

# Development and Implementation of an Objective Structured Clinical Examination to Provide Formative Feedback on Communication and Interpersonal Skills in Geriatric Training

Patricia O'Sullivan, MD,\*<sup>†</sup> Serena Chao, MD, MSc,<sup>‡</sup> Matthew Russell, MD,<sup>‡</sup> Sharon Levine, MD,<sup>‡</sup> and Anne Fabiny, MD<sup>§</sup>

Teaching and assessment of communication and interpersonal skills, one of the American Council for Graduate Medical Education–designated core competencies, is an important but difficult task in the training of physicians. Assessment of trainees offers an opportunity to provide explicit feedback on their skills and encourages learning. This article describes a pilot study in which clinician-educators affiliated with the geriatrics training programs at Beth Israel Deaconess Medical Center and Boston University Medical Center designed and piloted a novel Objective Structured Clinical Examination (OSCE) to assess the communication and interpersonal skills of medical, dental, and geriatric psychiatry fellows. The OSCE consisted of three stations where geriatricians and standardized patients evaluated candidates using specifically designed checklists and an abbreviated version of the Master Interview Rating Scale. Communication skills were assessed through performance of specific “real life” clinical tasks, such as obtaining a medical history, explaining a diagnosis and prognosis, giving therapeutic instructions, and counseling. Interpersonal skills were assessed through the effect of the communication between doctor and standardized patient on fostering trust, relieving anxiety, and establishing a therapeutic relationship. This pilot study demonstrated that the OSCE format of assessment provides a valid means of evaluating the communication and interpersonal skills of interdisciplinary geriatric trainees and provides a valuable forum for formative assessment and feedback. Given that geriatricians and non geriatricians involved in elder care both need communication and interpersonal skills, this novel OSCE can be used for assessment of these skills in trainees in diverse healthcare subspecialties. *J Am Geriatr Soc* 56:1730–1735, 2008.

From the \*Medical Department, Hebrew Senior Life, Boston, Massachusetts;<sup>†</sup>Division of Geriatrics, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts; <sup>‡</sup>Geriatrics Section, Department of Medicine, School of Medicine and Boston Medical Center, Boston University, Boston, Massachusetts; and <sup>§</sup>Division of Geriatrics, Cambridge Health Alliance, Harvard Medical School, Boston, Massachusetts.

Address correspondence to Patricia O'Sullivan, MD, Department of Gerontology, LMOB 1A, Beth Israel Deaconess Medical Center, 11 Francis Street, Boston, MA 02215. E-mail: posulliv@bidmc.harvard.edu

DOI: 10.1111/j.1532-5415.2008.01860.x

**Key words:** formative assessment; summative assessment; Objective Structured Clinical Examination (OSCE); competencies; communication skills; interpersonal skills

Medical practitioners and patients are keenly aware that medical knowledge is not the only element of competence that needs to be taught to trainees to facilitate establishment of therapeutic relationships with patients. The recognition of the importance of communication and interpersonal skills is reflected in the inclusion of these skills as components of “competence.” Competence may be defined as “the habitual and judicious use of communication, knowledge, technical skills, clinical reasoning, emotions, values, and reflection in daily practice for the benefit of the individual and community being served.”<sup>1</sup> The concept of “competence” has gained recognition since the American Council for Graduate Medical Education (ACGME) introduced six “core competencies,” or domains of competency, as a guide to teaching and evaluating students in graduate medical education.<sup>2,3</sup> These six domains of educational outcome are patient care, medical knowledge, practice-based learning and improvement, interpersonal and communication skills, professionalism, and systems-based practice. The six core domains offer a framework with which program directors can assess competence in trainees. Although all clinicians require interpersonal and communication skills, clinical geriatricians require these skills at a high level, given the complexity of the cases and the potential barriers to effective communication between doctors and their patients and the patients’ families. Although important, these skills may be difficult to cultivate if there is limited time during which the trainee directly observe the trainer or the trainer directly observes the trainee. The historic apprenticeship model of medical training, having the luxury of a higher trainer:trainee ratio, may have allowed for greater refinement of communication skills than the current training system allows for.

