

Faculty and the Observation of Trainees' Clinical Skills: Problems and Opportunities

Eric S. Holmboe, MD

ABSTRACT

The clinical skills of medical interviewing, physical examination, and counseling remain vital to the effective care of patients, yet research continues to document serious deficiencies in clinical skills among students and residents. The most important method of evaluation is the direct observation of trainees performing these clinical skills. Standardized patients and other simulation technologies are important and reliable tools for teaching clinical skills and evaluating competence and will be incorporated in the near future as part of the United States Medical Licensing Examination. Standardized patients and simulation, however, cannot and should not

replace the direct observation by faculty of trainees' clinical skills with actual patients. Faculty are in the best position to document improvement over time and to certify trainees have attained sophisticated levels of skill in medical interviewing, physical examination, and counseling. Unfortunately, current evidence suggests significant deficiencies in faculty direct observation evaluation skills. The author outlines the nature of the problems in clinical skills and their evaluation by faculty and ends with recommendations to improve the current state of faculty skills in evaluation.

Acad Med. 2004;79:16–22.

Recently the National Board of Medical Examiners (NBME), responsible for the administration of the three-part United States Medical Licensing Exam (USMLE), announced that they would institute a clinical skills exam (CSE) as a component of Part 2 of the USMLE for students in the medical school class of 2005.¹ The CSE will use ten to 12 different encounters with highly trained standardized patients (SPs) to test whether a fourth-year medical student has acquired a minimum level of competence in communication and physical examination skills. Although there has been substantial resistance from some medical organizations, a Harris poll found strong support among the public for such an examination.^{2,3}

However, the responsibility to certify competence in these clinical skills among residents in the United States will continue to fall solely on the residency program director and

associated teaching faculty. Observation and documentation of the clinical skills of residents is required by the Accreditation Council for Graduate Medical Education as part of its new outcomes-based accreditation project.^{4,5} Unlike Canada, no specialty in the United States includes a direct assessment of clinical skills as part of licensure or certification. The introduction of the USMLE CSE provides an opportunity to reflect on the states of the practice and, equally important, the evaluation of clinical skills among trainees.

In an era of rapidly advancing medical information and technology, how important are these clinical skills to the successful care of patients? If important, what is the quality of these skills among trainees and practicing physicians? With the availability of SPs and other simulation technology for evaluation, what role should physician–educators play in the evaluation of clinical skills?

I wrote this article to respond to these questions and to demonstrate that proficiency in clinical skills does matter, that the state of such skills among trainees and practicing physicians remains substandard, and that the main responsibility for teaching and evaluating these skills should remain with physician–educators. Medical educators must not abdicate this responsibility to SP and simulation; rather, such

Dr. Holmboe is associate professor of medicine and director of faculty development and evaluation, Department of Medicine, Yale University School of Medicine, New Haven, Connecticut.

Correspondence and requests for reprints should be addressed to Dr. Holmboe, Department of Medicine, Waterbury Hospital, Pomeroy 6, 64 Robbins Street, Waterbury, CT 06721; telephone: (203) 573-6573; fax: (203) 573-6707; e-mail: {eholmboe@wtbyhosp.chime.org}.

approaches should complement the teaching and evaluation of clinical skills.

THE IMPORTANCE OF CLINICAL SKILLS

Despite the explosion of technological advances, the clinical skills of medical interviewing, physical examination, and counseling remain vital to the successful care of patients. Despite documented limitations,⁶ studies over the last several decades have continually reaffirmed that a well-conducted medical interview and physical examination are still important diagnostic tools available to clinicians. Hampton and colleagues⁷ demonstrated that the medical history produced the final diagnosis in the majority of patients, with laboratory investigation providing the final diagnosis in only one of 80 consultations. More recently, Peterson et al.⁸ demonstrated that among 80 patients presenting to a primary care clinic with a previously undiagnosed condition, the history produced the correct final diagnosis in 76% of the visits. In his textbook, McGee⁶ highlights the valuable predictive value of many commonly used interview instruments, such as the mini-mental status exam, in evaluating patients. Using autopsies of 400 patients as the standard, Kirch and Schaffi⁹ compared the accuracy of the history and physical examination with diagnostic imaging in producing the correct diagnosis. They found that the combination of history and physical examination was accurate in 70% of the cases, whereas diagnostic imaging produced a correct diagnosis in only 35% of them.⁹ Chimowitz et al.¹⁰ have demonstrated the importance of the bedside examination in the accurate diagnosis of neurological disorders.

