

TWELVE TIPS

Twelve tips for implementing tools for direct observation of medical trainees' clinical skills during patient encounters

KAREN E. HAUER¹, ERIC S. HOLMBOE² & JENNIFER R. KOGAN³

¹University of California San Francisco, USA, ²American Board of Internal Medicine, USA, ³Penn Medicine, University of Pennsylvania School of Medicine, USA

Abstract

Background: Direct observation of medical trainees by their supervisors with actual patients is essential for trainees to develop clinical skills competence. Despite the many available tools for direct observation of trainees by supervisors, it is unclear how educators should identify an appropriate tool for a particular clinical setting and implement the tool to maximize educational benefits for trainees in a manner that is feasible for faculty.

Aims and methods: Based on our previous systematic review of the literature, we provide 12 tips for selecting and incorporating a tool for direct observation into a medical training program. We focus specifically on direct observation that occurs in clinical settings with actual patients.

Results: Educators should focus on the existing tools for direct observation that have evidence of validity. Tool implementation must be a component of an educational program that includes faculty development about rating performance, providing meaningful feedback, and developing action plans collaboratively with learners.

Conclusions: Educators can enhance clinical skills education with strategic incorporation of tools for direct observation into medical training programs. Identification of a psychometrically sound instrument and attention to faculty development and the feedback process are critical to the success of a program of direct observation.

Background

Direct observation of medical trainees with actual patients by supervising physicians constitutes a key component of clinical education and evaluation. Undergraduate and graduate medical trainees learn fundamental clinical skills including history-taking, physical examination, communication, and patient counseling skills in large part through the care of patients that is accompanied by feedback from more experienced physicians. The organizations that accredit training programs in the United States, including the Liaison Committee on Medical Education (LCME) and the Accreditation Council for Graduate Medical Education (ACGME) require supervisors to observe medical trainees with patients as part of the training program. Similarly, the United Kingdom created a series of work-based assessments for their new Foundation program (first 2 years post-medical school; PMETB 2008). Direct observation can enhance the level of supervision of trainees to help ensure that patients receive high-quality care (Kilminster et al. 2002).

Unfortunately, trainees are infrequently observed by their supervisors during clinical interactions with patients (Howley & Wilson 2004). Without adequate observation and feedback, trainees miss opportunities to learn fundamental clinical

skills and patients may not receive safe and effective care. Not surprisingly, then, studies have shown that practicing physicians commonly manifest inadequate skills in physical examination and communication (Mangione et al. 1995; Braddock et al. 1999; Vukanovic-Criley et al. 2006; Mitka 2008).

Direct observation as an educational tool

As medical trainees proceed through training, they ideally acquire the skills to progress from novice to expert. The process of skills acquisition in medicine or any field can occur through deliberate practice, in which the learner practices a skill repetitively and carefully, with feedback and opportunity to implement changes under the guidance of a more experienced individual (Ericsson et al. 1993). Without such intentional practice, clinical skills learning can be relegated to a random set of clinical encounters without attention to trainees' learning needs or progress. Therefore, clinical curricula should ideally incorporate intentional observation of trainees with patients followed by developmentally appropriate feedback on trainees' strengths and weaknesses. Feedback should

Correspondence: K. E. Hauer, Department of Medicine, University of California, San Francisco, 505 Parnassus Avenue, M1078, Box 0120, San Francisco, CA 94143-0120, USA. Tel: 1 415 476 1964; fax: 1 415 502 7544; email: karen.hauer@ucsf.edu

include the correction of errors and action planning for future clinical encounters (Hattie et al. 2007).

Direct observation as an assessment tool

Clinical training programs are expected to determine that their trainees are competent to move to the next level of training and ultimately to independent practice. In order to make a decision about a trainees' competence, the program must possess information about the trainees' performance in practice. The use of multiple assessment tools including tools for direct observation of trainees with patients allows for summative assessment of competence in multiple domains, particularly patient care and communication skills. While other methods of assessment such as end-of-rotation global evaluations are popular and convenient, they suffer from recall bias. Faculty who complete global evaluations may not have observed the clinical skills they are asked to rate, and they may thus infer competence from more easily observed activities such as verbal presentations (Burdick & Schoffstall 1995; Howley & Wilson 2004).