Personality and culture may play an important and confounding role in communication skills. Because of the intrinsically stylistic and subjective nature of communication and interpersonal skills, objective assessment of these domains of competence in trainees in a reproducible fashion is problematic.<sup>4</sup> It is plausible that “real” or standardized patients (SPs) are in a better position to facilitate and assess these skills in trainees. Because poor interpersonal or communication skills are a major cause of poor performance in physicians, this study aimed to develop a reliable and valid assessment tool to evaluate geriatric trainees in these domains. This pilot study was designed to develop and use a series of novel Objective Structured Clinical Examination (OSCE) stations with a “geriatric” theme to evaluate the interpersonal and communication skills of fellows in geriatrics. The scenarios were problem-based and strived to simulate cases that arise in daily clinical practice, where good communication and interpersonal skills are essential to delivering good medical care. Given the universal importance of these critical skills, the stations were applicable to trainees of different subspecialties and not just within the field of geriatric medicine. This is important, because the majority of older adults currently receive their medical care from nongeriatricians. Trainers need to try and solve the challenge of training and assessing nongeriatricians in the domain of geriatric medicine and in particular in the realm of certain competencies. The OSCE may provide an innovative and unique forum for assessing interdisciplinary trainees.

## METHODS

### Study Aims, Design, and Participants

The purpose of this pilot study was to objectively evaluate the communication and interpersonal skills of trainees undergoing fellowship training in clinical geriatrics. The study was designed to evaluate the candidates in a single site on a single day using identical test conditions and evaluators. The interpersonal and communication skills of 17 multidisciplinary fellows in geriatrics were evaluated. There were 11 medical, three dental, and three psychiatry fellows from five fellowship training programs in the Boston area.

### Development of the OSCE

The three stations were designed to explore scenarios of particular relevance to geriatrics, namely, elder abuse, medication-error disclosure, and caregiver stress. The stations were also designed to be equally relevant and plausible to fellows in different subspecialties by exploring scenarios with universal themes in which interpersonal and communication skills employed by the physician were paramount. The specific details of the OSCE stations can be requested by contacting the first author (POS).

### Use of SPs

Six SPs were recruited from a local panel of SPs. They undertook 4 hours of training in advance of the OSCE. This consisted of an explanation of the goals, structure, and content of each OSCE station. The SPs were taught how to use and score an abbreviated version of the Masters Interview Rating Scale (MIRS), a rating scale designed to evaluate interpersonal skills. SPs evaluated trainees on six MIRS items: organization of the interview, types of questions asked, development of rapport, pacing of the interview, whether trainees allowed “patients” to ask questions, and the quality of the closing of the interview (Figure 1).

### Use of Preceptors

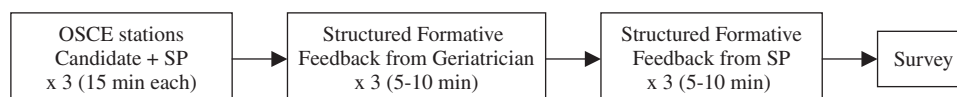
Preceptors were recruited from local faculty of geriatric medicine from the Boston University Medical Campus and Harvard Medical School. Faculty recruitment was augmented using leadership support and encouragement. The preceptors were explicitly instructed through e-mail on how to score the candidates using a station-specific checklist. Checklists were designed (blueprinted) to reflect concepts specific to each individual case. Each checklist consisted of 14 to 20 content items that it was felt were necessary to be discussed during the course of the interview. The specific blueprinted checklists for the three stations can be requested by contacting the first author (POS). Table 1 illustrates a sample checklist for the elder abuse station.

### Structure of the OSCE

Each fellow participated in three OSCE stations. The OSCE was conducted over 2 hours in one session and in one location. Each station consisted of an initial 15-minute interaction between the candidate and an SP. A preceptor (an experienced geriatric care provider) observed (by closed-circuit television) and evaluated the encounter using the standardized 14- to 20-item checklist.