Effective physician-patient communication has also been shown to improve health outcomes. Stewart¹¹ reviewed 21 studies involving physician-patient communication interventions and found that 16 studies reported positive effects on patient outcomes, including adherence. Levinson and colleagues¹² found that physicians with malpractice claims were less likely to use patient-centered interviewing skills than were physicians without malpractice claims. Another study found that patient satisfaction is significantly lower when physicians neglect psychosocial aspects of the interview.¹³ Finally, Bordage¹⁴ recently noted that poor data-collection skills were still a significant factor in diagnostic errors. In summary, medical interviewing, physical examination, and counseling remain the most important and effective diagnostic and therapeutic tools.

THE STATE OF TRAINEES' CLINICAL SKILLS

Despite the growing body of literature demonstrating the importance of clinical skills, studies continue to document

serious deficiencies. Although much has been learned about effective interviewing techniques, deficiencies in interviewing skills persist, and in the views of some may actually have declined.¹⁵⁻¹⁹ Furthermore, communication skills do not appear to improve even after completion of residency training. In a study using unannounced SPs, Ramsey et al.²⁰ found that a group of primary care physicians asked only 59% of essential history items. Braddock and colleagues found that among 1,057 counseling sessions involving primary care physicians and surgeons, only 9% of encounters met basic criteria for effective informed decision making.²¹ Other studies have shown that physicians fail to elicit over half of patient complaints and that many of the public complaints about physicians relate to communication problems.²²⁻²⁶ Another study highlighted the importance of patient non-verbal communication and how many physicians fail to recognize these important cues.²⁷

Errors are also common in physical examination skills. In a 1976 study, Weiner and Nathanson²⁸ documented numerous types of errors among trainees. Wray and Friedland²⁹ noted in a 1983 report that residents committed at least one physical examination error in 58% of the patients examined. Deficiencies in cardiac auscultatory skills among trainees were documented as early as the 1960s.^{30,31} Mangione and colleagues^{32,33} recently demonstrated that poor cardiac and pulmonary physical examination skills continue to plague students and residents. Other studies have reported poor skills by residents for many aspects of the basic physical exam.³⁴⁻³⁷

As a result of these known deficiencies in the basic clinical skills of history taking, physical examination, and counseling, there has been a significant push by medical educators and accrediting agencies to reemphasize both the training and evaluation of clinical skills.³⁸⁻⁴² Again, in the recent Harris poll over 95% of respondents rated communication and diagnostic skills as very or extremely important.³ Without accurate evaluation of clinical skills, which can only be accomplished by direct observation, improvement in the clinical skills of physician trainees is unlikely.

FACULTY AND CLINICAL SKILLS

Lack of Direct Observation of Trainees by Faculty

Perhaps the biggest problem in the evaluation of clinical skills is the lack of observation by faculty of trainees. For decades, faculty have taken at face value the veracity of the history and physical examination presented on inpatient and outpatient rounds without ever watching the trainee actually perform them. Two of the most prominent physician-scientists and physician-educators of the 20th century, the late

Alvan Feinstein and George Engel, strongly advocated direct observation of the history and physical examination skills of trainees over 30 years ago.^{43,44} Dr. George Engel commented in a 1976 editorial:

Evidently it is not deemed necessary to assay students' (and residents) clinical performance once they have entered the clinical years. Nor do clinical instructors more than occasionally show how they themselves elicit and check the reliability of clinical data. To a degree that is often at variance with their own professed scientific standards, attending staff all too often accept and use as the basis for discussion, if not recommendations, findings reported by students and housestaff without ever evaluating the reporter's mastery of the clinical methods utilized or the reliability of the data obtained.⁴⁵