Barriers to direct observation

Implementing and sustaining a program of direct observation is challenging in the current clinical environment. Faculty face competing demands to meet clinical productivity expectations. Graduate medical trainees also have high clinical demands in the setting of work hour restrictions that compress their available time for service and education. Even with adequate time for observation, trainees are often reluctant to be observed and faculty may feel ill prepared to offer meaningful feedback.

On a programmatic level, educators interested in incorporating direct observation must identify an appropriate tool for their program, either one of the many existing tools or a new tool created locally. They need to create time for direct observations and feedback in crowded training schedules. Observers and trainees need guidance in how to conduct the observed encounters in a meaningful way that does not disrupt patient care. Program directors must then determine how to use the information that is collected to benefit individual trainees and, where appropriate, the curriculum overall. These 12 tips will guide educators in implementing tools for direct observation of trainees' clinical skills by their supervisors in actual patient encounters.

Tip 1

Define competencies and objectives for the program to guide use of a tool for direct observation

A tool for direct observation should be introduced as part of a program of clinical skills education and assessment. Regardless of the level of learner, directors of educational programs must define a framework of competencies and objectives that outlines the specific clinical and communication

skills expected of learners in the program. This framework can then guide the articulation of developmental benchmarks characterizing the expected level of performance at specific intervals during the training program. Behavioral objectives regarding the expectations for important clinical skills in practice with patients must be included. This information is essential for implementation of a tool of direct observation because it informs both observers and trainees about the behaviors to be observed and assessed. In addition, clear performance objectives can also guide selection of the most appropriate tool for the particular clinical setting.

Tip 2

Determine whether the purpose of the direct observation program is formative or summative assessment

Assessment of clinical skills can be done for formative or summative assessment. When a program for direct observation is designed for formative assessment, observations and ratings are used for feedback to trainees about areas in which they are performing well and areas requiring additional attention. Typically, these assessments do not count for grading or promotion. Direct observation tools implemented for summative assessment entail observations and ratings that are used in evaluating the trainee. The decision about whether direct observations will be used for formative or summative assessment will influence the type of tool selected and the number of times a trainee should be assessed. Whereas most tools have been developed and used for formative assessment, some have been designed and tested for summative use (Kogan et al. 2009). Some tools developed for summative assessment could be used formatively and *vice versa*, but the purpose of the observation must be clear to both the observers and the learners.

Tip 3

Identify an existing tool for direct observation rather than creating a new one

Given the large number of existing tools for direct observation, educators have a range of options for introducing direct observation into a training program. A recent systematic review of the literature on tools for direct observation of students, residents and fellows by supervisors identified 55 existing tools (Kogan et al. 2009). These tools have been implemented in inpatient and outpatient settings and in both medical and surgical specialties.

Existing tools share significant commonality of skill domains assessed. Most assess trainees' history-taking, physical examination, and communication or counseling skills, and many include multiple skill domains. Some tools focus on more specific clinical skills, such as palliative care or cardiac auscultation (Han et al. 2005; Torre et al. 2005). Most tools employ numerical rating scales supported by adjectives/adverb anchors. Given this commonality of competency

domains and rating scales, we recommend that strong consideration be given to using existing tools, with modifications as needed, rather than focusing on new tool development. Existing tools can be adapted using developmental benchmarks to reflect the skills that should be attained at different points across the educational curriculum. With this approach, feedback received and incorporated as a medical student could be built upon as an intern, resident, or fellow. For example, an intern might be observed conducting a primary care visit in the ambulatory setting and receive feedback about basic approaches to counseling for medication adherence, whereas a resident might be observed counseling about medication adherence with a challenging patient and receive feedback using the same form. Forms with developmental benchmarks and behavioral anchors in addition to numerical ratings can provide raters with a framework for their assessments that may promote greater rater accuracy and stringency and decrease rating biases.

Selection of an existing tool should be based, in part, on identifying a tool for which there is established validity. Validity encompasses the degree to which the tool actually measures the construct (in this case, clinical skills) being assessed (Messick 1995). Content validity refers to how and by whom the items on the tool were selected. A tool ideally will have demonstrated reliability of the items, which means that the measurements are consistent across different observers or the same observer at different time points. The construct validity of a tool may be demonstrated by comparing performance on the tool with performance on other assessments, such as end-of-month ward evaluations or clinical skills examinations with standardized patients. For high-stakes summative assessments, a well-studied tool that yields reliable and valid data is needed, whereas for low-stakes formative assessment, a tool may be useful for feedback despite lack of information about reliability and validity. Accuracy of ratings, defined as the observer correctly identifying skills performed well in addition to errors of omission or commission, is a necessary component of validity. Often, errors are missed by faculty (Noel et al. 1992), and accuracy of ratings of the trainee is influenced by multiple factors including the time lag between the observation and the recording of the ratings, the specific questions on the rating form, and rater training on the items being evaluated (Williams 2003).

