

นิพนธ์ต้นฉบับ

A comparative study of the 1993 Objective Structured Clinical Examination scores among three medical programs, Faculty of Medicine, Chulalongkorn University*

Boonnart Laisnitsarekul**

Kitpramuk Tantayaporn*** Jiruth Sriratanaban****

Laisnitsarekul B, Tantayaporn K, Sriratanaban J. A comparative study of the 1993 Objective Structured Clinical Examination scores among three medical programs, Faculty of Medicine, Chulalongkorn University. *Chula Med J* 1996 Apr;40(4): 289-298

This descriptive research aimed to study the raw data of the 1993 OSCE comprehensive examination for sixth year medical students from the conventional program, the MESRAP program and the CTPB program. Other objectives were to compare the OSCE scores among the three programs, to find the OSCE reliability and to find correlations between OSCE and MCQ scores. The study population was 148 sixth year medical students of the Faculty of Medicine, Chulalongkorn University in the 1993 year. There were 97 medical students from the conventional program, 39 medical students from the Medical Education for Students in Rural Areas Project (MESRAP) and 12 medical students from the Community-Targeted Problem-Based curriculum (CTPB). The OSCE was divided into 45 stations; 23 procedure stations and 22 question stations. Time used in each station was 3 minutes. In the procedure station,

* In the name of the 1993 Comprehensive Examination Committee, Faculty of Medicine, Chulalongkorn University.

** Medical Education Unit, Faculty of Medicine, Chulalongkorn University.

*** Department of Obstetrics and Gynecology, Faculty of Medicine, Chulalongkorn University.

**** Department of Preventive and Social Medicine, Faculty of Medicine, Chulalongkorn University.

the examiner used a checklist in direct observation. The score at each station was 10 points. After examination, the committee collected the data and calculated the reliability of the OSCE by Cronbach's α coefficient and other statistics such as frequency, total score, maximum score, minimum score, mean, standard deviation, percentage and they also compared the OSCE scores among the other programs by use of the F test and t-test.

The authors found that from 450 total scores, the maximum score was 372.9 (74.58%), the minimum score was 264.8 (52.96%), the mean was 332.21 (66.44%). The conventional medical program students, averaged higher OSCE scores than the MESRAP and CTPB program students, significantly different at level 0.01 ($p < .01$). The MESRAP students averaged higher OSCE scores than the CTPB students, significantly different at level 0.01 ($p < .01$). The reliability of the OSCE was 0.66. When determining correlations between the OSCE scores and the MCQ scores of the 1993 comprehensive examination, both OSCE and MCQ had correlation ($r = 0.61$) significantly at level 0.01 ($p < .01$) in every programs.

Key words: OSCE, Comprehensive examination, Conventional, MESRAP, CTPB, Reliability.

Reprint request: Laisnitsarekul B, Medical Education Unit, Faculty of Medicine, Chulalongkorn University, Bangkok 10330, Thailand.

Received for publication. January 10, 1996.



บุญนาท ลายสนิทเสรีกุล, กิจประมุข ตันตยาภรณ์, จิรุตม์ ศรีรัตนบัลล์. การเปรียบเทียบ
คะแนนสอบปฏิบัติการทางคลินิกแบบปรนัย ปี 2536 ระหว่างหลักสูตรแพทยศาสตรบัณฑิต
3 หลักสูตร ของคณะแพทยศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย. จุฬาลงกรณ์เวชสาร 2539
เมษายน; 40(4): 289-298