### Evaluation and Measurements

Before the OSCE assessment, it was made clear to the candidates that the skills being evaluated were interpersonal and communication skills rather than medical knowledge or clinical examination technique. After the encounter, the geriatric care provider entered the room and provided structured formative feedback, guided by the checklist parameters. Before the precepting geriatrician provided face-to-face oral feedback to the candidate, the SP rated the interpersonal skills of the candidate using the abbreviated MIRS. The candidate then received feedback from the SP, guided by parameters outlined in the MIRS. Finally, 30 minutes after completion of the OSCE, the candidates were surveyed using a questionnaire. The candidates were asked whether they felt that the OSCE was appropriately difficult, valid, and acceptable and had an educational effect and whether they felt that their performance in the OSCE



**Figure 1.** General outline of the Objective Structured Clinical Examination (OSCE). SP = standardized patient.

**Table 1. Sample Checklist of 16 Critical Tasks Used by Preceptor to Score Interpersonal Skills of Candidates During an Objective Structured Clinical Examination Station that Simulated the Scenario of Elder Abuse**

Critical Task Performed by Candidate	Yes/ No
1. Introduces himself or herself by name	
2. Explains his or her role or position	
3. Asks for or uses patient's name	
4. Asks about circumstances surrounding the bruised eye	
5. Asks whether there are any problems with family or household members	
6. Asks about past medical history and prior hospitalizations	
7. Asks about alcohol intake	
8. Asks about circumstances of previous fractures	
9. Asks whether anyone has hit or hurt the patient	
10. Asks whether patient is in immediate danger of harm from son	
11. States that he or she has a legal obligation to make sure that the patient remains safe and therefore must report the physical abuse to adult protective services	
12. Communicates a sense of urgency without frightening the patient	
13. Explains that a report to adult protective services will result in a caseworker being assigned to decide whether a case should be opened	
14. Explains that the case worker will interview the patient and abuser to determine the underlying psychological and social problems that are influencing the family dynamic and might be able to suggest community agencies to help the patient and his or her son	
15. Makes a plan for a follow-up appointment	
16. Makes an emergency plan	
<b>Total score</b>	<b>/16</b>

correlated well with how they perform in daily clinical practice.

**RESULTS**

**Preceptor Scores**

The OSCE yielded objective performance scores for all candidates. The preceptors compiled performance scores for each candidate in each station. Three separate preceptors each assigned a performance score to each candidate from one of the three stations. The maximum possible per-

formance scores were 16 from the elder abuse station, 14 from the medication error station, and 20 from the caregiver stress station, based on the blueprinted scoring system for each station. Hence, the maximum possible summated performance score was 50. The scores are shown in Figure 2 as a percentage of correctly performed "critical tasks" from a cumulative total of 50 such tasks.

**Standardized Patient Scores**

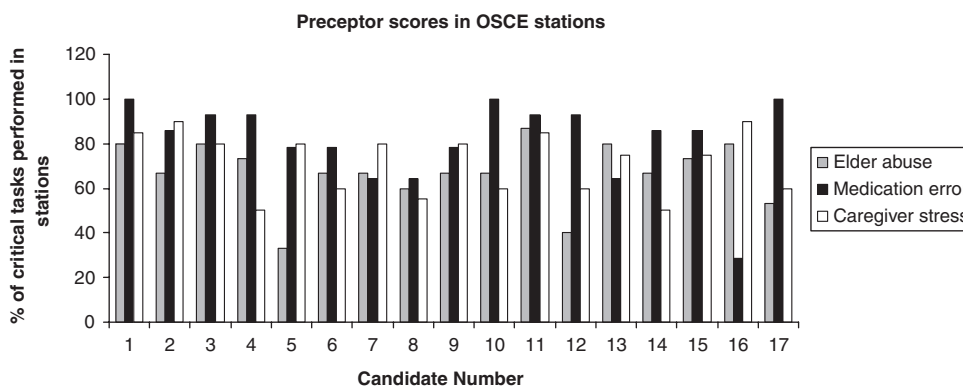
Three separate SPs each assigned a performance scores from one of the three stations to each candidate. The SP scores generated by the abbreviated MIRS from each of the three stations (30 points per station) were summed to give a total score (maximum of 90 points) (Figure 3).

