ORIGINAL RESEARCH

Bagcilar Med Bull 2021;6(1):7-13
**DOI:** 10.4274/BMB.galenos.2020.09.054

# Is Peer Assessment Reliable in Objectively Structured Clinical Examination?

## Objektif Yapılandırılmış Klinik Sınavda Akran Değerlendirmesi Güvenilir midir?

Hasan Dagmura[1], Emin Daldal[2], Hüseyin Bakır[3], Mehmet Fatih Daşıran[2], Zeki Özsoy[2], Osman Demir[4], İsmail Okan[3], Ertan Bülbüloğlu[3]

[1]University of Health Sciences Turkey, Kütahya Evliya Çelebi Training and Research Hospital, Clinic of Surgical Oncology, Kütahya, Turkey
[2]Gaziosmanpaşa University Faculty of Medicine, Department of General Surgery, Tokat, Turkey
[3]Gaziosmanpaşa University Faculty of Medicine, Department of General Surgery and Surgical Oncology, Tokat, Turkey
[4]Gaziosmanpaşa University Faculty of Medicine, Department of Biostatistics, Tokat, Turkey

## Abstract

**Objective:** The study was designed to evaluate the reliability of the peer assessment in the objectively structured clinical examination (OSCE) for the summative assessment of 4th grade students at the end of general surgery clerkship.

**Method:** The study was planned prospectively with the permission of the Dean of Medicine Faculty and approval of the ethics committee. The 6th grade students who were in the surgery rotation participated in the study as peer assessors (PA). Both peers and department of general surgery assessed the students. Pass/fail point was accepted as 60. The scores of OSCE and performance evaluation given by peers and faculty were compared statistically.

**Results:** Twenty-three students completed general surgery clerkship. Ten students (43.5%) were female. According to performance scores given by the faculty, 15 (65.2%) of the 23 students were successful, while all students were considered successful (having a grade of 60 or more) based on the scores of peer evaluation. There was a significant difference between the faculty members and PA with regard to the performance evaluation (p=0.008). The faculty members found five students (27.8%) successful in the OSCE (having a grade of 60 or more). However, ten students (43.5%) received a score of at least six from peer evaluation. Although there was a difference, it was not significant (p=0.063). Gender did not affect scoring in performance evaluation and OSCE application.

**Conclusion:** Although there was a difference between faculty members and peer evaluators in the performance evaluation, no difference was

## Öz

**Amaç:** Genel cerrahi kliniğinde 4. sınıf öğrencilerine "objektif yapılandırılmış klinik sınav" (OYKS) uygulanmasında akran değerlendirmesinin güvenilir olup olmadığını değerlendirmek ve öğrencinin servis içindeki performansının akran değerlendirmesi üzerine etkilerini araştırmak amacıyla bu çalışmayı tasarladık.

**Yöntem:** Çalışma, Tıp Fakültesi Dekanlığından ve etik kuruldan izin alınarak prospektif olarak planlandı. OYKS uygulamasına katılan 4. sınıf stajyer öğrenciler çalışma grubu olarak belirlendi. Altıncı sınıf öğrencileri ise akran değerlendirici olarak çalışmada yer aldı. Bu çalışma için akran değerlendiricilerin ve öğretim üyelerinin performans puanı vermeleri ve OYKS uygulamasına katılarak değerlendirmeleri istendi. Sonuçlar istatiksel yöntemlerle değerlendirildi.