การวิจัยเชิงบรรยายนี้มีวัตถุประสงค์ เพื่อศึกษาคะแนนสอบปฏิบัติการทางคลินิกแบบปรนัย
วิชาเวชศาสตร์ทั่วไป ปีการศึกษา 2536 ของนิสิตแพทย์ ชั้นปีที่ 6 คณะแพทยศาสตร์ จุฬาลงกรณ์
มหาวิทยาลัย ทั้งหลักสูตรโครงการปรกติ โครงการส่งเสริมการศึกษาแพทย์สำหรับชาวชนบท และ
โครงการผลิตแพทย์แนวใหม่ ตลอดจนเปรียบเทียบคะแนนสอบระหว่างหลักสูตร การหาค่าความ
เที่ยงของการสอบ รวมทั้งการหาความสัมพันธ์ระหว่างคะแนนสอบปฏิบัติการทางคลินิกแบบปรนัย
กับคะแนนสอบข้อสอบปรนัย ประชากรที่ศึกษาเป็นนิสิตแพทย์ชั้นปีที่ 6 คณะแพทยศาสตร์
จุฬาลงกรณ์มหาวิทยาลัย รวม 148 คน แยกเป็นนิสิตโครงการปรกติ 97 คน นิสิตโครงการส่งเสริม
การศึกษาแพทย์สำหรับชาวชนบท 39 คน นิสิตโครงการผลิตแพทย์แนวใหม่ 12 คน ดำเนินการ
สอบโดยจัดเป็น 45 สถานี แบ่งเป็นสถานีหัตถการ 23 สถานี และสถานีคำถาม 22 สถานี ใช้เวลา
สถานีละ 3 นาที ในสถานีหัตถการผู้ประเมินจะสังเกตโดยตรง และให้คะแนนตามแบบเลือกตรวจ
แต่ละสถานีมีคะแนนเต็มเท่ากับ 10 คะแนน ภายหลังจากสอบคณะกรรมการรวบรวมข้อมูล คำนวณ
หาค่าความเที่ยงด้วยสูตรสัมประสิทธิ์อัลฟา และค่าสถิติต่างๆ ได้แก่ ความถี่ คะแนนเต็ม คะแนน
สูงสุด คะแนนต่ำสุด คะแนนเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ร้อยละ ตลอดจนเปรียบเทียบคะแนน
สอบด้วยสูตรสถิติ F test และ t -test

ผลการวิจัยพบว่า จากคะแนนเต็ม 450 คะแนน นิสิตทำคะแนนได้สูงสุดเท่ากับ 372.9
คะแนนหรือร้อยละ 74.58 ต่ำสุดเท่ากับ 264.8 หรือร้อยละ 52.96 คะแนนเฉลี่ยเท่ากับ 332.21
หรือร้อยละ 66.44 โดยนิสิตโครงการปรกติโดยเฉลี่ยได้คะแนนสอบปฏิบัติการทางคลินิกแบบปรนัย
สูงกว่าโครงการอื่น อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และนิสิตโครงการส่งเสริมการศึกษาแพทย์
สำหรับชาวชนบท ทำคะแนนสอบปฏิบัติการทางคลินิกแบบปรนัยสูงกว่านิสิตโครงการผลิตแพทย์
แนวใหม่ อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 สำหรับความเที่ยงของการสอบครั้งนี้ เท่ากับ 0.66
เมื่อกำหนดหาความสัมพันธ์ระหว่างคะแนนสอบปฏิบัติการทางคลินิกแบบปรนัยและคะแนนข้อสอบ
ปรนัยของการสอบวิชาเวชศาสตร์ทั่วไป ประจำปีการศึกษา 2536 พบว่า มีความสัมพันธ์กันในทุก
หลักสูตร ($r=0.61$) อย่างมีนัยสำคัญทางสถิติที่ระดับ .01

The Objective Structured Clinical Examination (OSCE) was designed as an objective method of assessing a student's clinical skills where the areas tested and the evaluation criteria were determined in advance from course content and objectives. The OSCE was purported to have the potential for testing a wide range of knowledge and skills, and could evaluate a large number of students in one examination period. During an OSCE, students rotate around a series of timed stations. At some stations they are asked to take a focused history or perform some aspect of physical examination. These stations, where an observer is asked to score a student's performance, were called procedure stations. At other stations, students may be asked to answer short questions, to interpret patient data, or to record findings. Checklist or rating forms and scoring systems for each station were prepared in advance by the consensus of the program planners and instructors.⁽¹⁾

The OSCE provided a more valid examination than the traditional approach to clinical examinations. The examiners could decide in advance what was to be tested. The examiners could control not only the content but also the complexities of the examination - more straightforward cases for junior students, more advanced cases for senior students. The emphasis could be moved away from testing factual knowledge to testing a wide range of skills, including history taking. The examination was not only more valid but also more reliable. The variables of the examiner and the patient were to a large extent removed. With the large number of stations, a large sample of the student's skills could be

tested. The OSCE had the advantage that it could be used with large numbers of students. In a recent surgical final examination in Dundee, Scotland, 110 students were tested in one morning. The examiner was used efficiently in that he was used only for those parts of the examination, such as observing a student taking history or examining a patient, where there could be no substitute. The fact that the examiner had a clear checklist makes it possible to use more junior examiners.⁽²⁾