**Correlation of Preceptor and SP Scores**

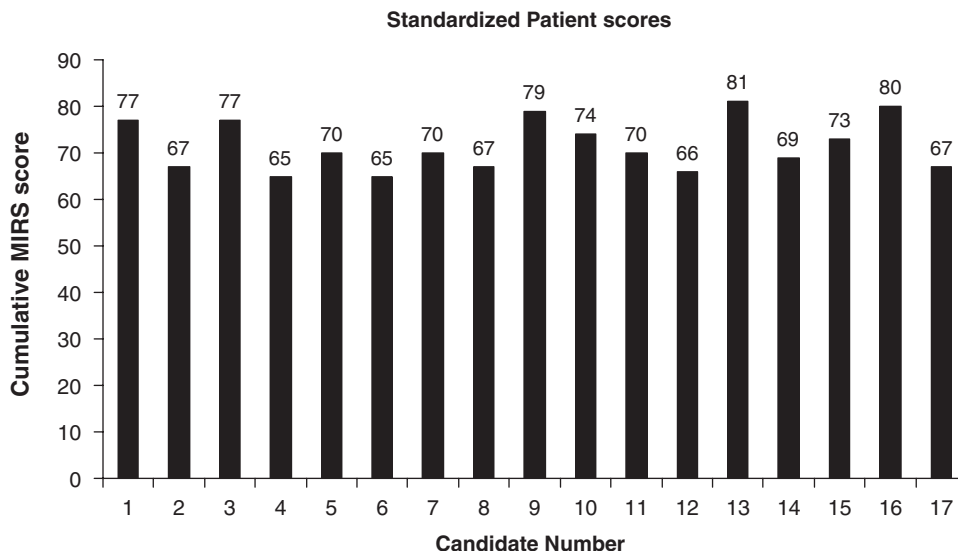
In general, the preceptor performance scores were taken to reflect the communication skills of the candidates, and the SP performance scores were taken to reflect the interpersonal skills of the candidates. Although these skills overlap in clinical practice, separate assessments from SPs and geriatricians using different assessment tools allowed the distinctive aspects of these competencies, as defined by the ACGME outcomes project, to be examined. Figure 4 displays the imperfect correlation between these scores.

**Survey of Candidates**

In general, the candidates felt that the OSCE identified deficits and had a positive educational effect. The candidates responded by choosing from a 5-point response scale (strongly disagree, disagree, unsure, agree, strongly agree). Thirteen candidates agreed or strongly agreed that their performance on the OSCE correlated well with their performance in clinical practice. All of the candidates felt that the level of difficulty was appropriate and that each scenario was plausible, although three candidates felt that the allotted time for each station was inadequate. Only one candidate felt that the OSCE was not a fair evaluation of the skills taught during fellowship, and all of the candidates felt that the OSCE helped to identify skills that they needed to improve on. Only one candidate felt that the feedback obtained at the end of the stations was not helpful and that the OSCE stations did not encourage them to learn more about the topics covered. Finally, despite these favorable



**Figure 2.** Summary of preceptor-generated performance scores of the candidates at the three stations. The scores were displayed as a percentage of maximal possible score in each scenario.



**Figure 3.** Summated standardized patient-generated performance scores of the candidates at the three stations. The scores are displayed as a numeric value out of a maximum possible cumulative score of 90. MIRS=Masters Interviewing Rating Scale.

responses, only 10 candidates agreed or strongly agreed that their fellowship director should use their performance on this OSCE when deciding whether they should graduate from fellowship training. Candidates reported that, although they recognized that the OSCE was worthwhile and an authentic test of their skills, it was challenging and stressful.