**Bulgular:** OYKS uygulamasına katılan 23 öğrenciden 10'u (%43,5) kız öğrenciydi. Genel cerrahi stajından sorumlu öğretim üyeleri tarafından verilen ortalama performans puanlarına göre, 23 öğrenciden 15'i (%65,2) başarılı oldu. Akran değerlendiricilerinin performans puanı ortalamalarına göre ise tüm öğrencilerin başarılı olduğu görüldü. Öğretim üyeleri ile akran değerlendiricilerin verdikleri performans puanları arasında istatistiksel olarak anlamlı fark bulunmuştur (p=0,008). Öğretim üyeleri ve akran değerlendiriciler tarafından stajyerlere verilen OYKS puanları karşılaştırıldı. Öğretim üyelerinin verdikleri puana göre, beş (%27,8) öğrenci altı veya daha yüksek puan alarak başarılı oldu. Akran değerlendiricilerin puanlamasına göre ise on (%43,5) öğrenci altı veya daha yüksek puan alarak başarılı oldu. Buna göre akran değerlendiricilerin OYKS'de daha fazla öğrenciyi başarılı bulduğu görüldü. Ancak iki grubun verdiği puanlarda

**Address for Correspondence:** Hasan Dagmura, University of Health Sciences Turkey, Kütahya Evliya Çelebi Training and Research Hospital, Clinic of Surgical Oncology, Kütahya, Turkey
**E-mail:** hassen@hacettepe.edu.tr **ORCID:** orcid.org/0000-0003-2289-5514 **Received:** 05.09.2020 **Accepted:** 30.09.2020

**Cite this article as:** Dagmura H, Daldal E, Bakır H, Daşıran MF, Özsoy Z, Demir O, Okan İ, Bülbüloğlu E. Is Peer Assessment Reliable in Objectively Structured Clinical Examination? Bagcilar Med Bull 2021;6(1):7-13.

©Copyright 2021 by the Health Sciences University Turkey, Bagcilar Training and Research Hospital
Bagcilar Medical Bulletin published by Galenos Publishing House.

## Abstract

observed in OSCE. In conclusion, OSCE assessment by peer evaluators is reliable.

**Keywords:** Education, objective structured clinical examination, peer assessment, reliability, surgery

## Öz

istatistiksel olarak anlamlı fark yoktu (p=0,063). Cinsiyetin performans değerlendirmesinde ve OYKS uygulamasında puanlamayı etkilemediği tespit edildi.

**Sonuç:** Çalışmamızda performans değerlendirmesinde öğretim üyeleri ile akran değerlendiriciler arasında fark olduğunu, OYKS uygulaması değerlendirmesinde ise farklılık olmadığını bulduk. Sonuç olarak akran değerlendirmesinin güvenilir olduğunu düşünüyoruz.

**Anahtar kelimeler:** Akran değerlendirmesi, cerrahi, eğitim, güvenilirlik, objektif yapılandırılmış klinik sınav

## Introduction

Physicians are expected to be active, lifelong learners nowadays. Medical education is constantly changing to enable physicians to develop such an understanding (1). The education of medical students should prepare them to cope with future problems and ensure that they have the necessary skills to become active and self-directed learners rather than passive recipients of the information. Therefore, a transition from time-based to competency-based education occurred in medical education and the medical curriculum was revised accordingly. This also led to the revision of student assessment tools (2). Planning an outcome-based education is as important for student motivation as ensuring quality in educational programs because it defines learning outcomes and forms a basis for curriculum decisions in contemporary education. Lecturers should establish the most appropriate assessment and evaluation system to evaluate the expected learning outcomes of the students. The selected assessment tools should be valid, reliable, and practical, and have an appropriate impact on student learning. An assessment profile should be produced for each student, aiming at the learning outcomes the student is expected to achieve (3).

Objectively structured clinical examination (OSCE) was defined by Harden et al. (4) in 1975 to evaluate the learning outcomes required from students effectively and objectively, to ensure the standardization of examinations, and to regulate the examination process. The OSCE is designed as a new assessment tool that allows candidates' clinical skills, attitudes, problem-solving skills, and knowledge practices to be assessed in an exam. It is a performance-based exam consisting of multiple stations. At each station, the examiners evaluate high-level thinking skills according to preformed blueprint. OSCE is widely used in the assessment of practical skills in medicine. It has several advantages: it is an objective tool in assessing the student, it has a pre-structured question and answer format, and it can assess knowledge, skills, and attitudes in the clinic. It is a complex and time-consuming task to prepare and perform OSCE smoothly. It aims to provide the standardization and to reduce the number of variables that may affect performance evaluation. Therefore, in a well-designed OSCE, students' grades should only be influenced by their performances (5).