Kassebaum and Eaglen⁴⁶ recently lamented that little had changed among medical schools over the last 25 years. While a growing proportion of medical schools now use SPs for some aspect of assessment, students continue to receive little direct observation from faculty. They also noted that by 1998 only 48% of medical schools used a standardized patient assessment in the fourth year. In the recent field trial of the USMLE CSE, nearly 40% of the 858 participating medical students reported they had been observed by a faculty member performing a history and physical four or fewer times during medical school. More concerning was the finding that residents and not faculty completed the majority of student observations.³

The state of affairs in residents' education is no better. Even in a recent trial of the mini-clinical evaluation exercise (mini-CEX) by the American Board of Internal Medicine, several of the participating programs had a difficult time getting teaching faculty to complete just four direct observations of each intern over the course of a year.⁴⁷ Two other studies in the early 1990s found faculty were not able to discriminate between the various dimensions of clinical competence. In fact, both studies suggested that faculty were only evaluating two dimensions of residents' performance: case-based medical knowledge and interpersonal skills.^{48,49} One of the hypothesized reasons for these findings in both studies was that faculty rarely, if ever, directly observed the performance of the resident they were evaluating.

Similarly, another study found poor correlation between ward evaluations by attending faculty and objective structured clinical examination scores by residents, suggesting that the ward evaluations were not based on direct observation of clinical skills.⁵⁰ Finally, a recent randomized controlled trial primarily designed to improve the specificity of written comments by faculty on a general medicine rotation found that residents reported they were infrequently ob-

served performing a history or parts of the physical examination during their ward rotation.⁵¹ Noting this difficulty getting faculty to teach and observe, Johnson and Boohan⁴⁰ recently argued in frustration that the training and evaluation of clinical skills should not be "left to teaching hospitals."

Quality of Faculty Observation Skills

Is the problem simply insufficient quantity of observation by faculty or is quality also suspect? Unfortunately, little data exist on the quality of faculty observation of clinical skills, but the available evidence suggests significant deficiencies exist. Noel and Herbers and their colleagues, in two important studies of the traditional complete CEX, found substantial deficiencies in the accuracy of faculty ratings.^{52,53} They demonstrated that faculty failed to detect 68% of errors committed by a resident when observing a videotape scripted to depict marginal performance. Use of checklists prompting faculty to look for specific skills increased accuracy of error detection from 32% to 64%, but the checklist did not improve overall accuracy of ratings of competence. More disturbing was that 69% of faculty still rated the overall performance of a resident depicting marginal performance as satisfactory or superior. Also, showing faculty a brief informational videotape about the traditional CEX and its purpose failed to improve the quality of ratings.⁵²

Elliot and Hickam⁵⁴ also noted that faculty observers did not reliably evaluate up to 32% of physical examination skills, especially involving the head, neck, and abdomen, among medical students. Kalet et al.⁵⁵ examined the reliability and validity of faculty observation of interviewing skills using videotapes of student performance on an objective structured clinical examination. They found that faculty were inconsistent in identifying the use of open-ended questions and empathy, and that the positive predictive value of faculty ratings for "adequate" interviewing skills was only 12%. Of the eight faculty raters involved in the study, six were instructors in the student medical-interviewing course.⁵⁵

Role of Faculty in the Evaluation of Clinical Skills

Given the shortcomings with faculty evaluations, should this responsibility be taken away from physician-educators? The technology of SPs has been in use and studied for over 30 years,⁵⁶⁻⁶⁰ and the capability of simulators continue to improve at a rapid pace.⁶¹ Much data exist on the reliability and validity properties of SPs. The Medical Council of Canada and the Educational Commission for Foreign Medical Graduates include clinical skills examinations as an integral component of the licensure process.^{62,63} As noted

previously, the NBME and the Federation of State Medical Boards will incorporate an SP-based clinical skills examination in the USLME starting in 2005. The development of SPs to evaluate clinical skills is unquestionably a major advance in competency assessment, and rigorous SP training and scoring methods support the reliability requirements needed for high-stakes examinations^{3,56-63}

There are limitations, however, in the application of SP-based methods for teaching and evaluation along the continuum of education and clinical practice. First, this technology is expensive and is not readily available in all areas and at all times. Second, SPs should supplement and not replace observation with actual patients. Standardized patients provide mostly a cross-sectional view of skills and only through the observation of actual patients can educators get a full picture of longitudinal changes and growth in clinical skills among trainees. Third, because most assessment instruments used for standardized patient exercises favor completeness over efficiency, SPs may have less validity with more advanced trainees.⁶⁴⁻⁶⁶ Finally, although the task for a trainee is often specified before even meeting the standardized patient, such information is often not available to a trainee or physician interacting with real patients.

