Dr. BDJ Silva Oration 2013

Assessing wisely the objective structured clinical examination (OSCE) – the way forward...

For assessing clinical skills at primary care examinations

Shyamalee Samaranayake[1]

Sri Lankan Family Physician, 2016, **32**, 17-20

Assessment drives learning. That is why Gibbs in 1999 said, "assessment is the most powerful lever teachers have, to influence the way students respond to courses and behave as learners"[1]. Assessment is an integral part of the teaching and learning process. It measures student achievement and what they have learnt. Assessment also provides a good feedback on teaching and training. It indicates whether the objectives of the programme have been achieved and whether the outcome of learning is as intended. Therefore, without an assessment the learning will not take place optimally.

Assessment can be formative or summative. Formative assessments are done during a training programme. They can give feed back for trainees as well as trainers. It is used to monitor student progress, to identify learning needs, which parts need strengthening and to reshape the curriculum. Summative assessments are carried out at an exit point and are being used to award grades and qualifications.

The quality of an assessment is determined by its validity and reliability[2]. Validity refers to how well the assessment measures what it is supposed to measure[2]. Are we assessing what we intended to measure? Are we fulfilling the objectives of the examination? Validity can be measured in several ways. Face validity refers to the compatibility of the assessment with the curriculum; on the face of it, does is it match with the curriculum. Content validity is the extent to which the assessment measures a representative content of the syllabus. Are we assessing an adequate sample from the curriculum? Construct validity is the measurement's ability to differentiate between the good candidates who show their competence and perform well from the poorly performing candidates.

Reliability of an assessment is the degree to which the assessment gives consistent and reproducible results[2]. If we give the same assessment to 2 similar groups of candidates, are they producing similar results?

According to the skills assessment framework described by Miller, there are 4 different levels that can be used to assess candidates[3]. They are knows, knows how, shows how and does. Higher the level, better the assessment becomes.

The highest level of competence "Does" can be assessed using work place based assessment (WPBA) tools. WPBA is very good as a formative assessment, mainly to give feed back and to improve the level of competence. However, this is not very practical as a summative assessment.

The objective structured clinical examination, the OSCE belongs to the 2nd best level in the Miller's pyramid the "shows how". It is a performance based assessment and is a very effective way to assess clinical competence at summative examinations.

The OSCE is objective because all the candidates are presented with the same test with same marking grid or similar cases if it takes place over several days. It is structured because the case and the marking scheme for each station is structured. It is a clinical examination as it can be used to assess skills, behaviours, attitudes and application of knowledge.

Harden and colleagues from the University of Dundee first described the OSCE in 1975. According to Harden "The OSCE is an approach to assess the clinical competence in which the components of competence are assessed in a planned or structured way with attention being paid to the objectivity of the examination"[4].

First OSCE was reported in Dundee and Glasgow in 1979. Introduction of simulated patients enhanced the evolution of OSCE. OSCE was adopted initially in North America. United Kingdom adopted OSCE as an assessment method in 90's and now it is the principle method of clinical skills assessment in medical schools and licensure bodies across America, New Zealand, Australia and Canada.

If we compare the traditional clinical examination methods, short and long cases with OSCE, both have their advantages and disadvantages.

Long and short cases are not labour intensive, no special training is required for patients and examiners and the patients are selected from the wards. It is possible to use a bigger variety of patients, especially children.

[1] *Senior Lecturer, Faculty of Medical Sciences, University of Sri Jayawardenapura.*

However, there are several limitations as well. Getting a sample of cases that are representative of the curriculum is difficult. Different candidates are tested on cases that are different and therefore the objectivity is low. As the history is not structured, patient's story can vary from candidate to candidate depending on their approach. Marking mainly depends on the individual examiners perception as there is no structured marking grid. Because of these reasons, validity and reliability is low in the traditional assessments.

On the other hand, OSCE represent a larger sample from the syllabus, all the candidates face the same level of difficulty. As the marking is structured, examiner bias is minimal. Therefore OSCE is a more valid and is a reliable tool for assessment for performance[5].

OSCEs also have some drawbacks. It does not assess a candidate's performance in a real life situation with a real patient. With the allocated limited time, it cannot assess how a candidate will face a complex situation. It is also not possible to use children due to practical reasons. OSCE is very much labour intensive and to have a proper OSCE, it is important to have a team of dedicated people. However when we compare the long and short cases with Objective Structured Clinical Examination, OSCE gives a better quality overall assessment.