Tip 4

Create a culture that values direct observation

Direct observation of trainees with patients occurs infrequently and may not be institutionally valued as necessary for the development and mastery of clinical skills. Attention to the institutional culture is crucial for the success of a program for direct observation of clinical skills. A system-wide change regarding the culture of supervision and assessment can be both challenging and intimidating. Existing models for transforming organizational beliefs, such as Kotter's (1995) Eight Step Model, have been successfully applied in educational settings through faculty development (Steinert et al.

2007). For example, emphasizing the need for direct observation with evidence documenting how poor clinical skills negatively impact quality of care, lead to diagnostic errors (Hasnain et al. 2001; Graber et al. 2005), reduce patient satisfaction, and decrease the likelihood that patients achieve goals for chronic conditions can promote buy-in (Simpson et al. 1991; Stewart 1995). Describing how feedback may minimize diagnostic errors and promote the development of expertise can be helpful (Graber 2008). Faculty educational "champions" should be identified to role model the importance of observation, teach others to do it well, and mentor other faculty in the process. Creating and then communicating a clear and concise vision of optimal clinical skills observation and assessment standardizes expectations. Finally, successful programs for direct observation can remove obstacles to the process and motivate faculty who regularly observe trainees with patients through awards, salary incentives, and dedicated time to teach (Steinert et al. 2007).

Tip 5

Conduct faculty development on direct observation

Ultimately, tools for direct observation are only as good as the individuals using them to assess learners. Faculty must be skilled at observation and accurate in their ratings. In contrast to the numerous tools developed, there is a paucity of information regarding best practices to train raters to use them. With few exceptions (Lane & Gottlieb 2000; McKinley et al. 2000; de Haes et al. 2001; Ross 2002; Holmboe et al. 2004a), most tools for direct observation have been implemented with minimal or no faculty preparation or training.

Observers need training in behavioral observation which encompasses how to prepare themselves, the learners and patients for direct observation. Although the relative efficacy of various faculty development approaches remains unclear, some strategies and best practices have been identified. Training about the actual observation can include teaching observers where to sit during observations and how to minimize interruptions during observation. Observers must also learn to rate learners' performance reliably and discriminate between performance levels (Shumway & Harden 2003). Effective faculty development requires orientation to the rating form and to learners' expected performance levels against predefined performance benchmarks. Faculty might discuss the rating form and performance expectations for particular clinical skills (Holmboe et al. 2004a). Using videotaped examples of learner performance, observers can develop evaluation standards for superior, satisfactory, marginal, and unsatisfactory performances to calibrate their ratings with other raters and then practice ratings with more example encounters. While faculty development programs can improve rater comfort in direct observation and increase rater stringency (Holmboe et al. 2004a), brief one-time interventions may be unsuccessful (Cook et al. 2009). Ongoing faculty development, using more longitudinal approaches, will probably be needed to help faculty acquire high levels of

observation skill and to recalibrate their ratings over time (Hemmer & Pangaro 2000).

Tip 6

Build meaningful feedback into the direct observation process and train faculty to provide effective feedback

To make the direct observation process useful to learners, faculty must be able to provide meaningful feedback. By observing learners with patients, assessing their performance, and providing feedback, faculty help learners to improve skills, and learn new patterns of performance (Duffy et al. 2004). Although faculty may be skilled at providing clinical care, they may have very little experience or expertise in guiding learners to develop the same skills. Therefore, faculty development should emphasize core principles of effective feedback including the importance of timely feedback (i.e., occurring soon after the observation of the trainee) that is delivered in a private and comfortable setting (Ende 1983; van de Ridder et al. 2008). Feedback should include rich narrative information. Learners value high-quality narrative feedback, both verbal and written, more highly than checklist-based feedback because it is individualized and focused on the learner's specific areas for improvement (Govaerts et al. 2005). Faculty development can emphasize the importance of sharing specific positive, reinforcing feedback that communicates what went well and how an effective encounter is constructed. Particular emphasis should be placed on giving clear, specific, behaviorally based constructive feedback that focuses on a behavior that the learner has the capacity to change. Faculty development can highlight the importance of trainee self-assessment, and observers can be taught how to ask trainees to identify skills performed well, gaps in skills, and other areas in which feedback is desired. This information provides the observer with information about the learner's insight about clinical skills development. Discrepancies between the observer's and the learner's understanding of the learner's performance can stimulate discussion and help in framing individualized learning plans.

