Use of a Standardized Patient Exercise to Assess Core Competencies During Fellowship Training

Curtis T. Barry, MD
Uri Avissar, MD
Maureen Asebrook, AB
Michael A. Sostok, MD
Kenneth E. Sherman, MD, PhD
Stephen D. Zucker, MD

Abstract

Background The Accreditation Council for Graduate Medical Education requires fellows in many specialties to demonstrate attainment of 6 core competencies, yet relatively few validated assessment tools currently exist. We present our initial experience with the design and implementation of a standardized patient (SP) exercise during gastroenterology fellowship that facilitates appraisal of all core clinical competencies.

Methods Fellows evaluated an SP trained to portray an individual referred for evaluation of abnormal liver tests. The encounters were independently graded by the SP and a faculty preceptor for patient care, professionalism, and interpersonal and communication skills using quantitative checklist tools. Trainees' consultation notes were scored using predefined key elements (medical knowledge) and subjected to a coding audit (systems-based practice). Practice-based learning and improvement was addressed via verbal feedback from the SP and self-assessment of the videotaped encounter.

Results Six trainees completed the exercise. Second-year fellows received significantly higher scores in medical knowledge (55.0 ± 4.2 [standard deviation], $P = .05$) and patient care skills (19.5 ± 0.7, $P = .04$) by a faculty evaluator as compared with first-year trainees (46.2 ± 2.3 and 14.7 ± 1.5, respectively). Scores correlated by Spearman rank (0.82, $P = .03$) with the results of the Gastroenterology Training Examination. Ratings of the fellows by the SP did not differ by level of training, nor did they correlate with faculty scores. Fellows viewed the exercise favorably, with most indicating they would alter their practice based on the experience.

Conclusions An SP exercise is an efficient and effective tool for assessing core clinical competencies during fellowship training.

Curtis T. Barry, MD, is a Fellow in Gastroenterology in the Division of Digestive Diseases at the University of Cincinnati; **Uri Avissar, MD,** is Associate Professor of Medicine in the Section of Gastroenterology at Boston Medical Center; **Maureen Asebrook, AB,** is a Research Associate in the University of Cincinnati Simulation Center; **Michael A. Sostok, MD,** is Professor of Medicine, Assistant Dean for Medical Education, and Director of the Simulation Center at the University of Cincinnati; **Kenneth E. Sherman, MD, PhD,** is the Gould Professor of Medicine and Director of the Division of Digestive Diseases at the University of Cincinnati; and **Stephen D. Zucker, MD,** is Associate Professor of Medicine and the Director of the Gastroenterology Training Program in the Division of Digestive Diseases at the University of Cincinnati.

This study was supported by a Medical Education Project Seed Grant to Stephen D. Zucker from the University of Cincinnati College of Medicine. Preliminary reports of this work have been published in abstract form.

The authors gratefully acknowledge Amy Decker, CPC, for assisting with coding assessments and Gretchen Niebling for coordinating the exercise. We also thank Timothy Lewis, MD, for helpful advice.

Editor's Note: The online version of this article contains a sample case scenario, instructions for the standardized patients, instructions for the trainees, referral letter, patient encounter form, consultation note template, consultation note checklist, faculty patient care checklist, standardized patient assessment checklist, consultation note coding audit tool, trainee self-reflective questionnaire, postexercise fellow survey, and guidelines for the postexercise discussion session.

Corresponding author: Stephen D. Zucker, MD, Division of Digestive Diseases, University of Cincinnati, 231 Albert B. Sabin Way, ML 0595, Cincinnati, OH 45267-0595, zuckersd@email.uc.edu

Received June 25, 2009; revisions received October 28, 2009 and December 14, 2009; accepted January 21, 2010.

