

## Editorial

# Is OSCE valid for evaluation of the six ACGME general competencies?

Assessment of clinical competencies by using objective structured clinical examination (OSCE) is one of the current mainstays in medical education. According to the concept of Miller's pyramid, OSCE is defined as an assessment method belonging to the entity of "shows how", i.e. an assessment method designed for the evaluation of clinical skills.<sup>1</sup> In this issue of the Journal of the Chinese Medical Association, Yang et al.<sup>2</sup> have developed an OSCE program to assess the six general competencies, including medical knowledge (MK), interpersonal and communication skills (ICS), systems-based practice (SBP), patient care (PC), professionalism, and practice-based learning and improvement (PBLI), advocated by the Accreditation Council for Graduate Medical Education (ACGME). The authors assessed the reliability of the examination and concluded that OSCE could be a reliable method for determining postgraduate year 1 (PGY<sub>1</sub>) resident's acquisition of the ACGME core competencies. Besides reliability, a more important and interesting question to be answered is whether or not OSCE is valid for evaluation of the six ACGME general competencies?

OSCE was primarily designed for assessment of clinical skills such as history taking, physical examination, simple procedures, interpretation of laboratory results, patient management problems, communication, attitude etc.<sup>3,4</sup> According to the ACGME Toolbox of Assessment Methods,<sup>5,6</sup> OSCE is one of the most desirable methods for evaluation of the relevant skills for PC, ICS, and professionalism; it can also be one of the next best methods for evaluation of PBLI and SBP, but is not considered as a cost-effective method to evaluate MK.

The contents of the ACGME Toolbox of Assessment Methods are apparently reasonable. However, the validity evidence of OSCE on assessment of the ACGME general competencies is still limited and necessary for further verification. In a recent study, Berman et al.<sup>7</sup> investigated the validity of the New York City Rheumatology OSCE (ROSCE) on assessment of four of the six ACGME general competencies, i.e., PC, ICS, professionalism, and MK. The authors disclosed promising results; however, the merit of their work is not for providing new concept but solid evidence on the validity of OSCE. As stated by the authors, during the 5-year period of investigation, every attempt on design and evolution of ROSCE is made to create scenarios that mimic typical clinical encounters with realistic patient interactions to improve examination validity. In other words, to build up a valid OSCE

program for assessment of the ACGME competencies may be both time and effort consuming.

In contrast to the four competencies assessed by ROSCE, SBP and PBLI, the quality improvement (QI)-related competencies involved in multidisciplinary cooperation and/or complex behaviors are more difficult to be evaluated. Currently, there is still little validity evidence to support the usefulness of OSCE in assessing SBP and PBLI. The first report using a single OSCE station to assess the performance of medical students in root cause analysis (RCA) and communication of a prescription error was published in 2007.<sup>8</sup> Afterwards, Varkey et al.<sup>9</sup> further demonstrated validity evidence of OSCE on assessment of SBP and PBLI. Their study contains several implications worth pondering. First of all, OSCE should not be an independent assessment project; it must be an integrated part of one or more clinical curricula. In the study of Varkey et al. the OSCE they used was designed in compliance with a three-week QI curriculum. The blueprint for the OSCE was based on the main curriculum objectives: prescription error, negotiation, evidence-based medicine (EBM), team collaboration, RCA, quality measurement, Nolan's three-question model of QI and insurance systems. On the other hand, all the OSCE stations should be designed by teamwork. For improving the validity of their OSCE, Varkey et al. established an OSCE committee which consisted of three educational assessment experts and three other local institutional content experts. Finally, an ingenious or even well-established strategy for evaluation of validity of the OSCE is essential. It is because validity requires multiple sources of evidence to support or refute meaningful score interpretation.<sup>10,11</sup> For this reason, the validity analyses of Varkey's study contained five aspects of construct validity, including content evidence, response process, internal structure, relationship with other variables, and consequences, as described by Downing et al.<sup>11,12</sup>

After the validity of an OSCE has been established, the next step would be to use it to verify the efficacies of relevant training program(s). In the study of Yang et al.<sup>2</sup> the second goal was to compare core-competency acquisition before and after the internal medicine training program of PGY<sub>1</sub> residents. As the same OSCE stations were used to evaluate the efficacy of a teaching program, one potential bias—"practice effect" bias—should be taken into account. Consequently, a strategy to minimize this kind of bias should be conducted during the evaluation process. For example, "practice effect" bias could be

reduced by not providing PGY<sub>1</sub> residents with feedback after their initial OSCE, and not informing them that the second OSCE would have the same stations. In addition, this kind of bias can also be decreased by the time between testing.<sup>13</sup> Nevertheless, these measures are not proper for formative OSCEs.

Although Yang et al. have done a great job for the research and development of OSCE in Taiwan, as we take a glance at the education and assessment programs of western countries, obviously, we still have a long way to go.

Cho-Yu Chan

Department of Education,

China Medical University Hospital and School of Medicine,

China Medical University,

2, Yuh-Der Road, Taichung 404, Taiwan, ROC

E-mail address: d6537@mail.cmuh.org.tw

## References

1. Miller GE. The assessment of clinical skills/competence/performance. *Acad Med* 1990;**65**:S63–7.
2. Yang YY, Lee FY, Hsu HC, Huang CC, Chen JQ, Lee WS, et al. A core competence-based objective structured clinical examination (OSCE) in evaluating of clinical performance of postgraduate year-1 (PGY1) residents. *J Chin Med Assoc* 2011;**74**:198–204.
3. Harden RM, Stevenson M, Wilson DW, Wilson GM. Assessment of clinical competencies using objective structured clinical examination. *Br J Med Educ* 1975;**1**:447–51.
4. Ananthakrishnan N. Objective structured clinical/practical examination (OSCE/OSPE). *J Postgrad Med* 1993;**39**:82–4.
5. Accreditation Council for Graduate Medical Education. Outcome project. toolbox of assessment methods. <http://www.acgme.org/Outcome/assess/Toolbox.pdf>. [accessed 11. 03. 11].
6. ACGME competencies: suggested best methods for evaluation. <http://www.acgme.org/Outcome/assess/ToolTable.pdf>. <http://www.acgme.org/Outcome/assess/Toolbox.pdf> [accessed 11. 03. 11].
7. Berman JR, Lazaro D, Fields T, Bass AR, Weinstein E, Putterman C, et al. The New York city rheumatology objective structured clinical examination: five-year data demonstrates its validity, usefulness as a unique rating tool, objectivity, and sensitivity to change. *Arthritis Care Res* 2009;**61**: 1686–93.
8. Varkey P, Natt N. Novel use of the OSCE as a teaching and assessment tool of medical students' quality improvement skills. *Jt Comm J Qual Improv* 2007;**33**:48–53.
9. Varkey P, Natt N, Lesnick T, Downing S, Yudkowsky R. Validity evidence for an OSCE to assess competency in systems-based practice and practice-based learning and improvement: a preliminary investigation. *Acad Med* 2008;**83**:775–80.
10. American Educational Research Association/American Psychological Association/National Council on Measurement in Education. *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association; 1999.
11. Downing SM. Validity: on meaningful interpretation of assessment data. *Med Educ* 2003;**37**:830–7.
12. Downing SM, Haladyna TM. Validity threats: overcoming interference with proposed interpretations of assessment data. *Med Educ* 2004;**38**: 327–33.
13. Short MW, Jorgensen JE, Edwards JA, Blankenship RB, Roth BJ. Assessing intern core competencies with an objective structured clinical examination. *J Grad Med Educ* 2009;**1**:30–6.