Faculty participants should have the opportunity to role-play giving and receiving feedback. For instance, these role plays can occur with videotapes of encounters between patients and learners (Murphy et al. 2004; van de Ridder et al. 2008).

Tip 7

Require action planning after each direct observation

An action plan that characterizes steps the learner can take to improve is crucial to the effectiveness of feedback. After observation, the observer and learner should agree upon an action plan for the learner. Unfortunately, most tools for direct observation do not include space for open-ended comments or a specific space for an action plan, although one can be added after any observation (Kogan et al. 2009). An effective action plan is measurable, specific, and behavioral. The

learner should have the potential to accomplish the plan in a reasonable timeframe that correlates with the next scheduled observation and feedback.

Action plans must be followed with more observation and feedback over time to determine learning progress. If an observer does not have the opportunity to work with a particular trainee again, systems should be established to instill in the trainee the motivation to take responsibility for sharing the action plan or seeking feedback on the pertinent skills with future supervisors. Action planning should prompt reassessment in the context of developmental benchmarks within the program. A tool with a rating scale that can be calibrated to the level of learner can be used to reflect skills acquisition over time.

There are several potential barriers to action planning (Holmboe et al. 2004b). Faculty may not know how to construct specific plans with learners, or they may not want to invite the accountability for ensuring that the learner accomplishes the action plan. Similarly, learners may not engage with the faculty in meaningful discussion of action planning, or they may feel anxious, embarrassed, or unmotivated to incorporate suggestions for improvement. Structuring the feedback discussion and orienting the learners and faculty to the process can overcome these barriers.

Tip 8

Orient learners to direct observation and feedback

In addition to training faculty, learners also should be oriented to the process of direct observation and feedback. The learner role in the feedback process has received little consideration in studies describing the implementation of direct observation tools (Kogan et al. 2009). However, feedback is a dynamic, two-way process, and increased attention is being given to the learner role (Bing-You & Trowbridge 2009).

Learners should understand the importance of direct observation and feedback for their developing clinical skills expertise. The expectation that they will be observed with patients should be made explicit. Learner familiarity with the direct observation process seems to decrease the anxiety associated with observation (Malhotra et al. 2008). Consistent with adult learning theory, learners can be encouraged to identify skills they wish to improve to help focus direct observation. Some learners may not wish to expose perceived areas of weakness, but educators can suggest approaches for developing skills without unduly highlighting shortcomings. Additionally, learners can be trained to recognize, receive, solicit and respond to feedback (Bing-You & Trowbridge 2009). Learners should be taught to seek directed feedback using strategies that increase the likelihood the feedback they receive will be specific, addressing both strengths and areas requiring attention (Eva & Regehr 2008). Learners who have been educated about the importance of an action plan may be more proactive in eliciting one when it is not spontaneously offered. Furthermore, since learners ideally should self-assess prior to feedback, educators can teach and encourage reflective practice using questions to promote reflection about knowledge, skills, attitudes, feelings, biases, and assumptions

(Westberg 2001). Strategies to help learners understand their affective response to constructive feedback may also be useful.

Tip 9

Apply the tool multiple times per trainee

Many studies of tools for direct observation of trainees describe applying the tool with each learner only once or a few times (Kogan et al. 2009). This finding underscores the unreliable use of these assessment tools in medical training (van der Vleuten & Schuwirth 2005). Substantial sampling across contexts and assessors is essential to achieve score reliability with any instrument (van der Vleuten & Schuwirth 2005). For high-stakes summative evaluation, at least 10–12 assessments with a tool are needed; however, a minimum of 4 assessments may suffice if trying to determine minimal competence (Norcini et al. 2003). Notably, the miniCEX is the most well-studied tool, with multiple assessments per learner in multiple studies. A program of multiple observations must be feasible within the context of the overall training program and the clinical setting, and medical educators should consider how to ensure that tools are implemented frequently across clinical contexts to enhance reliability. Tools for brief structured observations offer promise as feasible mechanisms to facilitate frequent observation and feedback for formative assessment (Lane & Gottlieb 2000; Kuo et al. 2005). For example, during a brief structured observation, an observer watches a trainee with a patient for less than 5 min at any point during an encounter and offers one to three feedback points after the encounter concludes.