Peer assessment (PA) is considered an important tool in medical education. In this process, students of similar levels evaluate the learning outcomes of their peers, which also contributes to their training. PA, defined as peers' evaluations of their friends' achievements, learning outcomes or performances, is increasingly used in modern medical education (6). It can be used to encourage students to participate in educational activities and clarify the assessment criteria, improve team performances, or identify individual efforts (7). Students who perform their internship in the same clinical setting during their medical education have the advantage of observing their peers while performing their duties. Therefore, it was stated that peers had a higher chance of observing each other's performances than faculty members (8).

PA in OSCE stations that can be used in the evaluation of almost all professional competency areas can be an effective model in the OSCE. This study aimed to evaluate the reliability of the PA in the OSCE that was performed by our clinic without interruption and to investigate the effects of student performances in the service on the PA.

## Materials and Methods

Members of the General Surgery Department, Faculty of Medicine, Gaziosmanpaşa University decided to conduct an OSCE in the evaluation of fourth-year interns studying at the faculty of medicine. For this purpose, the faculty members made the necessary preparations (training,

observation, and literature survey) and initiated the OSCE in our clinic in the 2017-2018 academic year.

### Ethical Committee Approval

Permission no: 17713155-100 was obtained from the Deanship of the Faculty of Medicine to conduct studies on the OSCE for the fourth-year students and to ensure the participation of the sixth-year students (peer assessors) for the PA. The study was designed prospectively after obtaining the permission of Non-interventional Clinical Research from the Ethical Committee of Gaziosmanpaşa University, Faculty of Medicine (registration number: 19-KAEK-149 date: 28/05/2019).

The study was conducted with fourth-year students who had been receiving their undergraduate education in the academic year 2018-2019. The working group involved 23 fourth-year students who participated in the OSCE.

### General Surgery Clerkship

In the fourth-year of medical education, the students receive eight-week training in the general surgery clinic as divided into four groups. The training involves three main objectives of learning: knowledge, skills, and attitude. The students receive not only theoretical and practical courses in the field of surgery, but also skills such as vascular access, nasogastric catheter application, Foley catheter placement, and suturing, and attitudes such as communication and professionalism.

### General Surgery Clerkship Assessment

At the end of the clerkship, the students' success is assessed by a multiple-choice exam, the OSCE, the evaluation of the portfolio (a list of interventional procedures requested during the clerkship), and the evaluation score of the instructor to the students' attitude during the eight-week clerkship period. The students who score at least 60 out of 100 are considered successful at Gaziosmanpaşa University, Medical Faculty. Four student groups are trained in our clinic every year. Each group consists of an average of 25 students. The success assessment is as follows: the multiple-choice theoretical exam accounts for 30 points, the portfolio evaluation for 10 points, the assessment of professionalism (faculty member's opinion about each student's attitudes during the clerkship period) for 10 points, and the OSCE for 50 points. The portfolio evaluating the students' practical and communication skills and volunteer participation was conducted by all faculty members three days before the end of the clerkship. The grade indicated the student's performance evaluation by the professors. The grade was given out of 10 points for each student. Students who scored at least six were found successful for the clerkship. We also asked the peer assessors (in this case 6$^{th}$ grade students or interns) to give an opinion grade to the fourth-year clerks as they had participated in many procedures during their educations. For this purpose, fourth-year clerks were evaluated one by one together with the faculty members and the opinion grade was given. The multiple-choice exam was held in the following day.

### Performance Evaluation

During the clerkship, faculty members in charge of the general surgery education observed the students in various settings such as the wards, the operating room, and during the lectures. They evaluated the students' performances according to their behaviors in various domains such as knowledge, attitude, communication, professionalism, and volunteering. In the same way, peer assessors (interns/six-grade students) also observed the students during their rotations of general surgery internship and formed opinions about them.