A distinction must also be made between the *performance* of a trainee in a testing situation and a trainee's *performing* when caring for actual patients. *Performing* denotes the ongoing and continuous interaction between a patient and physician, something SPs cannot evaluate effectively. The noted educator George Miller described this difference as the physician's ability to "show how" (e.g., demonstrate that they can execute the clinical skill) versus what the physician "does" with actual patients who are experiencing real pain, fears, and uncertainty.⁶⁷ Standardized patients are unquestionably a powerful tool to assess whether a trainee can execute a skill (performance, or "show how") and, in my view, a valid vehicle to test for minimum levels of competence in graduating medical students. However, given the diversity and range of patients seen during training, medical educators are in the best position to evaluate the ongoing and continuous care of patients by trainees performing with actual patients who are not following a predefined script.^{67,68} Furthermore, direct observation provides a valuable template for meaningful feedback to reinforce strong clinical skills and correct deficiencies.⁶⁹

Medical educators, especially those involved with residencies, must also realize that tests of knowledge by outside organizations such as the NBME do not and cannot replace the observation of clinical skills by faculty. Research has repeatedly demonstrated that a multiple-choice examination cannot attest to a trainee's proficiency in clinical skills. The addition of the CSE can only assure that a medical student has attained a basic level of clinical skills sufficient to begin

the next stage of residency. The CSE was designed mostly to detect serious deficiencies among students before they matriculate into their residencies and to provide meaningful information about clinical skills as a part of state licensure. Dr. Jordan Cohen of the Association of American Medical Colleges noted that the CSE helps to fulfill an obligation to the public that the profession is ensuring that graduates have acquired these skills as a result of their medical school training.⁷⁰ However, the CSE does not ensure that the student who enters residency will necessarily acquire the higher level of clinical skills and judgment required for independent practice.

Finally, evaluation is at the heart of professionalism for the medical educator. Medical educators have a moral and professional obligation to ensure that any trainee leaving their training program has attained a minimum level of clinical skills to care for patients safely, effectively, and compassionately. Medical educators should not wait for the results of a standardized CSE or other examinations to learn whether their trainees possess sufficient clinical skills. This responsibility cannot be abdicated to SPs, licensing boards, or computer simulators. Trainees recognize the importance of these clinical skills; trainees also recognize they are not always adequately prepared to care for patients after graduating from a residency program.^{71,72} Furthermore, the majority of trainees want effective evaluation and feedback from their faculty.^{73,74}

CHALLENGES AND RECOMMENDATIONS

The major challenge that lies ahead for medical educators is how to ensure that the educators themselves possess strong clinical skills but also have the necessary skill to effectively observe, evaluate, and provide feedback to trainees' regarding clinical skills. The available research suggests that there is much work to be done to improve medical educators' evaluations clinical skills. Research in clinical teaching has demonstrated that teaching skills can be improved with faculty development that will result in positive outcomes for learners.⁷⁵⁻⁷⁸ Thus there is good reason to believe that properly designed training can also lead to improvements not only in faculty evaluation skills, but also in the faculty's own clinical skills.^{38,78} One thing is clear: brief faculty interventions without periodic reinforcement will not produce meaningful changes.⁵¹⁻⁵³

There are many barriers facing medical educators in this task. Time and financial pressures on clinical faculty continue to increase while government support for graduate medical education continues to decline.^{79,80} Many academic medical centers and medical schools face significant financial difficulties. Work-hour reforms in medical residencies, although long overdue, add further burdens to already overex-

tended clinical faculty. These challenges cannot, however, become excuses not to improve the evaluation of clinical skills by faculty. The public has spoken and expects us to move forward and improve. Clinical skills are important to patients, and they are clearly important to physicians who desire to be successful professionals. Therefore, I propose four broad-based recommendations as a starting point:

