

Title: Improving the reliability and accuracy of the Objective Structured Clinical Examination (OSCE) at final MBBS level

Author: Christina Tan

Context and setting: The Objective Structured Clinical Examination (OSCE) is used for summative assessment at final year level in the undergraduate medical programme, and was first run in March 2003. It enables minor disciplines to be involved, and covers assessment of communication skills, data interpretation leading to patient management issues, and psychomotor skills.

Why the idea was necessary: There were no detailed records of clinical core competencies expected by clinical disciplines at the end of the Bachelor of Medicine & Bachelor of Surgery (MBBS) programme, and no analysis of whether these core competencies had been assessed in previous OSCEs prior to 2006. Previous observations of examiners in OSCEs prior to 2006 noted that they were inconsistent in marking checklists and some demonstrated inappropriate behaviour such as prompting and teaching. Examiner training was therefore required to improve the reliability and accuracy of this assessment process.

What was done: Clinical disciplines were asked to provide 5 core competencies expected of graduating doctors after completing the 5-year MBBS programme. Individual stations in past OSCEs could then be reviewed to determine if these competencies were assessed. Training workshops were organised for faculty who would be involved as OSCE examiners. During the March 2006 Final MBBS OSCE, examiners were observed to determine if trained examiners' assessment rating skills had improved and that there was inter-rater reliability.

Evaluation of results and impact: Survey forms were not returned by all clinical disciplines (the return rate was 75%) and retrospective analysis of core clinical competencies in past OSCEs was therefore not done.

During the training workshops, the "examiners" showed wide variation in scoring and some awarded marks without any justification or for reasons which were not listed on the station checklists. After comparison and discussion, scores showed more consistency with less variation. Feedback on examiner behaviour was provided by workshop facilitators.

Observers in an OSCE reported wide variations in practices between different examiners (whether trained or untrained) examining the same OSCE station in parallel tracks and inappropriate behaviour. However, the examiners who underwent the training workshop prior to the examination were on the whole consistent in their behaviour, with no or minimal prompting.

These preliminary results suggest that the examiner training workshops are helping examiners be more aware of and more consistent in their behaviour during OSCEs. A pre-OSCE briefing for all examiners would help provide clear criteria of marking and improve the consistency of rating, as well as serve as a reminder on examiner conduct. Some prompting was done by examiners because the OSCE question was not clear, giving rise to issues of question design and vetting that need to be addressed in a different training workshop. Training and selection of standardised patients also needs to be more rigorous.

Examiner training and feedback through observation in OSCEs needs to be continued.

Correspondence:

Christina Tan, MBBS, MRCP (UK), FRCGP (UK)

Department of Primary Care Medicine, University of Malaya, Malaysia