### Before the OSCE

All students who were going to take the OSCE were informed by the faculty members about it at the beginning of the clerkship. We established five stations for the OSCE in the general surgery clinic. Students were informed about the application of the OSCE as well as the exam area. The rules to be followed during the OSCE were explained.

### The OSCE Stations

Five stations were created in the OSCE application (Figure 1). During the creation of these stations, the exam topics were classified and the students drew lots for the OSCE questions in the last week of education period. The questions consisted of following topics for each station; the first station included basic topics such as fluid electrolyte, hemostasis, shock, surgical infections, and trauma. The second station included topics from the field of oncologic surgery such as the esophagus, stomach, and breast cancer. The questions at the third station were related to gastrointestinal diseases such as diverticulitis, acute appendicitis, and hemorrhoids. The questions at the fourth station were prepared from the field of endocrine surgery such as thyroid, parathyroid, and adrenal. The questions at the last station addressed the skills that needed to be acquired during a surgical clerkship, such as suturing, obtaining patient consent, and abdominal examination. Ten questions were prepared for each group, five questions out of 50 were drawn by lot and asked in the exam.

### The OSCE Application

We asked the students to manage the patient through prepared scenarios at the four stations of the OSCE. Management of hyperparathyroidism, perianal abscess, soft tissue infection, and soft tissue sarcomas were the questions determined by lot. We also asked the students to perform an abdominal examination on a model at the practical station. Each station was rated out of ten points. Eight points were given for information and management, two points were given for the smoothness of the presentation order and the self-confidence of the student.

Before the OSCE, we conducted an evaluation with the responsible faculty members. There was one member of the department of general surgery and one of the peer assessors at each station. The faculty members at each station briefed the peer assessors on the question and assessment sheet and informed them about the procedure.

The OSCE application started with a ringtone after the preparations were completed. Students were given five minutes at each station. When the bell rang by surgery resident, the students changed the station. All 23 students completed the OSCE. The faculty members and the peer assessors evaluated the students at each station. Only the grades of the faculty members for students were taken into account as pass/fail point.

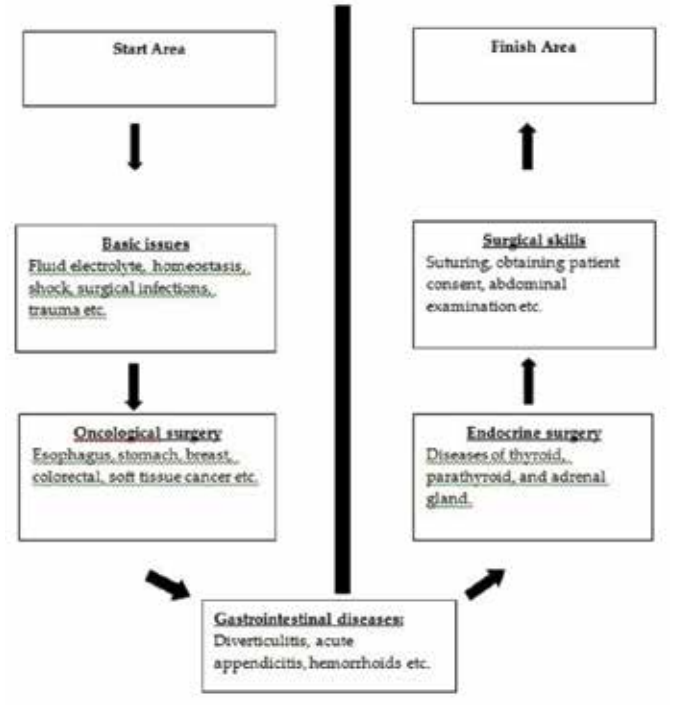



**Figure 1.** Schematic view of the OSCE stations
*OSCE: Objectively structured clinical examination*

### After the OSCE

At the end of the OSCE, we evaluated the exam with all the faculty members and the peer assessors. We compared the performance and the OSCE scores of the peer assessors with those of the faculty members for the same student to examine whether the performance assessment of the peer assessors affected the outcome. The average of the scores obtained at the five stations was taken and students with a mean score of six or above were considered successful. We also investigated whether the gender of the students was important in the performance and OSCE assessments of faculty members and peer assessors.