1. The evaluation of clinical skills by faculty must be given higher priority by medical school deans, department chairs, and residency program directors. This emphasis must translate into meaningful policies that encourage and reward the observation of clinical skills by faculty. For example, precepting in clinical setting should be counted toward productivity. A greater emphasis on faculty development in clinical skills and evaluation will also be needed, and recent research demonstrates that such faculty development can be effective.⁷⁸
2. Medical schools and residency programs should consider developing education champions and core groups of clinician-educator faculty who serve as the experts for clinical skills teaching and evaluation. Evaluation champions could potentially promote change like their counterparts in quality of care. These champions and core groups, once trained, can serve as mentors for other teaching faculty and can be role models of effective direct observation behaviors.⁸¹⁻⁸⁴
3. Medical schools and residencies will need to develop systems to "evaluate their evaluations." Although more direct observation is clearly desirable and will help to improve reliability, it does not ensure better validity and accuracy of direct observations. In essence, medical schools and residencies will need to measure the outcomes of their training of clinical skills.
4. Finally, national and regional organizations involved in certification, evaluation, and accreditation should provide meaningful resources for research and development of new faculty-training programs involving the evaluation of clinical skills. Few medical schools and fewer residencies will possess enough resources to develop such programs on their own.

In summary, improving the evaluation of clinical skills by faculty should be a high priority for medical schools and particularly residency programs. Residency may be the last real opportunity to ensure that the trainee has adequate clinical skills, especially because there are no clinical skills examinations for graduating residents. Therefore, the education community, in partnership with other stakeholders, should set a research agenda and develop mechanisms to provide meaningful support for these activities. We have

over 25 years of data documenting the myriad of problems; it is time to act more aggressively and proactively as an educational community to correct these problems. Although the need for better evaluation of clinical skills by faculty will have to compete with other needs and interests, encouraging such competition is long overdue.

This work was supported in part by a grant from the Robert Wood Johnson Foundation. The author is indebted to Dr. Louis Pangaro of the Uniformed Services University of the Health Sciences for his conceptual advice and review, and to Dr. Herbert Chase of Yale University for his thoughtful review of the manuscript.

REFERENCES

1. United States Medical Licensing Examination. An analysis of U.S. student field trial and international medical graduate certification test results for the proposed USMLE clinical skills examination (www.usmle.org/news/cse/ceftresults). Accessed 5 February 2003.
2. House of Delegates, American Medical Association. Resolution 308; proposed implementation of clinical skills assessment exam (www.ama-assn.org). Accessed 5 February 2003.
3. Steege/Thomson Communications. Clinical skills exam: frequently asked questions (www.usmle.org). Accessed 5 February 2003.
4. Accreditation Council for Graduate Medical Education. The outcomes project (www.acgme.org). Accessed 5 February 2003.
5. Batalden P, Leach D, Swing S, Dreyfus H, Dreyfus S. General competencies and accreditation in graduate medical education. An antidote to overspecification in the education of medical specialists. *Health Aff.* 2002;21:103-111.
6. McGee S. *Evidenced-based Physical Diagnosis*. Philadelphia: WB Saunders, 2001.
7. Hampton JR, Harrison MJG, Mitchell JRA, et al. Relative contributions of history-taking, physical examination, and laboratory investigation to diagnosis and management of medical outpatients. *BMJ.* 1975; 2:486-9.
8. Peterson MC, Holbrook JH, Hales DV, Smith NL, Staker LV. Contributions of the history, physical examination, and laboratory investigation in making medical diagnoses. *West J Med.* 1992;156: 163-5.
9. Kirch W, Schafiq C. Misdiagnosis at a university hospital in four medical eras. Report on 400 cases. *Medicine.* 1996;75:29-40.
10. Chimowitz MI, Logigian EL, Caplan LR. The accuracy of bedside neurological diagnoses. *Ann Neurol.* 1990;28:78-85.
11. Stewart MA. Effective physician-patient communication and health outcomes: a review. *CMAJ.* 1995;152:1423-33.
12. Levinson W, Roter DL, Mullooly JP, Dull VT, Frankel RM. Physician-patient communication: the relationship with malpractice claims among primary care physicians and surgeons. *JAMA.* 1997;277: 553-9.
13. Roter DL, Stewart M, Putnam SM, Lipkin MJr, Stiles W, Inui TS. Communication patterns of primary care physicians. *JAMA.* 1997;277: 350-6.
14. Bordage G. Why did I miss the diagnosis? Some cognitive explanations and educational implications. *Acad Med.* 1999;74(10 suppl):S138-S143.
15. Platt FW, McMath JC. Clinical hypocompetence: the interview. *Ann Intern Med.* 1979;91:898-902.
16. Meuleman JR, Caranasos GJ. Evaluating the interview performance of internal medicine interns. *Acad Med.* 1989;64:277-9.

