Logistics behind the OSCE in undergraduate medical education and an experience at a Caribbean Medical School

Sireesha Bala Arja[1], Sateesh B. Arja[1], Jesse Ramey[1], Tarig Fadlallah Altahir Ahmed[1], Kumar Ponnusamy[1], Venkata Raghavendrarao Morusupalli[1]

Corresponding author: Dr Sireesha Bala Arja drbala@avalonu.org
Institution: 1. Avalon University School of Medicine
Categories: Assessment, Learning Outcomes/Competency, Clinical Skills

Received: 21/09/2018
Published: 16/10/2018

Abstract

Current interest regarding holistic assessments in medical education is high. This is especially the case for use during clinical rotations and post-graduate residency programs. Holistic assessments consist of the complete process of assessing students in history taking, data gathering, and physical examination of the patient. In addition to knowledge of the medical sciences, communication and patient interaction skills are a critical part of these drills. Holistic assessment strategy differs from objective structured clinical examinations (OSCE) as it is a contiguous process compared to OSCE which assesses in compartments. The compartmental approach of OSCE is still effective in providing if the examinations are conducted correctly. To date, OSCE is the most comprehensive student evaluation technique following clinical rotations. Conducting the OSCE and standardized patient assessments are challenging for medical schools with limited resources, but achievable with proper planning. Those resources include standardized patients, physicians, faculty members, and adequately equipped physical facilities. Both the standardized patients and examiners require training. In spite of the challenges, this program has observed that proper use of OSCE exams done with a logistical approach is feasible and beneficial in our medical teaching environment.

Keywords: OSCE; assessments; feedback; clinical skills; clinical rotations.

Introduction

The assessment of clinical competency in medical education is crucial for preparing tomorrow’s doctors. The underpinning theory behind competency-based medical education is medical graduates are expected to be fit to practice rather than only fit to pass assessments or examinations. In this context, holistic assessments given to the
medical students play a vital role to make sure that these graduates will become competent physicians. Medical graduates are expected to be committed to patient's safety and public health while maintaining the highest standards of professionalism. Currently, there is a lot of discussion regarding testing medical students with holistic assessments as contrasted with assessing skills in segments.

A holistic assessment is an assessment of the student's ability to effectively and professionally assess a patient, using appropriate communication and interpersonal skills. The assessment should include the authentic and complete process of taking a focused history including data gathering, combined with an accurate physical examination. The final step of this assessment consists of using this information to formulate a differential diagnosis and an appropriate management plan. The goal is to simulate patient clinical assessment under the same conditions the closer that will be encountered in practice after graduation.

Holistic assessment differs from objective structured clinical examinations (OSCE) station in that it assesses the complete process of patient care rather than isolated elements. Despite the limitations of OSCE’s segmental approach, it currently is the method of choice used to assess clinical students after clinical rotations. Provided the OSCEs are conducted in right manner, they do produce valid and reliable results. OSCE is a comprehensive and robust assessment assessing students’ competencies including medical knowledge, patient care, communication, interpersonal skills, professionalism, and advocacy skills of the students.

Background

Avalon University School of Medicine is located on the Caribbean island of Curacao of the Netherland Antilles. The clinical skills course curriculum during the basic sciences program on the island utilizes a standardized patient program (Cleland, Abe and Rethans, 2009). During the basic sciences, pre-clinical students receive two-hundred forty hours of clinical skills instruction provided by multiple instructors who are medical doctors.

As part of the curriculum, students are required to do a mandatory 72 weeks of clinical rotations after the basic sciences program. The university needs that students are required to clear the National Board of Medical Examiners (NBME) comprehensive basic sciences exam with a minimum score of 68% before moving on to clinical rotations. Students complete 48 weeks of core rotations and 24 weeks of elective rotations across the teaching hospitals in the USA. Forty-eight weeks of core rotations include Internal Medicine, Family Medicine, Obstetrics and Gynecology, Pediatrics, Psychiatry, and Surgery. Students can select the electives of their choice.

Summative assessment methods during basic sciences program include multiple choice questions, objective structured practical examinations (OSPE), oral examinations, and standardized patient-based (SP-based) assessments. In the clinical rotations, summative assessments and grading comprise of preceptor evaluations based on their observations of students during rotations, NBME shelf exams (MCQs- clinical knowledge and reasoning), and OSCE at the end of all core rotations.

Methods and Materials

We reviewed the literature available including Pubmed, EBSCO, and other resources on OSCE and analyzed the literature. In addition to this literature review, our own experiences in conducting OSCE for the two years were studied. We then compiled this information.
Discussion

1. Objective Structured Clinical Examinations (OSCE)

1.1 Purpose of the OSCE
OSCE is an objective method of assessment (Harden and Gleeson, 1979) to test the student’s ability to perform and accurately record the findings and observations derived from physical examinations and history taking. The objective of the OSCE exam is to develop the tools to measure the competency of trainees in history taking or data gathering, physical examination skills, communication skills, clinical acumen, and professionalism. Skills that can be assessed by OSCE are such as physical examination, communication skills, and areas like the critical appraisal of the evidence. Communication skills are assessed against a checklist of required behaviors like doctor introduces himself, listens, explains, responds, summarizes, and plan, etc. (Blaney and Dacre, 2017).