DOI: 10.4300/JGME-D-09-00001.1

Introduction

In 1999, the Accreditation Council for Graduate Medical Education (ACGME) introduced a set of 6 essential competencies for all training programs: patient care, medical knowledge, practice-based learning and improvement, interpersonal and communication skills, professionalism, and systems-based practice.[1,2] Assessing proficiency in each of these competencies is a program requirement for all core programs and the majority of fellowship programs, and the degree to which trainees achieve these competencies is a primary accreditation standard. Yet the means by which fellowship programs achieve these mandates are imprecise because few of the endorsed assessment techniques have been adapted or validated for subspecialty training, and the small size of many programs (an average of 5 trainees annually) is a barrier to statistical analysis and tool validation. Program directors face the challenge of identifying, implementing, and substantiating specialty-specific instruments to evaluate trainee performance.[3]

The standardized patient (SP) encounter is a well-established and reliable tool for teaching and appraising competency.[4] The advantages of SP-based exercises are that they can be tailored to assess several clinical skills

simultaneously, they permit comparisons to be made across multiple trainees, and they allow for the longitudinal assessment of program effectiveness. Standardized patient encounters can also provide more immediate, informed, and impartial feedback to the trainee than can be achieved with patient surveys. For these reasons, development of an SP repertoire is an attractive approach for evaluating competency.[5]

Standardized patient encounters have been widely implemented in medical school curricula,[4,6] residency programs,[7–11] and the United States Medical Licensing Examination to measure clinical competency.[12] They generally take the form of an objective structured clinical examination (OSCE), which consists of brief (5- to 10-minute) stations designed to assess a discrete skill. Less common is the simulated clinical encounter (SCE), which encompasses a longer medical scenario that evaluates multiple clinical skills under conditions that closely mimic an actual patient encounter. Although some programs report using OSCEs to evaluate core competencies during subspecialty training,[13–15] the use of an SCE to assess clinical competency during fellowship has not been described previously. Given the dearth of established subspecialty assessment tools, the aim of the present study was to develop, implement, and provide preliminary validation of an SP exercise that evaluates all core clinical competencies within the context of gastroenterology fellowship training.

Methods

An SCE was developed under the auspices of the University of Cincinnati Simulation Center. The study protocol was assigned exempt status by the University of Cincinnati Institutional Review Board.

Case Development

A case scenario, modeled on the initial office encounter for a patient evaluated in the faculty clinic for abnormal liver tests, was used to construct a deidentified "script." Because a definitive diagnosis was ultimately established by liver biopsy, follow-up in the form of clinical course, laboratory results, and histology report was available, thereby heightening the relevance of the feedback provided to the trainees.

Recruitment and Training of Standardized Patients

Standardized patients were recruited from an available pool of candidates based on the required attributes. At an initial 1-hour training session, SPs were provided a copy of the script for memorization and an overview of the structure and purpose of the exercise. One week prior to the SCE, a mock encounter between a faculty preceptor and one of the SPs was held, with the other SPs observing. An ensuing group discussion addressed queries and suggestions regarding attire, behavior, and makeup (to simulate jaundice), and adjustments were made to ensure consistent responses and accurate symptom portrayal. The SPs also received instruction regarding the provision of feedback to the trainees, which encompassed a discussion of the nature and type of verbal feedback and a detailed review of the SP checklist questionnaire.

Execution of the Exercise

The SCE was designed to simulate a new patient consultation that is typical of what trainees encounter in their ambulatory continuity clinics. Fellows were provided with oral and written instructions detailing the objectives and format of the SCE along with a brief referral letter containing pertinent clinical, laboratory, and imaging information. The trainees were given 1 hour to complete the encounter (consisting of a specialty-focused history, physical examination, and the consultation note). Following the introduction, fellows received an encounter form containing the patient's vital signs, and were directed to an examination room in which the SP was waiting. The encounters were observed live and were also videotaped. An overhead message indicated when 40 minutes had elapsed to help trainees budget their remaining time.