**Reliability**

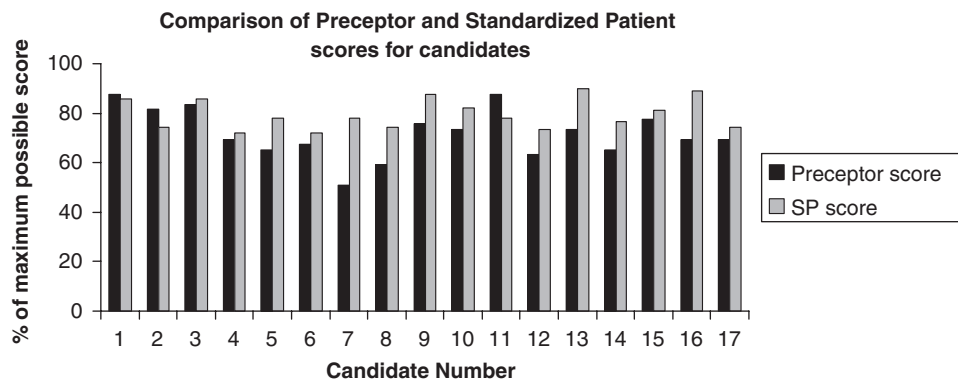
Although the candidates typically scored high on each of the stations, psychometric analysis of the preceptor-generated and SP-generated scores showed poor interstation reliability, a finding that was not surprising given the low number of stations piloted (n = 3). In addition, there was poor correlation between preceptor- and SP-generated scores within the three OSCE scenarios. Again this was not surprising given the limited number of stations used. Of particular interest because of the novelty of this multidisciplinary OSCE, there was no statistically significant difference between how trainees in medical, dental, and psychogeriatrics scored on communication or interpersonal skills.

**Validity**

Authors and preceptors of the OSCE agreed on its validity. Thirteen of 17 candidates felt that their performance on the OSCE correlated well with their perceived performance in daily clinical practice. The OSCE topics were blueprinted, or matched to fit curricular learning objectives. Commonly seen, potentially serious, and often undertaught curricular content was specifically chosen. The authors felt that the pilot OSCE was “content valid” in that it evaluated domains that were specifically intended to be assessed. Thirteen of the 17 candidates felt that the assessment resembled real-life clinical practice (was “face valid”). The pilot nature of this OSCE precluded judgment of other forms of test validity such as construct, predictive, and consequential validity. Given the lack of criterion standard assessment of communication and interpersonal skills, the criterion validity of the OSCE is unknown.

**Cost**

The cost of this pilot study of three stations was estimated at approximately \$110 per candidate. This includes SP



**Figure 4.** Summary of standardized patient- and preceptor-generated performance scores of the candidates at the three stations. SP = standardized patient.

training and performance, SP parking fees, and food. Preceptor and clinical skills center staff time was donated. The addition of scenarios and candidates would involve additional cost per unit time and per SP.

### Acceptability

The candidates received the OSCE well and perceived it to be an acceptable and valid test of interpersonal and communication skills.

Overall, the pilot study proved acceptable to candidates, had a positive educational effect, and was deemed valid by candidates and preceptors but had poor reliability scores, as was predictable, given the low number of stations. The interactive, theatrical nature of this memorable form of clinical assessment facilitates reflective learning.

### DISCUSSION

This was a novel OSCE specifically developed to assess the interpersonal and communications skills of interdisciplinary trainees in different subspecialties of geriatric training. In 1990, a framework or “pyramid” for assessing clinical competence was proposed.<sup>5</sup> At the lowest level of the pyramid is knowledge (knows), followed by competence (knows how), performance (shows how), and action (does).<sup>6</sup> The above-described pilot OSCE is an example of performance (shows how) assessment. The ultimate goal for a valid assessment of clinical competence is to test what the doctor does in the work place (an “in vivo” assessment). This simulates real clinical practice, where many extraneous factors influence clinical performance, including time constraints, motivation, physical and mental health of practitioner, and personalities of doctor and patient. Valid assessment is ideally a longitudinal process, involving multiple formats of assessment.