17. Beaumier A, Bordage G, Saucier D, Turgeon J. Nature of the clinical difficulties of first year family medicine residents under direct observation. *Can Med Assoc J*. 1992;146:489-97.
18. Sachdeva AK, Loiacono LA, Amiel GE, Blair PG, Friedman M, Roslyn JJ. Variability in the clinical skills of residents entering training programs in surgery. *Surgery*. 1995;118:300-9.
19. Pfeiffer C, Madray H, Ardolino A, Williams J. The rise and fall of student's skill in obtaining a medical history. *Med Educ*. 1998;32:283-8.
20. Ramsey PG, Curtis R, Paauw DS, Carline JD, Wenrich MD. History-taking and preventive medicine skills among primary care physicians: an assessment using standardized patients. *Am J Med*. 1998;104:152-8.
21. Braddock CH, Edwards KA, Hasenberg NM, Laidley TL, Levinson W. Informed decision making in outpatient practice: time to get back to basics. *JAMA*. 1999;281:203-10.
22. Stewart MA, McWhinney IR, Buck CW. The doctor-patient relationship and its effect upon outcome. *J R Coll Gen Pract*. 1979;29:77-82.
23. Richards T. Chasms in communication. *BMJ*. 1990;301:1407-8.
24. Katz J. *The Silent World of Doctor and Patient*. New York: Free Press, 1984.
25. Simpson M, Buckman R, Stewart M, et al. Doctor-patient communication: the Toronto consensus conference. *BMJ*. 1991;303:1385-7.
26. Shapiro RS, Simpson DE, Lawrence SL, Talky AM, Sobocinski KA, Schiedermayer DL. A survey of sued and nonsued physicians and suing patients. *Arch Intern Med*. 1989;149:2190-6.
27. Suchman AL, Markakis K, Beckman HB, Frankel R. A model of empathic communication in the medical interview. *JAMA*. 1997;277:678-82.
28. Weiner S, Nathanson M. Physical examination: frequently observed errors. *JAMA*. 1976;236:852-5.
29. Wray NP, Friedland JA. Detection and correction of house staff error in physical diagnosis. *JAMA*. 1983;249:1035-7.
30. Butterworth JS, Reppert EH. Auscultatory acumen in the general medical population. *JAMA*. 1960;174:32-4.
31. Raferty EB, Holland WW. Examination of the heart: an investigation into variation. *Am J Epidemiol*. 1967;85:438-44.
32. Mangione S, Nieman LZ. Cardiac auscultatory skills of internal medicine and family practice trainees: a comparison of diagnostic proficiency. *JAMA*. 1997;278:717-22.
33. Mangione S, Burdick WP, Peitzman SJ. Physical diagnosis skills of physicians in training: a focused assessment. *Acad Emerg Med*. 1995;2:622-9.
34. Li JTC. Assessment of basic examination skills of internal medicine residents. *Acad Med*. 1994;69:296-9.
35. Johnson JE, Carpenter JL. Medical house staff performance in physical examination. *Arch Intern Med*. 1986;146:937-41.
36. Fox RA, Clark CLI, Scotland AD, Dacre JE. A study of pre-registration house officers' clinical skills. *Med Educ*. 2000;34:1007-12.
37. Todd IK. A thorough pulmonary exam and other myths. *Acad Med*. 2000;75:50-1.
38. Turnbull J, Gray J, MacFayen J. Improving in-training evaluation programs. *J Gen Intern Med*. 1998;13:317-23.
39. Duffy DF. Dialogue: the core clinical skill. *Ann Intern Med*. 1998;128:139-41.
40. Johnson BT, Boohan M. Basic clinical skills: don't leave teaching to the teaching hospitals. *Med Educ*. 2000;34:692-9.
41. Cunnington JPW, Hanna E, Turnbull J, Kaigas TB, Norman GR. Defensible assessment of the competency of the practicing physician. 1997;72:9-12.
42. Long DM. Competency-based residency training: the next advance in graduate medical education. *Acad Med*. 2000;75:1178-83.
43. Feinstein AR. *Clinical Judgement*. Baltimore: Williams & Wilkins, 1967:1-71, 291-349.