The OSCE as an assessment tool has the advantages of acceptable validity, reliability (Newble, 2002), and minimization of examiner bias (Batra and Norcini, 2016). Reliability of OSCE can be improved by students passing through the same and multiple numbers of stations and students observed by different examiners. Students should strictly go through the stations. All students are allotted the same number of stations and the same amount of time per station. Assessors are provided with checklists or global ratings. The clinical content of each station determines the validity of the OSCE.

OSCE can be used as both formative and summative assessments for both postgraduate and undergraduate levels. OSCE can be successfully used where the purpose of the test is to assess the trainee for a defined level of competence and can decide to pass/fail (Batra and Norcini, 2016). It can also be used as norm based-testing where the purpose of the assessment is to identify the trainees with higher than average ability. The other use of OSCE is to provide feedback to the trainee so that he/she can improve his competency levels in the future practice.

The validity of OSCE is evidenced by Miller’s pyramid (Miller, 1990) and student “shows how” level can be assessed by OSCE (Norcini and Troncon, 2016). There is evidence that communication skills, as assessed by a national OSCE examination, are related to patient complaints (Tamblyn et al., 2007)

1.2 Logistical considerations of OSCE
OSCE stations need the space to conduct the examination. We allotted ten rooms (10 stations). Each room has three chairs for the student, the examiner, and the standardized patient and a table. There is a sitting area for students to wait comfortably for their turn, arrangements for standardized patients and faculty.

Specific tools including computers, measuring tapes, torch lights, neurological hammers, nasal speculum, stethoscopes, otoscopes, ophthalmoscopes, and sphygmomanometers are required. Examination equipment for conducting exams also includes a timer bell, necessary furniture, and mannequins for simulation.

1.3 Specifications for the OSCE
It is essential to create a blueprint or table of specifications for the OSCE. The validity of any test depends mainly on the content. Our school follows rubrics for OSCE.

1.4 Human resources
We employ trained academic faculty, standardized patients, administrative staff, allied health professionals, and one IT specialist.
1.5 Timeline
In OSCE it is required to test the competency of students in physical examination, communication and interpersonal skills, and history taking at the same time. Students should be able to spend a minimum of 15 minutes per station. We do ten stations of OSCE with 15 minutes spent per station. Two-three minutes is allowed for rotating between stations. Ten students can be processed simultaneously.

1.6 Feedback to the trainees
At the end of the OSCE, the trainee receives a pass/fail decision to determine graduation and also receives the feedback from the examiner. Feedback is provided in the areas of strengths and weaknesses. Even though it is a summative assessment, feedback is provided to the trainee who will get benefited in the future practice. Feedback is the most important modulator of trainees’ learning behavior. The primary focus of feedback should not be the criticism of performance but for learning.

Qualitative data is based on the professional judgment of the examiners (written narrative comments from the examiners). Quantitative data is based on the performance of the trainee on checklists. It is recommended to have both qualitative and quantitative data to be on OSCE grading sheet. It is shown that global rating has shown more reliability than checklists (Hodges and McIlroy, 2003). But it is reassuring that scores by the checklists are correlating with global ratings (Norcini and Boulet, 2003). The institution can decide whether to use either global ratings or checklists based on the local context. Giving individual feedback is essential. Personal feedback is so critical especially for poorly performing trainees. The written comments from examiners should include both strengths and weaknesses. Providing feedback on the same day of the assessment is essential. Generally, feedback is provided during the post-examination review.

1.7 OSCE –passing criteria and standardization of examination procedures:
The different components that can be assessed in OSCE include history taking, physical examination skills, communication skills, clinical acumen including lab values interpretation and patient management, and professionalism. Either checklists or global ratings can be used. It is required to pass a minimum number of stations, and it is also expected to pass mandatory stations along with the minimum number of stations.

For practice-based exams like the OSCE, the borderline groups’ method, contrast groups method or borderline regression standard setting method can be used to determine a passing score for each station. In the borderline group method, a candidate’s performance on a station will be categorized as either borderline, competent or non-competent. The average of all borderline candidates’ scores is used to determine the passing score for that station (De Champlain and Revest, 2016). Students are required to pass a designated minimum number of stations. Some stations need mandatory passing.

2. Organizing the OSCE:

2.1 OSCE day
Trained standardized patients are made available. The OSCE day consists of 10 encounters or 10 stations. All evaluators have the evaluation grid for each station before the OSCE examination begins (Harden, 1990).