In the past, assessment of student competence was determined only by written examination. The clinical examination was regarded by many examiners as of key importance in the assessment of a student's competence to practice medicine and the cornerstone in qualifying examinations. For the 1987 academic year 8 Thai medical schools recommended improving the assessment of clinical competence. Thus the Faculty of Medicine of Chulalongkorn University started to use OSCE in the comprehensive examination in that year⁽³⁾. Since then there has been no studies on the effectiveness of OSCE for the comprehensive examination. Thus the 1993 Comprehensive Examination Committee determined to study the OSCE during the 1993 academic year.

Objectives

1. To study the OSCE scores of the 1993 Comprehensive Examination for:-
 - 1.1 All sixth year medical students
 - 1.2 Conventional program students
 - 1.3 MESRAP students
 - 1.4 CTPB students

2. To determine the reliability of OSCE
3. To compare the OSCE scores among curriculum
4. To find correlations between OSCE and MCQ for the 1993 comprehensive examination

Definition

Objective Structured Clinical Examination⁽⁴⁾ OSCE is a performance assessment. It does not just assess what students know but what they can do. The OSCE assesses both process and product. It assesses not only the completeness and quality of a clinical history taken from a patient by a student, but also the student's technique in taking the history.

Materials and Method

Population: There were 148 sixth year medical students from three programs. The

number of medical students from the conventional, MESRAP and CTPB programs were 97, 39 and 12 students, respectively.

Stations: There were 45 stations (23 procedure stations and 22 question stations) from 11 clinical departments. There were 8 stations from the Department of Medicine, 7 stations from the Department of Surgery, 8 stations from the Department of OB-GYN, 5 stations from the Department of Pediatrics, 4 stations from the Department of Preventive and Social Medicine, 3 stations from the Department of Orthopedic Surgery, 2 stations from the Department of Ophthalmology, 2 stations from the Department of Otolaryngology, 2 stations from the Department of Anesthesiology, 2 stations from the Department of Psychiatry, and 2 stations from the Department of Forensic Medicine. [Table 1]

Table 1. Type of stations, the 1993 Comprehensive Examination OSCE

Department	Procedure station	Question station	Total
Medicine	7	1	8
Surgery	6	1	7
OB-GYN	5	3	8
Pediatrics	2	3	5
Preventive Med.	-	4	4
Orthopedic.	-	3	3
Ophthalmology	1	1	2
ENT	-	2	2
Anesthesiology	2	-	2
Psychiatry	-	2	2
Forensic Med.	-	2	2
Total	23	22	45

Methods:

1. The committee planned two half-days for assessing the students' clinical competence. The OSCE would be six hours long and comprised of two types of station. There would be 23 three minute procedure stations and 22 three minute question stations.

2. Each student performed for 3 minutes at each station. The student would start at a different point in the station circle and, at each time-change signal, move on to the next station. Thus the student who started at station 1 moved to station 2, the student at station 23 moved to station 1 and so on.

3. The examiner used a checklist to check off the items done by the examinee. Checklists contained a list of items (generally history questions/findings, examination maneuvers, and/or information to be communicated to the patient, depending upon the type of station), and the examiner's task was simply to check off the items done by the examinee. The examiners directly scored examinee performance. The total score for each station was ten points.

4. After the examination, the committee collected data and analysed the examinee's scores. They calculated the reliability of OSCE by Cronbach's α coefficient formula, compared the OSCE scores among the three programs by ANOVA and found correlations between OSCE

and MCQ by the Pearson product moment correlation coefficient formula.