Tip 10

Develop systems that accommodate direct observation of clinical skills

A system is essential to accommodate a program for direct observation of trainees' clinical skills with actual patients in the clinical setting. Time for observation must be identified, recognizing that faculty time observing trainees may leave them unavailable for other revenue-generating or scholarly activities. Educators can creatively minimize a potential loss of time and productivity by embedding direct observation within their usual patient care responsibilities. For example, a faculty member might join a trainee for part of morning rounds to see a patient, rather than seeing that patient independently later in the day. In the clinic, faculty who have parallel schedules with trainees they are supervising can conduct some of their own patient visits and intermittently observe the trainee in the room with the patient as part of the precepting interaction. Buy-in from senior leadership may help secure funds to support observation time and faculty development.

Developing systems that facilitate the completion of direct observation forms and their subsequent tracking are important to the success of a plan for direct observation. While paper systems have been used, web and personal digital assistant (PDA) based assessments are alternatives that make forms accessible with easy methods to monitor completion rates

(Torre et al. 2007). Whether the forms are “housed” with the learner or the observer needs to be decided, considering the pros and cons to each approach. Learners may gain feelings of empowerment over their learning with the responsibility to hold the form and seek assessments of their skills. The administrative time needed for implementation and maintenance of the program, often not specifically cited in studies, should be determined so that it can be supported accordingly.

Tip 11

Measure outcomes of the direct observation of clinical skills program

As with any educational program, evaluating outcomes of the intervention and the effect on learners is essential. Trainee and observers' attitudes and satisfaction with the observation process can be measured, commonly using ratings on the direct observation tools themselves. Alternatively, evaluations of the process can be completed at a later date, or focus groups can be used to acquire such information. While assessment of satisfaction is essential, it should not be considered sufficient. Self-assessed improvement in knowledge or skills has been used to evaluate whether a tool for direct observation improves learner outcomes (Kogan et al. 2009). However, in light of the inaccuracy of self-assessment (Davis et al. 2006) more objective measurements of learning are needed. Demonstrating that trainees apply skills learned through direct observation in future patient care, and that those behaviors enhance patient care quality, are challenging but important goals.

Tip 12

If a new tool is developed for use, try to assess its validity

Although many tools exist for the direct observation of clinical skills, educators may want to develop a new tool for direct observation based on specific programmatic goals. In that event, educators should try to evaluate the construct validity of the new tool and disseminate those findings. Potentially useful information for dissemination would include a description of the tool and how it was implemented, and different aspects of construct validity (Messick 1995). Adding to what is known in the literature about the validity of tools for direct observation of clinical skills could help medical educators who should be selecting formative and summative assessment tools based, in part, on their potential to enhance the educational program (Downing 2003; Shumway & Harden 2003; van der Vleuten & Schuwirth 2005; Cook et al. 2009). Assessment of new tools may require collaboration with experts in psychometrics, evaluation and assessment.

Conclusion

Many tools have been developed to promote bedside observation and assessment of trainees' clinical skills. For maximal

impact, a tool should be implemented as a key curricular component of an educational program of clinical skills education and assessment. Educators may select from a variety of existing tools, focusing on tools that have established psychometric strength. Faculty development enriches the process of direct observation by focusing the observers on the skills to be observed and expected levels of performance. Observers and learners must also be trained about meaningful feedback, which should include specific behavioral observations about areas for reinforcement and improvement, learner self-assessment, and a collaboratively developed action plan with follow up. A sustainable program of direct observation requires a reconceptualization by faculty on how to conduct observations, dedicated time and clinical space for observations to occur and supported faculty time to conduct observations and feedback in clinical settings.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the article.

Notes on contributors

KAREN E. HAUER, MD is professor of Medicine and Gold Headed Cane Endowed Education Chair in the Department of Medicine at the University of California, San Francisco. She is director of Internal Medicine Clerkships and director of Student Assessment for the School of Medicine.

ERIC S. HOLMBOE, MD is senior vice president and chief medical officer at the American Board of Internal Medicine and American Board of Internal Medicine Foundation. He is also adjunct professor of Medicine at Yale University, and adjunct professor at the Uniformed Services University of the Health Sciences.