After the encounter, fellows formulated a consultation note using a standard template. Concurrently, SPs completed a computerized checklist assessing the fellow's performance. At the conclusion of the exercise, fellows submitted their notes, received 10 minutes of confidential verbal feedback from the SP, and were then given unlimited time to watch a computer-based, viewer-controlled audio/video file of their encounter while completing a self-reflective questionnaire.

Core Competency Assessment Tools

Assessment tools were developed for each of the ACGME-defined core competencies. Medical knowledge was quantified by a blinded faculty member who evaluated the trainees' consultation notes for predefined key elements that were compiled in a checklist format. Four key components of the consult documentation were assessed, each including essential elements (devised by experienced faculty), with a single point awarded for the inclusion of each individual element within the note. Points were distributed across the key components as follows: history (25), physical examination (10), assessment (26), and plan (24).

Patient care, professionalism, and interpersonal and communication skills were graded independently by the SP and by a single faculty preceptor with 25 years of experience mentoring and training students, residents, and fellows. Because of the observational nature of the analysis, the preceptor could not be blinded. All 3 competencies were evaluated simultaneously using a composite assessment tool. The faculty preceptor used a patient care checklist derived from the Harvard Medical

School Communications Skills Form.[16] Performance categories comprised 3 elements, with a single point awarded for an element that was satisfactorily addressed by the trainee. An SP checklist adapted from the University of Illinois at Chicago College of Medicine Patient-Centered Communication and Interpersonal Skills scale[8,16] and the American Urological Association SP assessment tool[3] rated trainee performance using a 3-point scale (3 = excellent, 2 = satisfactory, and 1 = needs improvement) and included a brief narrative highlighting the trainee's strengths and weaknesses.

Systems-based practice was evaluated via a coding audit of the trainees' consultation notes performed by an experienced billing specialist using a standard coding tool employed by the University of Cincinnati Medical Center. Current procedural terminology codes for a new outpatient consultation were assigned based on the elements of the history, physical examination, and complexity of medical decision making documented in the note. Reimbursement was calculated based on current Medicare fee scales. Specific feedback regarding how the documentation could be better optimized was provided to each fellow.

Practice-based learning and improvement was addressed by having the trainees review the videotape of their encounter and complete a self-reflective questionnaire developed by the Royal College of Physicians and Surgeons of Canada and designed to encourage introspection on personal strengths and weaknesses.[17] Additionally, at the conclusion of the postexercise discussion session, fellows completed an anonymous postexercise survey evaluating the utility of the SCE and its potential impact on their clinical practice.

Validation of Assessment Tools

In an attempt to validate the faculty consultation note and patient care assessment tools, individual trainee evaluations were correlated with the scores they received on the national Gastroenterology Training Examination (GTE) administered in April 2008. Although the GTE has traditionally been considered a tool for assessing medical knowledge, it has previously been shown that examination scores correlate with measures of patient care, communication, and professionalism.[18,19] Because the results of the GTE were distributed in late May, subsequent to the completion and grading of the SCE, participants and evaluators were blinded to the scores.

Provision of Feedback to the Trainees

Trainees received individualized and global feedback covering all core competencies through a variety of formats and mechanisms. Immediate critique of patient care, professionalism, and interpersonal and communication skill competencies was provided via the face-to-face meeting with the SP, and feedback regarding practice-based learning and improvement was provided by the self-reflective questionnaire completed by each trainee.

Additional feedback was provided during a 1-hour discussion session held 10 days after completion of the SCE. The consultation note checklists and coding audits, the assessments of the encounter by the faculty and the SP, and the fellows' self-reflective questionnaires were scored and tabulated in preparation for the session, which was moderated by an experienced faculty preceptor and attended by all of the participating fellows. It was designed to address each of the ACGME core competencies. Global feedback was provided in aggregate form, and individualized feedback was conferred via confidential packets containing formative evaluations of each trainee's personal performance. With regard to medical knowledge, fellows were furnished with copies of their consultation note along with the scored faculty checklist. To facilitate a comparison of their appraisal with that of an experienced faculty member, a deidentified copy of the original consultation note for the patient on whom the exercise was based was also provided. The packets also contained the scored faculty and SP assessments of the encounter, addressing patient care, professionalism, and interpersonal and communication skills; copies of their individual self-reflective questionnaires (practice-based learning and improvement); and a summary of the coding audit of their consultation note, including the Medicare reimbursement value of the consultation and suggestions for improvement (systems-based practice).