Preparing for an Objective Structured Clinical Examination begins with formation of an expert group to identify content areas in the curriculum and the skills that need to be tested. The next step is to make a blueprint of these and to identify possible cases to be presented as OSCEs. Proper blueprinting is essential to get a representative sample of the curriculum and to assess variety of skills.

Then comes the development of OSCE stations. To facilitate this process, examiners need to be trained especially on identifying the skills to be tested and developing the marking grid.

A case comprises of 3 main components, instructions for the candidate, details of the case and the marking grid.

Instructions for the candidate should be clear and precise. What to do and what not to do should be specified clearly. The amount of information that should be mentioned in the instruction sheet depends on the complexity of the problem and the type of the examination. For an undergraduate OSCE, more information should be given in a simplified manner. Some postgraduate examinations only provide the name and the age of the patient. However most examinations tend to provide other relevant information like blood pressure, body mass index in the candidate information sheet.

Details of the case should include a synopsis of the case and key issues and skills to be tested should be identified. The candidate should be able to address all the issues during the allocated time. Skills, which could be tested in that particular station should be stated briefly, and these should match the expected outcomes of the teaching programme.

Information to the simulators should be given in a structured manner for them to understand easily. This section should have adequate information for the simulated patients (SP) to take up their roles. Appropriate and short opening statement should be mentioned, so the SP would be able to memorise and repeat the same word for word with all the candidates to give them the similar opportunity at the start. Comprehensive account about the problem with all the positive details and relevant negative details according to the differential diagnosis and patients concerns should be given to make a clear picture about the problem.

Domains of the marking grid should be decided according to listed skills to be tested. For undergraduate exams, a checklist to see whether a task is performed or not can be used. However at postgraduate level, examiners should assess each domain using the marking grid only as a guide.

Ideally all the cases should be piloted before putting in to a case bank. People may be reluctant to do this due to confidentiality issues. Piloting has to be done by a group of examiners who are familiar with the level of the exam and its objectives. During the piloting, cases are refined to include all necessary information to make it a real life situation, a believable scenario. This exercise is also important to check the timing of the station, so that the examiners can adjust the number of issues to match the allocated time period. Timing of the OSCE is very important; it should not be too long or too short. The marking grid is reviewed and revised to match the skills to be tested. Piloting is the key to have a proper OSCE and attempts should be made to do this prior to all the OSCEs. Confidentiality statements can be obtained from the examiners and simulators to make them aware about their obligations.

There are 3 parties involved in the OSCE; examiners, simulated patients or real patients and candidates. All 3 parties should be trained properly to get a better assessment.

Examiners have to be trained on OSCE writing, most importantly on developing the marking grid. They should also be trained on the marking. All the examiners should have some consensus about the level expected from the candidates; hence they should be briefed about the level descriptors, depending on the examination. This will create consciousness among examiners as to what level of competencies should be there to award a satisfactory passing grade. The level of competence expected from

postgraduates and undergraduates are different. The same examiner could appear at both undergraduate and postgraduate exams. Therefore they should be very clear in their minds about the standard expected.

Getting all the examiners together to practice marking using videos will help them to think in a similar manner in assessing. This will improve the reliability of the assessment.

Simulated patients are proven to be better than real patients in an OSCE setting. The information given by the simulators are uniform to all the candidates and this will improve the objectivity of the examination. Simulated patients should be trained carefully on taking up the patient's role especially being consistent, on expressing emotions and revealing clinical signs. It is very important to have a pool of simulated patients and training them on a regular basis to improve their role playing skills as everything in the examination depends on how they act their role. However, real patients can be used to test examination skills, particularly for eliciting physical signs.

Candidates too need prior training on how to face an OSCE. Ideally this should be done at a mock OSCE with an individual feedback to familiarise them with the skills necessary, level of competence expected and how to manage their time.

Then comes the actual conducting of the OSCE. A valid clinical test should assess the components in clinical competence; history taking and examination, management, patient education and even advance communication skills like counselling and breaking bad news. Using the blueprint, examiners need to identify a representative sample from the curriculum for the OSCE. This will increase the face validity and content validity. Selecting cases with different difficulty indices will improve the construct validity. There is enough evidence to prove that increasing the number of stations and the duration of the examination will improve the reliability. Having a wider sample of cases with different skills to be tested and an adequate number of stations with a reasonable duration will increase the reliability of the OSCE. Therefore to ensure the validity, reliability and objectivity, great care should be taken in planning the OSCE.

