



## Clinical supervisor evaluations during general surgery clerkships

Tzu-Chieh Yu, Benjamin R. L. Wheeler & Andrew G. Hill

To cite this article: Tzu-Chieh Yu, Benjamin R. L. Wheeler & Andrew G. Hill (2011) Clinical supervisor evaluations during general surgery clerkships, *Medical Teacher*, 33:9, e479-e484, DOI: [10.3109/0142159X.2011.590558](https://doi.org/10.3109/0142159X.2011.590558)

To link to this article: <https://doi.org/10.3109/0142159X.2011.590558>



Published online: 19 Aug 2011.



Submit your article to this journal [↗](#)



Article views: 945



View related articles [↗](#)



Citing articles: 1 View citing articles [↗](#)

## WEB PAPER

# Clinical supervisor evaluations during general surgery clerkships

TZU-CHIEH YU, BENJAMIN R. L. WHEELER &amp; ANDREW G. HILL

The University of Auckland, New Zealand

## Abstract

**Background:** Clerkship performance is commonly evaluated by consultant surgeons who seldom supervise medical students directly. In contrast, surgical residents and interns frequently supervise students and provide essential teaching but are not tasked with evaluating them.

**Aim:** To prospectively investigate and compare the accuracy of general surgery clerkship performance evaluations by clinical supervisors of differing seniorities.

**Method:** Between September 2008 and May 2010, clinical supervisors of varying seniorities independently evaluated 57 fourth-year medical students using a multi-dimensional performance evaluation tool. Total evaluation grades and subtotal grades for clinical ability were correlated to the results of a validated surgical objective structured clinical examination (OSCE).

**Results:** In this study, 85 clinical supervisors provided 427 student performance evaluations. Total evaluation grades awarded by consultant surgeons had weak correlation to student OSCE results ( $r=0.27$ ,  $p < 0.05$ ) and associated subtotal grades for clinical ability had no correlation. In comparison, the equivalent sets of grades awarded by residents and interns had moderate correlations to OSCE results ( $r=0.49$  and  $r=0.54$ ,  $p < 0.01$ ).

**Conclusions:** Validity of clinical supervisor evaluations during general surgery clerkships vary according to assessor seniority. Including performance evaluation grades by surgical residents and interns may enhance the overall validity of this common clerkship evaluation tool and improve its summative and formative assessment value.

## Introduction

An important goal of clinical clerkships is to provide student evaluations that are reliable, valid and those which deliver meaningful feedback (Griffith & Wilson 2008). Such evaluations also need to assess clinical competency in a summative manner and fulfill administrative requirements (accreditation and progression through the medical programme). During surgical clerkships, combinations of different student evaluations are often utilised and may include subjective performance evaluations and objective written, oral and directly observed clinical examinations.

A commonly used student clerkship performance evaluation is the clinical supervisor evaluation. It often makes up a significant proportion of the overall clerkship evaluation grade and its advantages include low administrative and implementation costs, flexibility and time efficiency (Neufeld & Norman 1985). In theory, these assessments can be considered a form of direct observational evaluation because grades are based on observed student performance during routine clinical interactions.

In reality, however, senior clinicians and faculty rarely have opportunities to observe medical students performing core clinical skills such as history taking and physical examination (Stillman et al. 1991; Kassebaum & Eaglen 1999; Howley & Wilson 2004). The situation is exacerbated during surgical clerkships when the teaching and supervising responsibilities

## Practice points

- Clinical supervisor evaluations during general surgery clerkships can lack validity because consultant surgeons rarely have the opportunity and time to directly observe medical students.
- In comparison, residents and interns frequently supervise students and are able to provide reliable and accurate clerkship performance evaluations.
- Inviting surgical residents and interns to provide clinical supervisor evaluations on clerkship performance may improve the overall validity of this form of student evaluation.

of consultant surgeons (attending surgeons) frequently conflict with operative and research responsibilities. The validity and reliability of clinical supervisor evaluations during clerkships remain an ongoing concern for programme directors, clerkship coordinators and medical students, who perceive these to be the least fair of all evaluations (Duffield & Spencer 2002). Various cognitive, social and environmental factors have been described as contributing to unwanted sources of evaluation score variation (bias) including differing levels of leniency/stringency amongst raters, inflation of performance ratings, rating range restrictions, the central tendency phenomenon and the halo effect where the rater of a multi-dimensional

*Correspondence:* T.-C. Yu, South Auckland Clinical School, Middlemore Hospital, The University of Auckland, Private Bag 93 3111, Otahuhu, Auckland, New Zealand. Tel: 64 9 276 0076; fax: 64 9 276 0066; email: wendy.yu@auckland.ac.nz

performance evaluation gives similar ratings across all dimensions rather than distinguish amongst single dimensions (Risucci et al. 1992; Williams et al. 2003).