Statistical and Cost Analyses

Data were analyzed using a computer-based statistical package (Statistix 9.0, Analytical Software, Tallahassee, FL). Differences between group scores were analyzed by analysis of variance with Scheffe comparison of means. Correlations were assessed by Spearman rank. The reported cost of the exercise reflects actual expenditures.

Results

Six of the 8 fellows enrolled in the gastroenterology training program at the University of Cincinnati completed the SCE. Three were in their first year of fellowship training, 2 were in their second year, and 1 was in the final year of training. The SCE was held 10 months into the academic training cycle (April 2008) and was executed in 2 back-to-back 90-minute sessions employing the same 3 SPs. Three trainees were randomly assigned to each session, the overlapping nature of which prevented communication between participants.

Medical Knowledge Assessment

Evaluation of medical knowledge via a checklist-based assessment of the fellows' consultation notes (TABLE 1) showed that mean scores increased from the first (46.2 ± 2.3) to the second (55.0 ± 4.2) year of training

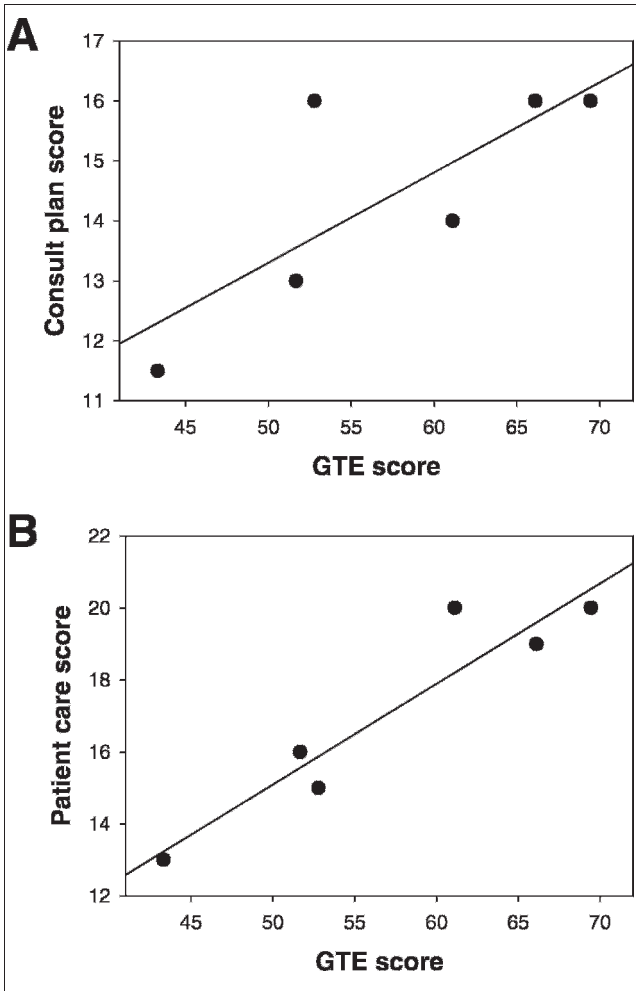

FIGURE **Association Between the Results of the Gastroenterology Training Examination (GTE) and Faculty Medical Knowledge and Patient Care Checklist Scores**

Panel A displays a plot of the score received on the "plan" component of the consultation note checklist versus the trainee's GTE score. The solid line reflects a linear fit of the data (r^2 = 0.604). In panel B, GTE scores are plotted against the trainee's faculty checklist score (r^2 = 0.886).