### Statistical Analysis

Data were expressed in frequency and percentage. The McNemar test was used to compare the categorical data between the groups. Pearson correlation coefficient was used for correlation between variables. A p-value <0.05 was considered significant. Analyses were performed using SPSS 19 (IBM SPSS Statistics 19, SPSS inc. an IBM Co., Somers, NY).

## Results

Twenty-three students participated in the OSCE, ten (43.5%) were female students. Based on the average performance evaluation scores given by the faculty members responsible for the general surgery clerkship, 15 of the 23 students (65.2%) were successful. All students were successful according to the peer assessors' performance evaluation score averages. There was a significant difference between the faculty members and peer assessors regarding the performance evaluation scores (p=0.008) (Table 1).

We evaluated the scores given to the students by the faculty members and the peers in the OSCE (Table 2). The faculty members found five students (27.8%) successful in the OSCE by getting a score of six or above. As for the peer assessors, ten students (43.5%) received a score of at least six from them. Therefore, the peer assessors found

**Table 1. Comparison of performance scores between faculty members and peer assessors**

| | Performance scores given by the faculty members (n=23) | Performance scores given by the peer assessors (n=23) | p |
|---|---|---|---|
| **<6** | 8 (34.8%) | 0 (0%) | 0.008 |
| **≥6** | 15 (65.2%) | 23 (100%) | |
| **Total** | 23 (100%) | 23 (100%) | - |

more students successful in the OSCE. However, there was no statistically significant difference (p=0.063) (Table 2). We analyzed the performance evaluation scores given by the faculty members concerning gender and found no statistically significant correlation (Table 3) (p=0.379). The peer assessors gave a performance score indicating that all the students were successful.

There was no significant correlation between the scores given by the faculty members and the gender of the students (p=0.618) (Table 4). Again, the gender of the students did not play a role in the OSCE scoring by the peer assessors (p=0.222) (Table 5).

**Table 2. Comparison of OSCE scores between faculty members and peer assessors**

| | OSCE scores given by the faculty members (n=23) | OSCE scores given by the peer assessors (n=23) | p |
|---|---|---|---|
| **<6** | 18 (78.3%) | 13 (56.5%) | 0.063 |
| **≥6** | 5 (21.7%) | 10 (43.5%) | |
| **Total** | 23 (100%) | 23 (100%) | - |

OSCE: Objectively structured clinical examination

**Table 3. Distribution of performance scores given by the faculty members by gender**

| Performance scores given by the faculty members | Female | Male | p |
|---|---|---|---|
| **<6** | 2 | 6 | 0.379 |
| **≥6** | 8 | 7 | |
| **Total** | 10 | 13 | - |

**Table 4. Distribution of OSCE scores given by the faculty members by gender**

| OSCE scores given by the faculty members | Female | Male | p |
|---|---|---|---|
| **<6** | 7 | 11 | 0.618 |
| **≥6** | 3 | 2 | |
| **Total** | 10 | 13 | - |

OSCE: Objectively structured clinical examination

**Table 5. Distribution of OSCE scores given by the peer assessors by gender**

| OSCE scores given by the peer assessors | Female | Male | p |
|---|---|---|---|
| **<6** | 4 | 9 | 0.222 |
| **≥6** | 6 | 4 | |
| **Total** | 10 | 13 | - |