JENNIFER R. KOGAN, MD is assistant professor of Medicine and director of Undergraduate Education for the Department of Medicine at the University of Pennsylvania School of Medicine. She is clerkship director of the Internal Medicine Core Clerkship and course director of the Internal Medicine Sub-Internship.

References

- Bing-You RG, Trowbridge RL. 2009. Why medical educators may be failing at feedback. *JAMA* 302:1330–1331.
- Braddock III CH, Edwards KA, Hasenberg NM, Laidley TL, Levinson W. 1999. Informed decision making in outpatient practice: Time to get back to basics. *JAMA* 282:2313–2320.
- Burdick WP, Schoffstall J. 1995. Observation of emergency medicine residents at the bedside: How often does it happen? *Acad Emerg Med* 2:909–913.
- Cook DA, Dupras DM, Beckman TJ, Thomas KG, Pankratz VS. 2009. Effect of rater training on reliability and accuracy of mini-CEX scores: A randomized, controlled trial. *J Gen Intern Med* 24:74–79.
- Davis DA, Mazmanian PE, Fordis M, Van Harrison R, Thorpe KE, Perrier L. 2006. Accuracy of physician self-assessment compared with observed measures of competence: A systematic review. *JAMA* 296:1094–1102.
- de Haes JC, Oort F, Oosterveld P, Ten Cate O. 2001. Assessment of medical students' communicative behaviour and attitudes: Estimating the reliability of the use of the Amsterdam attitudes and communication scale through generalisability coefficients. *Patient Educ Couns* 45:35–42.
- Downing SM. 2003. Validity: On meaningful interpretation of assessment data. *Med Educ* 37:830–837.
- Duffy FD, Gordon GH, Whelan G, Cole-Kelly K, Frankel R, Buffone N, Lofton S, Wallace M, Goode L, Langdon L. 2004. Assessing competence in communication and interpersonal skills: The Kalamazoo II report. *Acad Med* 79:495–507.
- Ende J. 1983. Feedback in clinical medical education. *JAMA* 250:777–781.
- Ericsson K, Krampe RT, Tesch-Römer C. 1993. The role of deliberate practice in the acquisition of expert performance. *Psychol Rev* 100:363–406.
- Eva KW, Regehr G. 2008. "I'll never play professional football" and other fallacies of self-assessment. *J Contin Educ Health Prof* 28:14–19.
- Govaerts MJ, van der Vleuten CP, Schuwirth LW, Muijtjens AM. 2005. The use of observational diaries in in-training evaluation: Student perceptions. *Adv Health Sci Educ Theory Pract* 10:171–188.
- Graber ML. 2008. Taking steps towards a safer future: Measures to promote timely and accurate medical diagnosis. *Am J Med* 121:S43–S46.
- Graber M, Franklin N, Gordon R, Power D, Shanafelt T. 2005. Diagnostic error in internal medicine. *Arch Intern Med* 165:1493–1499.
- Han PK, Keranen LB, Lescisin DA, Arnold RM. 2005. The palliative care clinical evaluation exercise (CEX): An experience-based intervention for teaching end-of-life communication skills. *Acad Med* 80:669–676.
- Hasnain M, Bordage G, Connell K, Sinacore J. 2001. History-taking behaviors associated with diagnostic competence of clerks: An exploratory study. *Acad Med* 76:S14–S17.
- Hattie JA, Timperly H, Rogers HJ. 2007. The power of feedback. *Rev Res Educ* 77:81–112.
- Hemmer PA, Pangaro L. 2000. Using formal evaluation sessions for case-based faculty development during clinical clerkships. *Acad Med* 75:1216–1221.
- Holmboe ES, Hawkins RE, Huot SJ. 2004a. Effects of training in direct observation of medical residents' clinical competence: A randomized trial. *Ann Intern Med* 140:874–881.
- Holmboe ES, Yepes M, Williams F, Huot SJ. 2004b. Feedback and the mini clinical evaluation exercise. *J Gen Intern Med* 19:558–561.
- Howley LD, Wilson WG. 2004. Direct observation of students during clerkship rotations: A multiyear descriptive study. *Acad Med* 79:276–280.
- Kilminster S, Jolly B, van der Vleuten CP. 2002. A framework for effective training for supervisors. *Med Teach* 24:385–389.
- Kogan JR, Holmboe ES, Hauer KE. 2009. Tools for direct observation and assessment of clinical skills of medical trainees: A systematic review. *JAMA* 302:1316–1326.
- Kotter J. 1995. Leading change: Why transformation efforts fail. *Harv Bus Rev* 73:59–67.
- Kuo AK, Irby DI, Loeser H. 2005. Does direct observation improve medical students' clerkship experiences? *Med Educ* 39:518.
- Lane JL, Gottlieb RP. 2000. Structured clinical observations: A method to teach clinical skills with limited time and financial resources. *Pediatrics* 105:973–977.
- Malhotra S, Hatala R, Courneya CA. 2008. Internal medicine residents' perceptions of the Mini-Clinical Evaluation Exercise. *Med Teach* 30:414–419.
- Mangione S, Burdick WP, Peitzman SJ. 1995. Physical diagnosis skills of physicians in training: A focused assessment. *Acad Emerg Med* 2:622–629.
- Mckinley RK, Fraser RC, van der Vleuten C, Hastings AM. 2000. Formative assessment of the consultation performance of medical students in the setting of general practice using a modified version of the leicester assessment package. *Med Educ* 34:573–579.
- Messick S. 1995. Standards of validity and the validity of standards in performance assessment. *Educ Meas* 14:5–8.
- Mitka M. 2008. Many physician practices fall short on accurate blood pressure measurement. *JAMA* 299:2842–2844.
- Murphy KR, Cleveland JN, Skattebo AL, Kinney TB. 2004. Raters who pursue different goals give different ratings. *J Appl Psychol* 89:158–164.
- Noel GL, Herbers Jr JE, Caplow MP, Cooper GS, Pangaro LN, Harvey J. 1992. How well do internal medicine faculty members evaluate the clinical skills of residents? *Ann Intern Med* 117:757–765.
- Norcini JJ, Blank LL, Duffy FD, Fortna GS. 2003. The mini-CEX: A method for assessing clinical skills. *Ann Intern Med* 138:476–481.
- PMETB [Internet]. 2008. Postgraduate Medical Education Training Board. Standards for curriculum and assessment systems. United Kingdom. [Published 2009 November 20]. Available from: <http://www.pmetb.org>.