(P = .05), paralleling the results of the GTE examination (49.3 ± 5.2 vs 63.6 ± 3.5, P = .05). Because there was only a single third-year fellow, statistical comparisons could not be made. There was a significant (P = .03) correlation by Spearman rank (0.82) between the scores the trainees received on the consult note "plan" (reflective of trainee knowledge) and the GTE (FIGURE, panel A).

Evaluation of Patient Care, Professionalism, and Interpersonal and Communication Skills

Evaluation of the videotaped encounters using a 3-point patient care checklist showed that the mean aggregate score for second-year fellows was significantly higher than that for first-year fellows (19.5 ± 0.7 vs 14.7 ± 1.5, P = .04), with no differences between individual competencies identified. There was a highly significant correlation (0.84, P = .03) between checklist scores and GTE results (FIGURE, panel B). The mean score received by the trainees on the SP checklist was 27.0 ± 4.7 (82%), with no aggregate or competency-specific differences by year of training. The average total score and global score were 2.5 ± 0.43 and 2.5 ± 0.55, respectively, out of a possible 3 points. No correlation between the SP checklist score and either the faculty patient care checklist or GTE score was found.

Systems-Based Practice Appraisal

Consultation note audits supported a level IV new office consultation in 3 cases, a level III in 2 cases, and a level II in 1 case. Mean Medicare reimbursement for first- and second-year fellows was $126.74 ± $44.47 and $174.59 ± $0.00, respectively. Lower coding levels resulted from an insufficiently detailed history in all cases (inadequate family history in 2 cases, unsatisfactory history of present illness in 1 case, and <10 element review of systems in 1 case). All trainees appropriately documented a comprehensive physical examination and decision making of moderate complexity.

Assessment of Practice-Based Learning and Improvement

The mean score on the self-reflective questionnaire completed by the trainees after reviewing a videotape of their individual encounter was 22.3 ± 3.7 (83%), with no significant differences by year of training. There was no correlation between the self-reflective scores and checklist scores assigned by either the faculty or the SPs. As a group, the fellows thought they were least effective in using the medical literature to answer queries.

Trainee Feedback

During the postexercise feedback session, fellows discussed the differential diagnosis and management plan for the patient (medical knowledge). Follow-up clinical and laboratory data were then provided, the liver histology report was reviewed, and teaching points were highlighted. A subsequent review of the compiled results focused on effective behaviors and practices by the trainees, and also identified common pitfalls. The coding audit findings led to discussions about medical reimbursement, the importance of documentation, and effective communication with referring physicians (systems-based practice). Fellows received confidential packets containing the results of their individual assessments, after which their perceptions of the SCE were solicited.

Postexercise Survey

The results of the confidential survey assessing the trainees' opinion of the exercise are described in TABLE 2. All believed that performing an SCE annually would be beneficial. Aspects of the exercise deemed most useful

TABLE 1 **Comparison of Consultation Note Assessment Scores Between First- and Second-Year Trainees**[a]

Component	Maximum Score	First-Year Mean Score (SD)	Second-Year Mean Score (SD)	*P* Value
History	25	14.3 (2.1)	17.5 (3.5)	NS
Physical examination	10	6.5 (0.9)	7.8 (0.4)	NS
Assessment	26	11.8 (0.8)	14.8 (0.4)	.02
Plan	24	13.5 (2.3)	15.0 (1.4)	NS
Total	**85**	**46.2 (2.3)**	**55.0 (4.2)**	**.05**

Abbreviation: NS, not significant.

[a] Mean (SD) scores received for each component of the consultation note, along with the total aggregate score, are stratified by year of training. The maximum possible score and *P* value for the analysis of variance are also indicated.

included the review of the encounter videotape, the verbal feedback from the SP, and the educational content of the case.