The literature on the gerontology-specific OSCE is sparse. In 2002, one group reported, in abstract form, their experience with an educational curriculum in inpatient geriatrics for internal medicine housestaff.<sup>7</sup> They designed their OSCE stations to probe the medical aspects of geriatrics rather than assessment of interdisciplinary and communication skills. The same group subsequently described the use of the OSCE to formatively assess a curriculum in inpatient geriatric medicine, among rotating internal medicine house staff, in an acute care unit for elderly people in a large urban medical center.<sup>8</sup> Again, the emphasis of the OSCE stations was the assessment of candidates' ability to manage commonly encountered “medical” scenarios in clinical gerontology. Their pilot data show that, although more than 90% of learners performed satisfactorily on three of five procedure stations and four of five question-and-answer stations, discrepancies were apparent between self-reported knowledge results when compared with OSCE results.

Another group described the use of an OSCE station to evaluate geriatric functional assessment in third-year medical students as part of a vertically integrated geriatric curriculum.<sup>9</sup> Subsequently, a study described the use of geriatric-focused Observed Structured Video Examinations, 1- to 3-minute video triggers associated with a series of multiple choice and constructed response questions (e.g., fill in the blank), to assess residents' understanding of pro-

fessionalism, systems-based practice, communication, and practice-based learning.<sup>10</sup>

This article reports a study in which a novel OSCE was developed and piloted to evaluate the interpersonal and communication skills of interdisciplinary trainees in gerontology. Given how important communication and interpersonal skills are in gerontology and how effectively assessment drives learning, the assessment of these skills is an important part of the “hidden” curriculum when training junior doctors and undergraduates. Although the OSCE has been used to assess communication and interpersonal skills in medical, dental, and psychiatric trainees separately, to the best of the authors' knowledge, there is no literature on the interdisciplinary OSCE. The universal applicability of these particular core competencies makes this novel OSCE suitable for assessment of trainees in different subspecialty programs. The ability to assess important elements of the practice of geriatric medicine in combinations of subspecialties is important, because the majority of those who care for geriatric patients are not geriatricians. The ability to pool resources from different training programs is valuable if resources, and numbers of trainees for assessment, are limited. It was possible to combine the resources from fellowship programs and augment the sample size by assessing 11 medical, three dental, and three psychiatry fellows from five fellowship training programs in the Boston area.

As expected in a pilot study, poor reliability was observed on psychometric analysis of the test scores. Generally speaking, a minimum of 10 stations over 3 to 4 hours is required to achieve a reliability of 0.85 to 0.90, 0.8 being the minimum standard acceptable.<sup>11</sup> Also, it is likely that wider sampling of cases and structured assessment would improve reliability.<sup>12</sup> Although reliability was expectedly substandard, the majority of candidates found the assessment to be a valid representation of everyday practice. The OSCE scenarios were blueprinted to test the ability of candidates to use communication and interpersonal skills when assessing three common but potentially serious clinical situations (elder abuse, medication error, and caregiver stress). In this pilot study, satisfaction surveys indicated that the interactive nature of the intervention motivated house staff to learn. They confirmed the applicability of such an assessment, regardless of the subspecialty of candidates.

Although the OSCE is a valid and reliable form of assessment, given an adequate number of stations, it is often logistically demanding and expensive. Although the cost of conducting an OSCE can seem daunting, it must be remembered that it can form one of the most visceral and memorable teaching experiences of the year. A pilot study allows for identification of potential operational problems and pitfalls.

