methods
From: <http://www.biomedcentral.com/1472-6920/6/46>
27. Carlson J, Tomkowiak J, Stilp C. Using the Angoff method to Set defensible cutoff scores for standardized patient performance evaluations in PA education. *J Physician Assist Educ* 2009; 20(1):15-23
28. Norcini J, McKingley D.W. Standard setting. In: *A practical guide for medical teachers*. Churchill Livingstone 2009. P 311-317
29. Ben David MF. AMEE Guide No. 18: Standard setting in student assessment. *Medical Teacher* 2000; Vol. 22 (2): 120-130.