OSCE: Objectively structured clinical examination

## Discussion

Assessment and feedback by peers are becoming a valuable and increasingly recognized method used to enhance the student experience in medical schools around the world. In addition, PA has the potential to help prepare students for their professional lives (9). An advantage of PA is that although teachers have only limited time to observe each student, students have more opportunities to observe each other (10,11). PA can both be reliable and valid and can provide an effective learning experience for students (12). In our study, we used PA to determine the development and success of each student during the eight-week training. For this purpose, we asked the faculty members and peer assessors to give a performance evaluation score to the students. We found that the peer assessors who monitored the interns during clinical clerkship period gave higher performance evaluation scores to all compared to those of faculty members. Although it has the potential for an accurate and valid assessment, factors such as reliability, interpersonal relationships, interests, inter-group interaction, and equivalence may influence the assessment (13). Our study included sixth-year students as peer assessors to minimize the effects such as intra-group interaction, personal interest, and friendship. However, we found that they gave high scores for the performance assessment. This may have been the result of their empathy with the interns as they were also students.

In addition, we found that peer assessors gave students higher scores in the OSCE performances. More people were successful in the OSCE; however, there was no statistical difference between the scores given by the peer assessors and those by the faculty members. A neutral and objective assessment is an important foundation of the OSCE. For this purpose, standardizing the scoring system and having the same questions for each candidate make the assessment easier and more reliable. Some evidence suggests that the training of assessors reduces the difference in scoring (5). Some studies show that peer assessors give high scores in the OSCE. On the other hand, there are also some studies indicating lower-scoring by peer assessors than that of faculty members, although not statistically significant (14-17). Despite the high scores given by the peer assessors in our study, the results were similar to those of the faculty members due to the standardized scoring system and the training they received about the implementation of the exam. In the performance assessment, the scores given by peer assessors were statistically significant. This may be due to the fact that the OSCE evaluates knowledge and skills

rather than attitude and behaviors such as communication, professionalism, and volunteering evaluated with performance assessment.

Some studies found that male assessors give higher scores to female students, albeit not significantly (18). This raises the question of whether male supervisors are softer to female students. However, it was stated that female students had better communication skills which might also have affected this result (19). Gender bias was not detected in most of the studies (20,21). We found no difference in opinion grades and the OSCE in relation to gender.

### Study Limitations

Our study has some limitations worth mentioning: it was on a small scale. Assessing more students and having more assessors would increase the reliability. In addition, a higher number of stations can increase the effective evaluation in the OSCE. In our study, the number of stations was low. However, the educational institutions with limited resources still can apply the OSCE to the target evaluation.

We found that the peer assessors in our study gave higher scores in both performance assessment and the OSCE, and the difference was significant between the peer assessors and the faculty members in performance evaluation. This could either be due to the more time and place shared with students allowing the peer assessors to evaluate them promptly or to the biased attitude the peers have about them. A standardized performance assessment would increase reliability eliminating this interaction. Although peer assessors gave higher scores to the interns in the OSCE, there was no statistical difference.

## Conclusion

As a conclusion, we think that PA can be performed safely in the OSCE with adequate training and a standardized scoring system.

### Statement of Licensing Committee and Institution

Permission no: 17713155-100 was obtained from the Deanship of the Faculty of Medicine to conduct this study. Ethical committee approval for this study was obtained from the Local Research Ethics Committee registered under the number 19-KAEK-149 date: 28/05/2019. All methods were performed in accordance with the relevant guidelines and regulations of the institution.

### Ethics

**Ethics Committee Approval:** Ethical committee approval for this study was obtained from the Local Research Ethics Committee registered under the number 19-KAEK-149 date: 28/05/2019.

**Informed Consent:** Informed consent was obtained.

**Peer-review:** Externally peer-reviewed.

### Authorship Contributions

Concept: İ.O., E.D., E.B., Design: H.D., E.D., H.B., Data Collection or Processing: E.D., M.F.D., O.D., Z.Ö., H.B., Analysis or Interpretation: E.D., O.D., İ.O., E.B., Literature Search: H.D., H.B., M.F.D., Z.Ö., O.D., Writing: İ.O., E.D., E.B., H.D., Z.Ö., M.F.D.