Exercise Costs

The net cost of the SCE was $1190. The principal expenditure was for compensation of the SPs ($40.00 per hour plus parking), consisting of $275 for the coordinating SP and $213 each for the 2 additional SPs. The Simulation Center charges totaled $489, including room rental, consultative services, hiring and training of SPs, makeup, form preparation, and videotape processing. An additional $179 was allocated for lunches on the day of the exercise.

TABLE 2 **Postexercise Trainee Survey**[a]

Statement	Mean Score	SD
History was realistic	4.2	0.4
Physical examination was realistic	3.8	0.4
Effectively evaluated communication skills	4.0	0.0
Effectively evaluated examination skills	3.8	0.4
Valid way to test medical knowledge	4.0	0.0
Valid way to test professionalism	4.2	0.4
Valid way to test overall patient care	4.0	0.0
Videotaping was useful	3.8	0.4
Will change the way I practice	3.8	0.4
Objectives were met	4.2	0.4
Annual exercise would be useful	4.0	0.0

[a] Trainees responded to statements about the simulated clinical encounter using the following scale: 1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, and 5 = strongly agree. The mean (SD) scores for all 6 respondents are provided.

Discussion

We present our initial experience with the development and execution of an SCE for assessing all 6 core competencies in the context of subspecialty training. Although there have been a few reports describing the use of OSCEs to evaluate individual clinical skills during fellowship,[13–15,20] this report is the first to use an SCE format to assess competency during fellowship training. The SP encounter was designed to appraise advanced skills required for specialty practice; we also created a novel consult note assessment instrument and judiciously employed checklists adapted from existing tools in order to permit feedback to be both efficient and quantitative. Our intent was to provide a template that can be readily modified to address program-specific interests and objectives.

One limitation to our study is the low power reflecting the small number of participants. At the same time, our sample size is highly representative of what a typical subspecialty program must evaluate. Despite the small cohort, we were able to demonstrate improved competency in medical knowledge and patient care between the first and the second year of training. The correlation between GTE results and faculty checklist scores lends further credence to the validity of the exercise in assessing competency. Our findings are consistent with those of other investigators who have shown that written board examination scores closely parallel measures of patient care, communication, and professionalism.[8,18,19,21–23]

A second limitation of our study is that it does not directly control for interobserver variability. This concern is partly offset by employing a single faculty evaluator. On the other hand, the use of multiple SPs is crucial to executing a streamlined exercise, particularly for larger fellowship programs. It has been well established that, with proper training, different SPs portraying the same case yield consistent and reliable evaluation results.[12,24–27] As previously noted by other investigators,[14,26,28,29] we found no correlation between SP checklist scores and trainee experience, GTE scores, or faculty checklist results. Even

though this lack of correlation could be construed as evidence of bias by the unblinded faculty or of variability in SP scoring, the observation that the average patient care scores assigned by the faculty (82%) and the SPs (82%) were equivalent argues against these possibilities.

Finally, as is true of most available assessment instruments, the capacity of the present SCE to evaluate interprogram competency is unknown. To adequately address this issue would require the coordinated implementation of the exercise across multiple programs. The SP exercises unquestionably provide for more consistent grading than can be achieved with other recommended tools that rely on actual patients, such as satisfaction surveys and 360-degree assessments. Our findings support the attractiveness of using an SCE as a patient-based learning experience of self-reflection and self-evaluation.