Simulated patients should be well trained to get in to their respective roles and must have a number of rehearsals before facing the candidates to maintain the consistency. Western universities train their simulators days to weeks before the exam on the real cases; they are briefed about the roles and even the script is given to take home. In our setting, this is not happening due to confidentiality issues, however we should allocate enough time to prepare the simulated patients and should have few run-throughs before the SP faces the real candidate.

It is also important that the candidate should be briefed about the OSCE circuit and timing. This briefing should include details about the number of stations, duration per station and the frequency of notification during a station.

Examiners have to revisit the grading process to improve inter examiner correlation. They also should be reminded again about the level descriptors and the conduct during the exam.

Standard setting is another important aspect that has to be considered in any examination. Certain examinations may be easier; simulators may not perform as expected, some examiners may mark more stringently than others. Therefore; ideally the pass mark should be determined on the performance of the candidates. Standardising the pass mark will make the assessment perfect.

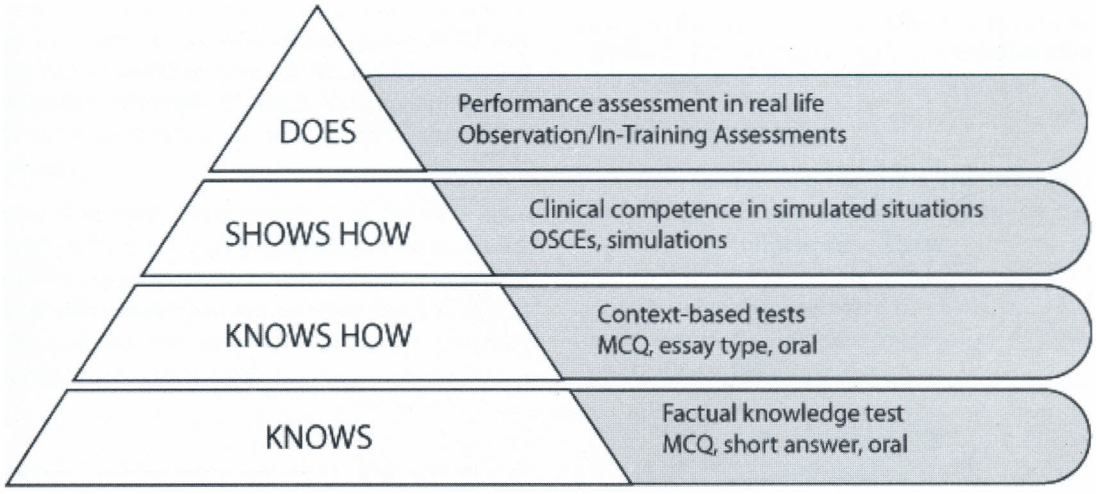

Figure 1. *The assessment of clinical skills/competence.*

In summary, OSCE is the best clinical assessment we can have for primary care as the summative assessment. This can be organised in a way to represent a primary care practitioner's day-to-day clinic with a good variety of problems. Further, OSCE is now considered all over the world as the best practical tool available for assessment of clinical competence. It is the only assessment where the full consultation could be observed and is the best way to assess communication skills, which is the most important skill for a primary care doctor. It is a performance-based assessment and a variety of cases can be given at the same examination to represent the wider range of problems seen in primary care. It is more reliable and valid than other tools of assessment. Therefore a properly constructed OSCE is the way forward in assessing clinical competence for all primary care examinations.

References

1. Gibbs G. Using assessment strategically to change the way students learn, in Assessment Matters in Higher Education, eds Brown S and Glasner A, Society for Research into Higher Education and Open University Press, Maidenhead, pp 41-53tre, 1999.
2. McAlpine M. Principles of assessment, Robert Clerk centre for Technological Education, University of Glasgow, 2002.
3. Miller GE. The assessment of clinical skills/competence/ performance. *Acad Med* 1990; **65**(9): 63-7.
4. Harden RM, Gleeson FA. Assessment of clinical competence using an objective structured clinical examination. ASME medical education booklet no **8**: 39-54.
5. A systematic review of the reliability of objective structured clinical examination scores. Michael T Brannick et al http://www.ncbi.nlm.nih.gov/pubmed/21988659#