44. Engel GL. The deficiencies of the case presentation as a method of teaching: another approach. *N Engl J Med*. 1971;284:20-4.
45. Engel GL. Are medical schools neglecting clinical skills? *JAMA*. 1976;236:861-3.
46. Kassebaum DG, Eaglen RH. Shortcomings in the evaluation of students' clinical skills and behaviors in medical school. *Acad Med*. 1999;74:842-9.
47. Norcini JJ, Blank LL, Arnold GK, Kimball HR. The mini-CEX (clinical evaluation exercise): a preliminary investigation. *Ann Intern Med*. 1995;123:795-9.
48. Thompson WG, Lipkin M, Gilbert DA, Guzzo RA. Evaluating evaluation: assessment of the American Board of Internal Medicine resident evaluation form. *J Gen Intern Med*. 1990;5:214-7.
49. Haber RJ, Avins AL. Do ratings on the American Board of Internal Medicine resident evaluation form detect differences in clinical competence? *J Gen Intern Med*. 1994;9:140-5.
50. Schwartz RW, Donnelly MB, Sloan DA, Johnson SB, Strodel WE. The relationship between faculty ward evaluations, OSCE and ABSITE as measures of surgical intern performance. *J Surg Res*. 1994;57:613-8.
51. Holmboe ES, Fiebach NF, Galaty L, Huot S. The effectiveness of a focused educational intervention on resident evaluations from faculty: a randomized controlled trial. *J Gen Intern Med*. 2001;16:1-6.
52. Noel GL, Herbers JE, Caplow MP, Cooper GS, Pangaro LN, Harvey J. How well do internal medicine faculty members evaluate the clinical skills of residents? *Ann Intern Med*. 1992;117:757-65.
53. Herbers JE, Noel GL, Cooper GS, Harvey J, Pangaro LN, Weaver MJ. How accurate are faculty evaluations of clinical competence? *J Gen Intern Med*. 1989;4:202-8.
54. Elliot DL, Hickam DH. Evaluation of physical examination skills. Reliability of faculty observers and patient instructors. *JAMA*. 1987;340:5-8.
55. Kalet A, Earp JA, Kowlowitz V. How well do faculty evaluate the interviewing skills of medical students? *J Gen Intern Med*. 1992;9:179-84.
56. Barrows HS. An overview of the uses of standardized patients on teaching and evaluating clinical skills. *Acad Med*. 1993;68:443-53.
57. Van der Vleuten CPM, Swanson DB. Assessment of clinical skills with standardized patients: state of the art. *Teach Learn Med*. 1990;2:58-76.
58. Sloan DA, Donnelly MB, Schwartz RW, Strodel WE. The objective structured clinical examination: the new gold standard for evaluating postgraduate clinical performance. *Ann Surg*. 1995;222:735-42.
59. Anderson MB, Stillman PL, Wang Y. Growing use of standardized patients in teaching and evaluation in medical education. *Teach Learn Med*. 1994;6:15-22.
60. Richards BF, Rupp R, Zaccaro DJ, et al. Use of standardized patient based clinical performance examination as an outcome measure to evaluate medical school curricula. *Acad Med*. 1996;71:S49-S51.
61. Issenberg SB, McGaghie WC, Hart IR, et al. Simulation technology for health care professional skills training and assessment. *JAMA*. 1999;282:861-6.
62. Brailovsky CA, Grand'Maison P, Lescop J. Construct validity of the Quebec licensing examination SP-based OSCE. *Teach Learn Med*. 1997;9:44-50.
63. Grand'Maison P, Brailovsky CA, Lescop J, Rainsberry P. Using standardized patients in licensing/certification examinations: comparison of two tests in Canada. *Fam Med*. 1997;29:27-32.
64. Ram P, van der Vleuten C, Rethans JJ, Grol R, Aretz K. Assessment of family physicians: comparison of observation in a multiple-station examination using standardized patients with observation of consultations in daily practice. *Acad Med*. 1999;74:62-9.