Results

1. From 450 OSCE total scores, 148 sixth year medical students received a maximum score of 372.9 (74.58%) and a minimum score was 264.8 (52.96%). The mean score was 332.21 (66.44%) and the standard deviation was 20.31 (4.06%). When identified for each program, the 97 conventional students received a maximum score of 372.9 (74.58%) and a minimum score of 296.8 (59.36%). The mean score was 340.30 (68.06%) and the standard deviation was 15.99 (3.20%). The 39 MESRAP students received a maximum score of 365.3 (73.06%) and a minimum score of 290.2 (58.04%). The mean score was 321.25 (64.25%) and the standard deviation was 16.94 (3.39%). The 12 CTPB students received a maximum score of 329.4 (65.88%) and a minimum score was 264.8 (52.96%). The mean score was 302.38 (60.48%) and the standard deviation was 16.73 (3.35%). [Table 2] When the program mean scores were compared, the mean score of the conventional program was highest and significantly different from the other programs. The mean score of the MESRAP program was higher than the mean score of the CTPB program, significantly at level 0.01 ($p < .01$). [Table 3]

Table 2. Scores of OSCE, the 1993 Comprehensive Examination.

Program	n	Maximum	Minimum	Mean	S.D.
All students	148	372.9 (74.58%)	264.8 (52.96%)	322.21 (66.44%)	20.31 (4.06%)
Conventional	97	372.9 (74.58%)	296.8 (59.36%)	340.30 (68.06%)	15.99 (3.20%)
MESRAP	39	365.3 (73.06%)	290.2 (58.04%)	321.25 (64.25%)	16.94 (3.39%)
CTPB	12	329.4 (65.88%)	264.8 (52.96%)	302.38 (60.48%)	16.73 (3.35%)

Table 3. The comparison of means, the 1993 Comprehensive Examination OSCE.

Program		
Conventional vs MESRAP vs CTPB	F = 39.974	p<.01
Conventional vs MESRAP	t = 6.128	p<.01
Conventional vs CTPB	t = 7.637	p<.01
MESRAP vs CTPB	t = 3.317	p<.01

2. From 11 clinical subjects, the 148 sixth year medical students recieved 100% scores in two subjects; Anesthesiology and Forensic Medicine. They also recieved OSCE scores between 90-99% in seven subjects; Medicine, Surgery,

Pediatrics, Obstetrics and Gynecology, Orthopedic Surgery, Ophthalmology and Psychiatry. For only two subjects, Oto-laryngology and Preventive and Social Medicine, did the students recieve OSCE scores between 78-85%. [Table 4]

Table 4. OSCE scores of each subject.

Department	Total score	Maximum	Minimum	Mean	S.D.
Medicine	80	77 (96.25%)	47.5 (59.38%)	63.03 (78.79%)	5.12 (6.40%)
Surgery	70	63.7 (91.00%)	41.6 (59.43%)	54.44 (77.77%)	4.37 (6.24%)
OB-GYN	80	76.9 (96.13%)	42.2 (52.75%)	64.82 (81.03%)	6.27 (7.84%)
Pediatrics	50	46 (92.00%)	14.5 (29.00%)	32.00 (64.00%)	6.74 (13.48%)
Preventive Med.	40	31.5 (78.75%)	13 (32.50%)	20.96 (52.40%)	3.38 (8.45%)
Orthopedic.	30	29 (96.67%)	12.5 (41.67%)	22.18 (73.93%)	3.26 (10.87%)
Ophthalmology	20	19.5 (97.50%)	11 (55.00%)	17.25 (86.25%)	1.83 (9.15%)
ENT	20	17 (85.00%)	6.5 (32.50%)	12.30 (61.50%)	2.05 (10.25%)
Anesthesiology	20	20 (100.0%)	14.5 (72.50%)	18.17 (90.85%)	1.2 (6.00%)
Psychiatry	20	18 (90.00%)	6 (30.00%)	12.75 (63.75%)	2.65 (13.25%)
Forensic Med.	20	20 (100.0%)	4 (20.00%)	14.31 (71.55%)	3.55 (17.75%)
Total	450	-	-	-	-

3. The reliability of the OSCE for the 1993 comprehensive examination, as calculated by Cronbach's α coefficient, equals 0.66. When the correlation between the OSCE scores and the MCQ scores was determined, both OSCE and MCQ correlated significantly at level 0.01 ($p < .01$) in every program. [Table 5]

Table 5. The correlation between OSCE and MCQ, the 1993 Comprehensive Examination.