**Conflict of Interest:** No conflict of interest was declared by the authors.

**Financial Disclosure:** The authors declared that this study received no financial support.

## References

1. Rosof AB, Felch WC. Continuing medical education: a primer. Westport: Greenwood Publishing Group, 1992.
2. Tabish SA. Assessment methods in medical education. Int J Health Sci 2010;2(2):3-7.
3. Shumway JM, Harden RM, Association for Medical Education in Europe. AMEE Guide No. 25: The assessment of learning outcomes for the competent and reflective physician. Med Teach 2003;25(6):569-584.
4. Harden RM, Stevenson M, Downie WW, Wilson GM. Assessment of clinical competence using objective structured examination. Br Med J 1975;1(5955):447-451.
5. Khan KZ, Ramachandran S, Gaunt K, Pushkar P. The objective structured clinical examination (OSCE): AMEE guide no. 81. Part I: an historical and theoretical perspective. Med Teach 2013;35(9):e1437-e1446.
6. Lindblom-ylänne S, Pihlajamäki H, Kotkas T. Self-peer-and teacher-assessment of student essays. Active Learn High Educ 2006;7(1):51-62.
7. Speyer R, Pilz W, Van Der Kruis J, Brunings JW. Reliability and validity of student peer assessment in medical education: a systematic review. Med Teach 2011;33(11):e572-e585.
8. Somervell H. Issues in assessment, enterprise and higher education: The case for self-peer and collaborative assessment. Assess Eval High Educ 1993;18(3):221-233.
9. Cushing A, Abbott S, Lothian D, Hall A, Westwood OM. Peer feedback as an aid to learning-What do we want? Feedback. When do we want it? Now! Med Teach 2011;33(2):e105-e112.
10. Arnold L, Shue CK, Kalishman S, Prislin M, Pohl C, Pohl H, et al. Can there be a single system for peer assessment of professionalism among medical students? A multi-institutional study. Acad Med 2007;82(6):578-586.

11. Nofziger AC, Naumburg EH, Davis BJ, Mooney CJ, Epstein RM. Impact of peer assessment on the professional development of medical students: a qualitative study. Acad Med 2010;85(1):140-147.

12. English R, Brookes ST, Avery K, Blazeby JM, Ben-Shlomo Y. The effectiveness and reliability of peer-marking in first-year medical students. Med Educ 2006;40(10):965-972.

13. Norcini JJ. Peer assessment of competence. Med Educ 2003;37(6):539-543.

14. Mavis BE, Ogle KS, Lovell KL, Madden LM. Medical students as standardized patients to assess interviewing skills for pain evaluation. Med Educ 2002;36(2):135-140.

15. Burgess A, Clark T, Chapman R, Mellis C. Senior medical students as peer examiners in an OSCE. Med Teach 2013;35(1):58-62.

16. Chenot JF, Simmenroth-Nayda A, Koch A, Fischer T, Scherer M, Emmert B, et al. Can student tutors act as examiners in an objective structured clinical examination? Med Educ 2007;41(11):1032-1038.

17. Basehore PM, Pomerantz SC, Gentile M. Reliability and benefits of medical student peers in rating complex clinical skills. Med Teach 2014;36(5):409-414.

18. Schleicher I, Leitner K, Juenger J, Moeltner A, Ruesseler M, Bender B, et al. Examiner effect on the objective structured clinical exam-a study at five medical schools. BMC Med Educ 2017;17(1):71.

19. Casey M, Wilkinson D, Fitzgerald J, Eley D, Connor J. Clinical communication skills learning outcomes among first year medical students are consistent irrespective of participation in an interview for admission to medical school. Med Teach 2014;36(7):640-642.

20. Denney ML, Freeman A, Wakeford R. MRCGP CSA: are the examiners biased, favouring their own by sex, ethnicity, and degree source? Br J Gen Pract 2013;63(616):e718-e725.

21. 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(1):103.