2.2 Training examiners: will be done before the OSCE exam in the form of workshops.

I. General information about the OSCE, including a definition, what it can assess, the process, the evidence supporting its use, etc.
II. Information specific to the OSCE for which the examiners are being trained including the purpose, the number, and type of stations being used, the length of each station, the nature of the cases, the nature of the scoring, etc.

III. The role of the examiner including knowing what to look for, how to complete the forms, minimizing intrusiveness and interference, etc.

IV. Practice scoring using videos (or volunteer students working through sample stations) supported by small group discussions.

The workshop should be interactive and based on a combination of didactic mini-lectures, small group exercises, and role play. Step 4, practicing scoring, should be the most substantial part of the workshop. It should also be augmented with written materials and periodic, brief refresher courses that can be done in person or online.

Just before an OSCE administration, a brief meeting with examiners is useful. This meeting should ensure that they are familiar with the practicalities of the administration, the station they will be observing, and the scoring rubric and form.

3. Quality assurance/evaluation of OSCE

3.1 Quality assurance of OSCE observers (examiners) and stations

For the evaluation of the examiner, we would recommend doing the following things. Compare the performance of all the students on this exam with all the examiners. Interview the students, standardized patients, and all examiners. Compare the performance of students on this exam with his/her past performance. Compare the performance of the students in the past with other examiners.

For reviewing the task or station, we would recommend doing the following things. Review the percentage of students taking action if it is a checklist. Review the average rating if it is a global rating. Consider the correlation between each task and total station and total test score or mean. If a student is not performing well on a task, it could be due to the content of the station or the task itself or rating forms itself or the observer or sometimes with the standardized patient. It could also be a problem with the original design of the assessment or blueprinting of assessment.

It is always good to use the following data for evaluation of the examinations. Descriptive statistics, including the mean and standard deviation for the total group and selected subgroups of interest, reliability of the scores, relationship of exam scores with other routinely collected local assessment data such as written tests, department level assessments, workplace assessments, etc., post-exam feedback from students, observers, standardized patients, administrators, etc., focus groups with students, observers, standardised patients, administrators, etc.

The variability of the observers comes from different reasons. It could be due to differences in the aspects of the performance they are rating or random error or how they use the rating form or how easy or hard they mark or whether they allow irrelevant aspects of individuals or their performance to influence their judgment. If the examiner proves to be more stringent marker than others, then we would recommend retraining the examiner. It is essential to train them on a regular basis, monitor examiners, provide them with feedback, and relieve them if there is no other alternative. It is very imperative to ensure the confidential report card for the examiners which contain information like individual performance, how many students they have examined, and the grades they have given compared to the performance of all other examiners as a group. It is also critical for the examiners to seek information regarding their performance from students, standardized patients, and administrators. If the task itself is not up to the mark, then it would be recommended to change the station.
3.2 Quality assurance of the standardized patient:
As with examiners, there can be variability among standardized patients even when they portray the same role. SPs also require coaching (Wallace, 2006). Unlike observers, SPs are usually non-professionals, so they must spend longer in training, and they are generally more compliant than experts. Also, it is not unusual for them to be required to reach training benchmarks before they can become part of a live exam. Nonetheless, even after administration, it remains vital for you to provide them with feedback, retrain them periodically, and relieve them of duty in instances where there is no alternative.

The first is to monitor their performance routinely by watching them or videotaping them regularly, to be sure that they are playing their role with consistency. This observation is often accompanied by "double scoring" where the person who observes the SP, completes the same rating form or checklist as well. These are compared to ensure that they are similar, and retraining is undertaken as needed.

The second strategy is to develop a report card of sorts for the SP, containing average grades over students and over time. Nonetheless, it is essential for you to investigate any concerns expressed by students, staff, or expert observers.

3.3 Quality assurance for the station as a whole
There are a few relatively straightforward statistics that can be calculated after administration and which address in a general way the quality of an OSCE station. We can review the percentage of students taking action if it is a checklist, review the average rating if it is a global rating, and consider the correlation between each item and total station and total test score or mean.

However, if a station is not performing well, it will be unclear whether it is a result of the content of the station itself, the rating forms that have been used, the observer, or, if relevant, the standardized patient. It could also be a problem with the original blueprinting and design. Correcting stations that are not performing well is a painstaking task. It might, in some cases, be better to design a new station.

Conclusion

Conducting the OSCE and standardized patient assessments are challenging for medical schools with limited resources, but achievable with proper planning. Those resources include standardized patients, physicians, and faculty members and adequately equipped physical facilities. Both the patients and physicians require training. In spite of the challenges, we concluded that proper use of OSCE exams done with a logistical approach is feasible and beneficial in the medical teaching environment.