In contrast to consultants, surgical residents and interns can spend significant amounts of time teaching and directly supervising medical students in a wide variety of different clinical settings (Lowry 1976; Pelletier & Belliveau 1999; Minor & Poenaru 2002; Whittaker et al. 2006). Residents and interns are also closer to medical students in terms of age and professional development, and generally feel a strong responsibility towards teaching and supervising students and junior colleagues (Wilkerson et al. 1986; Busari et al. 2002).

Although residents and interns in New Zealand are not routinely asked nor trained to provide performance evaluations for medical students, we hypothesised that they may have an advantage over consultant surgeons if asked to

evaluate the clinical performance of medical students during general surgery clerkships because of the amount of time they spend interacting with students. A prospective comparison study was conducted to investigate whether the seniority of assessing clinical supervisors had an impact on the validity and reliability of clinical supervisor evaluation grades during general surgery clerkships.

## Method

During the 6-year undergraduate medical programme at The University of Auckland, the Year 4 general surgery clerkship is 6 weeks in duration and medical students are distributed amongst four University-affiliated teaching hospitals. At each hospital, students are assigned to different general surgical teams for each 3-week half of the clerkship and encouraged to actively participate in all team activities including ward rounds, outpatient clinics, acute and elective surgical procedures and review of acute patients in the emergency department. Surgical teams have a standard structure: two or three consultant surgeons, two residents and two interns.

As well as objective evaluations such as written assignments and case histories, two sets of clinical supervisor evaluation grades make up 40% of each student's overall general surgery clerkship grade. Students are responsible for nominating two consultant surgeons from differing surgical teams to provide them with these performance evaluations at the completion of the clerkship.

The clinical supervisor evaluation is a multi-dimensional student evaluation tool consisting of 10 items describing core clerkship learning expectations grouped into three domains: (1) acquisition and application of medical knowledge; (2) professional, clinical and research skills; and (3) population health and primary health care (Table 1). Items are graded using a four-level Likert scale: 4 = 'Excellent', 3 = 'Satisfactory', 2 = 'Some Reservations' and 1 = 'Major Deficiencies', resulting in a total grade out of 40 marks. A standardised definition for each grade is included in the clinical supervisor evaluation form (Table 2). Evaluation items left blank by assessors are routinely filled with the grade '3', a basic imputation method adopted by the Department of Surgery.

**Table 1.** University of Auckland fourth-year general surgery clerkship learning objectives outlined in the clinical supervisor evaluation form.

<i>Domain 1 – acquisition and application of medical knowledge</i>
Clinical knowledge (e.g. knowledge of common symptoms, understanding of disease and approach to disease/patient management including therapeutics)
<i>Domain 2 – professional, clinical and research skills</i>
Clinical patient assessment (ability to take history, perform basic clinical examination)
Clinical patient assessment (generate problem list and differential diagnosis)
Relevant procedural skills (e.g. venesection/suturing/arterial blood gas)
Communication skills <sup>a</sup> (patients/families/healthcare team)
Reliability and dependability <sup>a</sup> (punctuality, carrying out of instructions, time management)
Initiative and enthusiasm <sup>a</sup> (gets involved, able to identify needs of the job, follows up without being prompted, thinks and plans ahead, shows commitment, asks questions of supervisors)
Responsibility for own learning <sup>a</sup> (evidence of reading up on cases, attending teaching sessions, curiosity)
<i>Domain 3 – population health and primary health care</i>
Health promotion (knowledge of risk factors for illness and preventative strategies)
Location of patient illness in continuum of care (ability to plan future care and patient transition between primary, secondary and tertiary care)

Note: <sup>a</sup>Learning objectives related to student 'professionalism'.

**Table 2.** University of Auckland fourth-year general surgery clerkship clinical supervisor evaluation form: Grades and definitions.

Scored grade	Acquisition and application of medical knowledge	Professional, clinical and research skills	Population health and primary health care
Excellent (4)	Outstanding underpinning knowledge which is applied to patients appropriately	Excellent professional, clinical, communication and procedural skills. Able to function well as a member of the team	Comprehensive understanding of (1) health promotion and disease prevention and (2) of the whole patient and issues surrounding their care
Satisfactory (3)	Good working knowledge which is usually well applied to patient care	Good standard of professional, clinical and communication skills. Works as a team member. Able to recognise gaps and correct them	Good awareness of health promotion and disease prevention. Understands the whole patient and their place in community/society
Some Reservations (2)	Several gaps in underpinning knowledge	Some deficiencies in professional, clinical and communication skills	Tends to see the patient in isolation. Limited awareness of health promotion and disease prevention
Major Deficiencies (1)	Major deficiencies in knowledge base or in the ability to apply knowledge	Major deficiencies in professional and clinical skills. Poor communication, inappropriate attitudes and limited insight	Rarely if ever considers the patient in the wider context. Poor awareness of health promotion and disease prevention

For this study, the clinical performance of fourth-year medical students were independently assessed by supervising consultants and by team residents and interns at Middlemore Hospital during the course of five General surgery clerkships in separate 3-month periods in 2008, 2009 and 2010. While consultants provided each medical student with clinical supervisor evaluation grades in the conventional manner, participating residents and interns were individually approached by study investigators and asked to complete the same clinical supervisor evaluation form, providing grades for students under their supervision during the 6-week clerkships.