Although the OSCE has been incorporated into the geriatric medical fellowships involved, questions for the future involve the challenge of expanding the OSCE and making it more reliable by adding more stations. Because this was a novel OSCE and formative assessment process, there was grading, although this is an important element to review on subsequent iterations. Although instructions on feedback can be given to preceptors, the issue of remediation is more difficult. Some might argue that basic communication skills are a prerequisite for “fitness to practice.” The question was raised whether students who fail to

demonstrate a minimal level of competence in this area should be allowed to progress to the next stage of the course and eventually graduate. Recently, one author described a program in which students who fail the communication skills OSCE stations, irrespective of their overall examination performance, undertake remedial studies in a compulsory 2-week directed study module followed by a four-station OSCE.<sup>13</sup>

In summary, the results supported the hypothesis that the OSCE can be conducted with interdisciplinary trainees (medical, dental, and psychiatric–medical) to provide valuable formative assessment on interpersonal and communication skills. The OSCE stations have universal applicability to a wide number of subspecialties. The candidates valued the immediate and specific feedback given, and it became one of the most powerful and memorable teaching sessions of their training.

#### ACKNOWLEDGMENTS

This project was financially supported by the Health Resources and Services Administration (HRSA). The advice and support of Charles Hatem, MD, Beth Lowen, MD, and Lori Newman, MA is appreciated.

**Conflict of Interest:** There was no outside funding or support for this study. The authors do not had any conflict of interests with a company whose products or services are directly related to the subject matter of the manuscript.

**Author Contributions:** Patricia O’Sullivan: study concept, acquisition of subjects and data, analysis and interpretation of data, and preparation of manuscript. Serena Chao and Matthew Russell: acquisition of subjects and data, and preparation of manuscript. Sharon Levine: study concept and design, and acquisition of subjects and data. Anne Fabiny: study concept and design, acquisition of sub-

jects and data, analysis and interpretation of data, and preparation of manuscript.

**Sponsor’s Role:** This study was embarked upon and funded by the affiliated fellowship programs to critically evaluate the competencies of trainees according to ACGME standards. No external sponsor played a role in the design, methods, subject recruitment, or data collection of the pilot study or in data analysis or manuscript preparation.

#### REFERENCES

1. Epstein RM, Hundert EM. Defining and assessing professional competence. *JAMA* 2002;287:226–235.
2. Batalden P, Leach D, Swing S et al. General competencies and accreditation in graduate medical education. *Health Aff (Millwood)* 2002;21:103–111.
3. ACGME Outcome Project- Enhancing Residency Education through Outcomes Assessment [on-line]. Available at <http://www.acgme.org/outcome/> Accessed December 30, 2007.
4. Makoul G, Curry RH. The value of assessing and addressing communication skills. *JAMA* 2007;298:1057–1059.
5. Miller GE. The assessment of clinical skills/competence/performance. *Acad Med* 1990;65:63–67.
6. Wass V, Van der Vleuten C, Shatzer J et al. Assessment of clinical competence. *Lancet* 2001;357:945–949.
7. Karani R, Callahan EH, Thomas DC. An unfolding case with a linked OSCE: A curriculum in inpatient geriatric medicine. *Acad Med* 2002;77:938.
8. Karani R, Leipzig RM, Callahan EH et al. An unfolding case with a linked Objective Structured Clinical Examination (OSCE): A curriculum in inpatient geriatric medicine. *J Am Geriatr Soc* 2004;52:1191–1198.
9. Supiano MA, Fitzgerald JT, Hall KE et al. A vertically integrated geriatric curriculum improves medical student knowledge and clinical skills. *J Am Geriatr Soc* 2007;55:1650–1655.
10. Simpson D, Helm R, Drewniak T et al. Objective Structured Video Examinations (OSVEs) for geriatrics education. *Gerontol Geriatr Educ* 2006;26:7–24.
11. Reznick RK, Blackmore D, Cohen R et al. An objective structured clinical examination for the licentiate of the Medical Council of Canada: From research to reality. *Acad Med* 1993;68(S1):S4–S6.
12. Schuwirth LW, van der Vleuten CP. The use of clinical simulations in assessment. *Med Educ* 2003;37(S1):65–71.
13. What to do about medical students with unsatisfactory consultation skills? *Med Teach* 2006;28:443–446.