Take Home Messages

- Even though holistic assessments are gaining importance in medical education and training medical students and residents, OSCEs can produce valid and reliable if they are conducted properly.
- The reliability of OSCEs can be achieved by letting the students go through the same and the multiple numbers of stations.
- The validity of the OSCEs is determined by the content of the stations and the blueprinting of the exam.
- Training of standardized patients and examiners is required before the objective structured clinical examinations.
- Continuous quality assurance of the examiners, standardized patients, and stations is required for achieving
the intended outcome of OSCEs.

Notes On Contributors

Dr. Sireesha Bala Arja, Associate Professor of Pharmacology and Associate Dean of Basic Sciences

Dr. Sateesh B. Arja, Associate Professor of Clinical Skills and Medical Education, Executive Dean

Mr. Jesse Ramey, 2nd-year medical student

Dr. Tarig Fadallah Altahir Ahmed, Assistant Professor of Medical Ethics and Clinical Skills, Chair of Promotions and Assessment Committee

Dr. Kumar Ponnusamy, Professor of Molecular Basis of Medicine

Dr. Venkata Raghavendrarao Morusupalli, Chair of Diseases, Immunity, and Therapeutics Department, Associate Dean of Student Affairs

Acknowledgements

None.

Bibliography/References

Batra, B. and Norcini, J. (2016). Developing objective structured clinical examinations (OSCEs). FAIMER-Keele Master's in Health Professions Education: Accreditation and Assessment. Module 1, Unit 3. 4th edition. FAIMER Centre for Distance Learning, CenMEDIC, London.

Blaney, D. and Dacre, J. (2017). Assessment: Practical Observations. FAIMER-Keele Master's in Health Professions Education: Accreditation and Assessment. Module 3, Unit 6. 4th edition. FAIMER Centre for Distance Learning, CenMEDIC, London.

Cleland, J.A., Abe, K., Rethans, J. (2009). 'The use of simulated patients in medical education: AMEE Guide No 42'. Medical Teacher, 31, 6, 477-486. https://doi.org/10.1080/01421590903002821

De Champlain, A. and Revest, T. (2016). Standard setting. FAIMER-Keele Master's in Health Professions Education: Accreditation and Assessment. Module 1, Unit 8. 4th edition. FAIMER Centre for Distance Learning, CenMEDIC, London.

Harden, R.M. (1990). 'Twelve tips for organizing an objective structured clinical examination (OSCE)'. Medical Teacher, 12,314, 259-264. https://doi.org/10.3109/01421599009006629

Harden, R.M. and Gleeson, F.A.(1979). 'Assessment of clinical competence using an objective structured clinical examination (OSCE)'. Medical Education, 13, 1, 39-54. https://doi.org/10.1111/j.1365-2923.1979.tb00918.x
Hodges B, McIlroy J.H. (2003). 'Analytic global OSCE ratings are sensitive to level of training'. Medical Education, 37, 11, 1012 – 1016. https://doi.org/10.1046/j.1365-2923.2003.01674.x

Miller G.E. (1990). 'The assessment of clinical skills/competence/performance'. Academic Medicine, 65, 563–67. https://doi.org/10.1097/00001888-199009000-00045

Newble, D.I. (2002). Assessing clinical competence at the undergraduate level. In: Medical Education Booklet, No.25. Edinburgh: Association for the Study of Medical Education.

Norcini, J. and Boulet, J. (2003). 'Methodological issues in the use of standardized patients for assessment'. Teaching and Learning in Medicine, 15, 4, 293–297. https://doi.org/10.1207/S15328015TLM1504_12

Norcini, J. and Troncon, L. (2016). Foundations of assessment. FAIMER-Keele Master’s in Health Professions Education: Accreditation and Assessment. Module 1, Unit 1. 4th edition. FAIMER Centre for Distance Learning, CenMEDIC, London.

Tamblyn, R., Abrahamowicz, M., Dauphinee, D., Wenghofer, E., et al, (2007). 'Physician scores on a national clinical skills examination as predictors of complaints to medical regulatory authorities'. JAMA, 298, 9, 993-1001. https://doi.org/10.1001/jama.298.9.993

Wallace, P. (2006). Coaching Standardized Patients: For use in the Assessment of Clinical Competence. First edition. New York; Springer Publishing Company.

Appendices

None.

Declarations

The author has declared that there are no conflicts of interest.

This has been published under Creative Commons "CC BY 4.0" (https://creativecommons.org/licenses/by-sa/4.0/)

Ethics Statement

The Research and Ethics Committee exempted this study from ethics approval as there is no involvement of human subjects and this study is completely based on literature review.

External Funding

This paper has not had any External Funding

AMEE MedEdPublish: rapid, post-publication, peer-reviewed papers on healthcare professions’ education. For more information please visit www.mededpublish.org or contact mededpublish@dundee.ac.uk.