65. Kopelow ML, Schnabl GK, Hassard TH, et al. Assessment of performance in the office setting with standardized patients: assessing practicing physicians in two settings using standardized patients. *Acad Med.* 1992;10:S19–S21.
66. Rethans JJ, Sturmans F, drop R, van der Vleuten C, Hobus P. Does competence of general practitioners predict their performance? Comparison between examination setting and actual practice. *BMJ.* 1991;303:1377–80.
67. Miller G. Invited reviews: the assessment of clinical skills/competence/performance. *Acad Med.* 1990;65(9 suppl):S63–S67.
68. Wass V, Van der Vleuten CP, Shatzer J, Jones R. Assessment of clinical competence. *Lancet.* 2001;357:945–9.
69. Holmboe ES, Williams FK, Yepes M, Norcini JJ, Blank LL, Huot S. The mini clinical evaluation exercise and interactive feedback: preliminary results. *J Gen Intern Med.* 2001;16(suppl 1):100.
70. Cohen JJ. Clinical skills under scrutiny (www.aamc.org/newsroom/reporter/dec02). Accessed 5 February 2003. *AAMC Reporter*, December 2002.
71. Wiest FC, Ferris TG, Gokhale M, Campbell EG, Weissman JS, Blumenthal D. Preparedness of internal medicine and family practice residents for treating common conditions. *JAMA.* 2002;288:2609–14.
72. Rich EC, Crowson TW, Harris IB. The diagnostic value of the medical history: perceptions of internal medicine physicians. *Arch Intern Med.* 1987;147:1957–60.
73. Ende J. Feedback in clinical medical education. *JAMA.* 1983;250:777–81.
74. Gil DH, Heins M, Jones PB. Perceptions of medical school faculty members and students on clinical clerkship feedback. *J Med Educ.* 1984;59:856–63.
75. Skeff KM. Evaluation of a method for improving the teaching performance of attending physicians. *Am J Med.* 1983;75:465–70.
76. Skeff KM, Stratos GA, Berman J, Bergen MR. Improving clinical teaching: evaluation of a national dissemination program. *Arch Intern Med.* 1992;152:1156–61.
77. Litzelman DK, Stratos GA, Marriott DJ, Skeff KM. Factorial validation of a widely disseminated framework for evaluating clinical teachers. *Acad Med.* 1998;73:688–95.
78. Holmboe ES, Huot SJ, Hawkins RE. Efficacy of direct observation of competence (DOC) training: a randomized controlled trial. *J Gen Intern Med.* 2003;18(1 suppl):244.
79. Levinson W, Rubenstein A. Integrating clinician-educators into academic medical centers: challenges and potential solutions. *Acad Med.* 2000;75:906–12.
80. Levinson W, Branch WTJr, Kroenke K. Clinician-educators in academic medical centers: a two-part challenge. *Ann Intern Med.* 1998;129:59–64.
81. Misch DA. Evaluating physicians' professionalism and humanism: the case for humanism connoisseurs. *Acad Med.* 2002;77:489–95.
82. Soumerai SB, McLaughlin TJ, Gurwitz JH, et al. Effect of local medical opinion leaders on quality of care for acute myocardial infarction: a randomized controlled trial. *JAMA.* 1998;279:1358–63.
83. Lomas J, Enkin M, Anderson GM, Hannah WJ, Vayda E, Singer J. Opinion leaders vs. audit and feedback to implement practice guidelines: delivery after a previous caesarian section. *JAMA.* 1991;265:2202–7.
84. Holmboe ES, Bradley ES, Mattera JA, Roumanis SA, Radford MJ, Krumholz HM. Characteristics of physician leaders working to improve the quality of care in acute myocardial infarction: a qualitative study. *Jt Comm J Qual Saf.* 2003;29:289–96.