Reviewer’s report

Title: How Well Do Second-Year Students Learn Physical Diagnosis? Observational Study of an Objective Structured Clinical Examination (OSCE)

Authors:

Claus Hamann (chamann@partners.org)
Kevin Volkan (kevin.volkan@csuci.edu)
Mary B Fishman (MBF3@gunet.georgetown.edu)
Ronald C Silvestri (rsilvest@caregroup.harvard.edu)
Steven R Simon (steven_simon@hms.harvard.edu)
Suzanne W Fletcher (suzanne_fletcher@hms.harvard.edu)

Version: 1 Date: 25 Oct 2001

Reviewer: Dr Tony Townsend

Level of interest: A paper whose findings are important to those with closely related research interests

Advice on publication: Accept after revision, which I do not need to see

This is a useful study and adds to our understanding of the use of an OSCE as an educational and assessment instrument. A major strength of this study is that it involved observation and analysis of data from 489 students over 3 years. The results are adequate to support the conclusions and provide additional support for the increasing evidence that a properly constructed OSCE is a reliable and valid instrument.

The title and abstract fairly describe the study but the running title is misleading in that the study is concerned with assessment of the learning of basic clinical skills rather than the actual learning of them. The suggested running title for the paper is 'learning physical diagnosis' which is even less accurate.

The method used is clearly described, as is the development of the OSCE instrument. However, some details of the questions used in the stations would be helpful, e.g. how can 30 questions relevant to differential diagnosis be asked about calf pain? Utilization of the OSCE as an educational tool was enhanced by the provision of immediate feedback at the end of each station but that cut the students' time for performance of the required tasks to 4 minutes for all but two stations. This is a very short time, particularly for some stations such as history-taking and differential diagnosis of hemoptysis. No comment has been made about whether or not students were able to complete the tasks in the allocated time. Data were adequately analyzed and the inclusion of multivariate analyses of factors associated with OSCE performance strengthens the observations.

It was of interest that despite significant differences in teaching sites for both station and clinical skills scores, the total OSCE score was not site dependent. This is likely to be a reflection of the number of stations used and the wide range of skills assessed. It was also of interest that the lowest scoring station, calf pain, carried the second to highest mark (highest for a 6 minute station) and that 30 marks were awarded for differential diagnosis. These results should lead to further assessment of the teaching sites and ongoing development of the OSCE stations.
The study confirms that the OSCE is a faculty intensive and organizationally complex assessment/educational tool with 445 faculty being involved over 3 days. Although the reliability of one of the skill scores and 7 of the station scores were less than 0.6, overall the OSCE has good reliability and compares well with other tools used for measuring clinical competence across a broad range of skills.

**Competing interests:**

None declared.