It may be argued that expanding the exercise to encompass all 6 competencies has the potential to dilute utility. Although it is true that some of the assessment tools employed in this exercise (eg, billing audit, consult note checklist) could be readily applied to an alternative clinical setting, the advantage of assessing all competencies in the context of a single, well-constructed SCE is in its uniformity with regard to background information, clinical presentation, and time constraints. The consistency of the encounter also facilitates a more direct and ready comparison of competency between individual trainees, across groups of trainees (by year of training as well as longitudinally), and, potentially, across training programs. We believe that there is value in developing methods to efficiently assess competency in light of the escalating regulatory burden on training programs[30] and the increasing program director turnover.[31,32]

From an implementation perspective, we identified 3 impediments to implementing an SP exercise. The first is the time and effort required for initial development and planning. We hope that this report will streamline the process for other programs by providing a comprehensive framework and relevant assessment tools. The second hurdle is access to appropriate facilities. For institutions that lack facilities specifically designed for such exercises, however, a standard examination room with a well-positioned video camera would suffice. A final obstacle is the cost, which approached $1200 for the present exercise. The principal fixed expense was for SP salaries (~$700) and the remainder was spent on initial development of the SCE and would be lower for subsequent iterations.

Conclusions

Because of its versatility, our SCE template can be readily adapted to suit most subspecialty programs and further modified to address psychosocial issues including substance abuse, noncompliance, or the breaking of bad news.[9] It could also be extended to not only assess, but also teach clinical competency. For instance, the faculty patient care assessment revealed that our trainees routinely failed to elicit the patient's beliefs, concerns, and expectations about their illness. A reinvigorated focus on these issues could help improve the effectiveness of our fellows in caring for their patients.

The present SCE focuses on hepatology knowledge, and we propose devising additional exercises to assess topics such as pancreaticobiliary disorders or inflammatory bowel disease. Ultimately, we envision integrating an SP exercise into our program curriculum on an annual basis to serve as an anchor from which to build a summative evaluation portfolio. As gastroenterology training is 3 years in duration, rotation of cases on a triennial cycle would minimize development time, reduce costs, and facilitate longitudinal assessment of educational effectiveness.

References

1. Reisdorff EJ, Hayes OW, Carlson DJ, Walker GL. Assessing the new general competencies for resident education: a model from an emergency medicine program. *Acad Med*. 2001;76:753–757.
2. Lee AG, Carter KD. Managing the new mandate in resident education: a blueprint for translating a national mandate into local compliance. *Ophthalmology*. 2004;111:1807–1812.
3. Miller DC, Montie JE, Faerber GJ. Evaluating the accreditation council on graduate medical education core clinical competencies: techniques and feasibility in a urology training program. *J Urol*. 2003;170:1312–1317.
4. Barrows HS. An overview of the uses of standardized patients for teaching and evaluating clinical skills. *Acad Med*. 1993;68:443–451.
5. Swing SR. Assessing the ACGME general competencies: general considerations and assessment methods. *Acad Emerg Med*. 2002;9:1278–1288.
6. Bennett AJ, Arnold LM, Welge JA. Use of standardized patients during a psychiatry clerkship. *Acad Psychiatry*. 2006;30:185–190.
7. Yudowsky R, Alseidi A, Cintron J. Beyond fulfilling the core competencies: an objective structured clinical examination to assess communication and interpersonal skills in a surgical residency. *Curr Surg*. 2004;61:499–503.
8. Yudowsky R, Downing SM, Sandlow LJ. Developing an institutional-based assessment of resident communication and interpersonal skills. *Acad Med*. 2006;81:1115–1122.
9. Quest TE, Otsuki JA, Banja J, Ratcliff JJ, Heron SL, Kaslow NJ. The use of standardized patients within a procedural competency model to teach death disclosure. *Acad Emerg Med*. 2002;9:1326–1333.
10. Biernat K, Simpson D, Duthie E, Bragg D, London R. Primary care residents self assessment skills in dementia. *Adv Health Sci Educ*. 2003;8:105–110.
11. Kissela B, Harris S, Kleindorfer D, et al. The use of standardized patients for mock oral board exams in neurology: a pilot study. *BMC Med Educ*. 2006;6:22.
12. Dillon GF, Boulet JR, Hawkins RE, Swanson DB. Simulations in the United States Medical Licensing Examination (USMLE). *Qual Saf Health Care*. 2004;13(suppl 1):i41–i45.
13. Chander B, Kule R, Baiocco P, et al. Teaching the competencies: using objective structured clinical encounters for gastroenterology fellows. *Clin Gastroenterol Hepatol*. 2009;7:509–514.
14. Boudreau D, Tamblyn R, Dufresne L. Evaluation of consultative skills in respiratory medicine using a structured medical consultation. *Am J Respir Crit Care Med*. 1994;150:1298–1304.
15. O'Sullivan PO, Chao S, Russell M, Levine S, Fabiny A. Development and implementation of an objective structured clinical examination to provide formative feedback on communication and interpersonal skills in geriatric training. *J Am Geriatr Soc*. 2008;56:1730–1735.
16. Makoul G. The SEGUE framework for teaching and assessing communication skills. *Patient Educ Couns*. 2001;45:23–34.
17. Hall W, Violato C, Lewkonia R, et al. Assessment of physician performance in Alberta: the Physician Achievement Review. *Can Med Assoc J*. 1999;161:52–57.
18. Ramsey PG, Carline JD, Inui TS, Larson EB, LoGerfo JP, Wenrich MD. Predictive validity of certification by the American Board of Internal Medicine. *Ann Intern Med*. 1989;110:719–726.