- uk/fileadmin/user/Standards_Requirements/PMETB_Scas_July2008_Final.pdf
- Ross R. 2002. A clinical-performance biopsy instrument. *Acad Med* 77:268.
- Shumway JM, Harden RM. 2003. AMEE Guide No. 25: The assessment of learning outcomes for the competent and reflective physician. *Med Teach* 25:569–584.
- Simpson M, Buckman R, Stewart M, Maguire P, Lipkin M, Novack D, Till J. 1991. Doctor-patient communication: The Toronto consensus statement. *BMJ* 303:1385–1387.
- Steinert Y, Cruess RL, Cruess SR, Boudreau JD, Fuks A. 2007. Faculty development as an instrument of change: A case study on teaching professionalism. *Acad Med* 82:1057–1064.
- Stewart MA. 1995. Effective physician-patient communication and health outcomes: A review. *CMAJ* 152:1423–1433.
- Torre DM, Sebastian JL, Simpson DE. 2005. A PDA-based instructional tool to monitor students' cardiac auscultation during a medicine clerkship. *Med Teach* 27:559–561.
- Torre DM, Simpson DE, Elnicki DM, Sebastian JL, Holmboe ES. 2007. Feasibility, reliability and user satisfaction with a PDA-based mini-CEX to evaluate the clinical skills of third-year medical students. *Teach Learn Med* 19(3):271–277.
- van de Ridder JM, Stokking KM, Mcgaghie WC, Ten Cate OT. 2008. What is feedback in clinical education? *Med Educ* 42:189–197.
- van der Vleuten CP, Schuwirth LW. 2005. Assessing professional competence: From methods to programmes. *Med Educ* 39:309–317.
- Vukanovic-Criley JM, Criley S, Warde CM, Boker JR, Guevara-Matheus L, Churchill WH, Nelson WP, Criley JM. 2006. Competency in cardiac examination skills in medical students, trainees, physicians, and faculty: A multicenter study. *Arch Intern Med* 166:610–616.
- Westberg J. 2001. Helping learners become reflective practitioners. *Educ Health (Abingdon)* 14:313–321.
- Williams RG, Klamen DA, McGaghie WC. 2003. Cognitive, social and environmental sources of bias in clinical performance ratings. *Teach Learn Med* 15:270–292.