Program	Correlation coeff.	
Three programs	0.612	$p < .01$
Conventional	0.423	$p < .01$
MESRAP	0.643	$p < .01$
CTPB	0.717	$p < .01$

Discussion

OSCE may be used for a number of different purposes. It can be used as a pass/fail examination to check whether students have achieved the standard appropriate to pass on to the next section of the course or to graduate. It can also be used to identify areas in which students have achieved competence and to find areas where further assessment is required to explore the competences in more detail. Van der Vleuten & Swanson⁽⁵⁾ summarized the results of OSCE studies in the years 1978-1990. They found that the test lengths varied from 1 to 16 hours and the time allotted to each station varied from 6 to 40 minutes. The number of stations varied from 4 to 40 stations and the reliability coefficient varied from 0.39 to 0.93. The OSCE of the 1993 comprehensive examination was fairly good and the number of stations was higher but the time allotted to each station was less than in their studies. For the typical OSCE assessing hands-on clinical skills with patients, at least four hours of testing time were necessary. At the sixth Ottawa conference on medical education,⁽⁶⁾ they reported that the OSCE with standardized patients had already been used for Medical Licensure in North America. NBME of the U.S.A. or MCC of Canada had done a lot of research in this area. The cut of point was 16 in 20 stations (criterion referenced). An average 17% of the candidates failed the OSCE. The obtained reliability was 0.61 (Cronbach alpha) and the reliability for a 500-item MCQ was 0.56 (Cronbach alpha).

The reliability and the validity (correlation with MCQ) of the OSCE for the 1993

comprehensive examination were good. But the correlation between OSCE and MCQ of the conventional program was less than the other programs. The hypothetical reasons may be (1) the conventional medical students learnt the contents of medical subjects better than the MESRAP and the CTPB programs and (2) In medical practice, the conventional medical students practiced less than the other programs.

The Faculty of Medicine of Chulalongkorn University has used OSCE since 1987 but had not analysed and published on its effectiveness. This study was the first paper on OSCE of the Faculty of Medicine of Chulalongkorn University. This was to determine if the Chulalongkorn University Medical School was on the right path regarding OSCE examinations. It was determined not only to be a reliable examination for skills required for medical students, but also an excellent preparation for medical students to reach international standards.

Summary

This descriptive research aimed to study raw data of the 1993 OSCE comprehensive examination for the sixth year medical students from the conventional program, the MESRAP program and the CTPB program. Other objectives were to compare OSCE scores among programs, to find the reliability of OSCE, and to find correlations between OSCE and MCQ. The authors found that from 450 total scores, the maximum score was 372.9 (74.58%), the minimum score was 264.8 (52.96%), the mean was 332.21 (66.44%). The conventional medical students averaged higher OSCE scores than the

MESRAP and CTPB students, significantly different at level 0.01 ($p < .01$). The MESRAP students had averaged higher OSCE scores than the CTPB students, significantly different at level 0.01 ($p < .01$). The reliability of the OSCE was 0.66. When determining correlations between the OSCE scores and the MCQ scores of the 1993 comprehensive examination, both OSCE and MCQ had significant correlations at level 0.01 ($p < .01$) in every program.

The OSCE may be an excellent examination for assessing the competency in the clinical performance which other examinations could not do. The reliability of the OSCE for the 1993 comprehensive examination was 0.66 and the correlation coefficient with the 1993 MCQ comprehensive examination was 0.61. It shown that the 1993 OSCE comprehensive examination was a good examination.

References

1. McMaster University. Evaluation Methods: A Resource Manual. Hamilton: Program for Educational Development, 1987: 31-2
2. Harden RM, Gleeson FA. Assessment of clinical competence using an objective structured clinical examination (OSCE). *Med Educ* 1979 Jan; 13(1): 41-54
3. Faculty of Medicine, Chulalongkorn University. Table of Specification for Comprehensive Examination. Bangkok: Division of Academic Affairs, 1993: 1
4. Harden RM. The OSCE - A 15 year retrospective. In: Hard IR, Harden RM, Marchais JD, eds. *Current Developments in Assessing Clinical Competence*. Montreal: Can-Heal Publications, 1992: 41
5. Van der Vleuten CPM, Swanson DB. Reliability issues in objective structured clinical examinations. In: Hard IR, Harden RM, Marchais JD, eds. *Current Developments in Assessing Clinical Competence*. Montreal: Can-Heal Publications, 1992: 66-78
6. Kengsakul K. Report of the Sixth Ottawa Conference on Medical Education, June 26-29, 1994, Toronto, Canada, 1994.