19 Brennan TA, Horwitz RI, Duffy FD, Cassel CK, Goode LD, Lipner RS. The role of physician specialty board certification status in the quality movement. *JAMA*. 2004;292:1038–1043.

20 Varkey P, Gupta P, Bennet KE. An innovative method to assess negotiation skills necessary for quality improvement. *Am J Med Qual*. 2008;23:350–355.

21 Shea JA, Norcini JJ, Kimball HR. Characteristics of licensure and certification examinations: relationships of ratings of clinical competence and ABIM scores to certification status. *Acad Med*. 1993;68:S22–S24.

22 Tamblyn R, Abrahamowicz M, Dauphinee D, et al. Physician scores on a national clinical skills examination as predictors of complaints to medical regulatory authorities. *JAMA*. 2007;298:993–1001.

23 Norcini JJ, Webster GD, Grosso LJ, Blank LL, Benson JA. Ratings of residents' clinical competence and performance on certification examination. *J Med Educ*. 1987;62:457–462.

24 Badger LW, DeGruy F, Hartman J, et al. Stability of standardized patients' performance in a study of clinical decision making. *Fam Med*. 1995;27:126–131.

25 Varkey P, Natt N, Lesnick T, Downing S, Yudowsky 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;83:775–780.

26 Stillman PL, Swanson DB, Smee S, et al. Assessing clinical skills of residents with standardized patients. *Ann Intern Med*. 1986;105:762–771.

27 Tudiver F, Rose F, Banks B, Pfortmiller D. Reliability and validity testing of an evidence-based medicine OSCE station. *Fam Med*. 2009;41:89–91.

28 Gimpel JR, Boulet JR, Errichetti AM. Evaluating the clinical skills of osteopathic medical students. *J Am Osteopath Assoc*. 2003;103:267–279.

29 McKinstry B, Wlaker J, Blaney D, Heaney D, Begg D. Do patients and expert doctors agree on the assessment of consultation skills? A comparison of two patient consultation assessment scales with the video component of the MRCGP. *Fam Pract*. 2004;21:75–80.

30 Dickinson TA, Clayton CP. The future of fellowship education in internal medicine. *Am J Med*. 2004;116:720–723.

31 Beasley BW, Kern DE, Kolodner K. Job turnover and its correlates among residency program directors in internal medicine: a three-year cohort study. *Acad Med*. 2001;76:1127–1135.

32 Dyrbye LN, Shanafelt TD, Thomas MR, Durning SJ. Brief observation: a national study of burnout among internal medicine clerkship directors. *Am J Med*. 2009;122:310–312.