As part of the study, investigators also divided the 10 learning objectives of the clinical supervisor evaluation form into two groups: six describing clinical knowledge and skills and four describing professional conduct. Each completed evaluation, therefore, provided three figures: a total grade out of 40 marks and two subtotal grades, one representing the student's 'clinical ability' and the other their 'professionalism'. These were out of possible 24 and 16 marks, respectively.

Clinical supervisor evaluations were grouped according to the assessor's seniority and initial statistical analysis involved confirming data normality and examining the internal consistency of this clerkship evaluation instrument by calculating Cronbach-alpha coefficients. Total evaluation grades and subtotal grades for 'clinical ability' and 'professionalism' were then compared on the basis of assessor seniority using one-way analysis of variance. Finally, to test for criterion validity, evaluation grades for individual students were grouped according to assessor seniority and the mean of these grades was correlated to results from a centralised objective structured clinical examination (OSCE). This analysis was achieved using Pearson correlation tests.

The clerkship OSCE was chosen as the most fitting measure of each student's grasp of core surgical knowledge and skills because it has objectivity and demonstrable external validity and internal reliability (Yu et al. 2011). It consists of 11 written stations and 4 clinical skills stations which require students to demonstrate surgical history taking and physical examination skills under direct observation by University faculty academics. Each written station entails four or five short-answer questions centred on a clinical scenario and tests surgical knowledge and clinical application. Interpretation of laboratory and radiology investigations is also included in the written stations. While clinical skills make up 120 marks, written questions make up the remaining 110 marks to give a total OSCE result of 230 marks.

All statistical analyses were conducted using SPSS Version 13.0 for Windows (SPSS Inc., Chicago, IL, USA) and statistical significance was defined by a *p*-value < 0.05. The study's ethical approval was granted by The University of Auckland Human Participants Ethics Committee.

## Results

A total of 85 clinical supervisors participated in this study, providing 427 clinical supervisor evaluations for 57 fourth-year medical students. The 15 consultant surgeons provided 113 performance evaluations, 40 residents 152 evaluations and 30 interns the remaining 162 evaluations. Consultant participation

rate was 100% while the participation rates for registrars and house officers were 93% and 86%, respectively. Unavailability as a result of coinciding annual leave or after-hour clinical duties during study data collection were main reasons for non-participation by residents and interns. Consultants provided a median of two clinical supervisor evaluations for each student, residents a median of two and interns a median of three evaluations.

Consultant surgeons, as a group, did not submit any performance evaluations which had missing grades. Residents and interns submitted a total of 50 (11.7%) incomplete evaluation forms missing one or two individual grades. Of these, residents contributed 18 and interns contributed 32. Out of the evaluations, 31 of the evaluations were missing one grade and the remaining 19 were missing two grades. 'Procedural Skills' was the evaluation item most commonly left blank.

The clinical supervisor evaluation instrument demonstrated satisfactory internal consistency when completed by assessors from all three differing seniorities (Table 3). Consultants, as a group, awarded students significantly higher total evaluation grades compared to interns (Table 4). Consultants also awarded significantly higher subtotal grades for clinical ability compared to both residents and interns. There were no differences between awarded subtotal grades for professionalism amongst the clinical supervisors.

Table 5 summarises results of the correlation analyses between clinical supervisor evaluation grades and individual student OSCE grades. Mean of total evaluation grades awarded by consultant surgeons had a weak correlation with the individual student's OSCE result, while mean of subtotal grades for clinical ability had no correlation with OSCE results.

**Table 3.** Clinical supervisor evaluation grades: internal consistency.

	Cronbach-alpha coefficients		
	Clinical ability	Professionalism	Total
Consultants	0.87	0.85	0.90
Residents	0.85	0.83	0.88
Interns	0.78	0.80	0.86

**Table 4.** Clinical supervisor evaluation grades: Means and confidence intervals.

	N	Mean grades (95% confidence intervals)		
		Clinical ability	Professionalism	Total
Consultants	113	19.5* (19.1–19.9)	14.3 (13.9–14.6)	33.8‡ (33.1–34.5)
Residents	152	18.8† (18.5–19.1)	14.0 (13.7 – 14.3)	32.8 (32.3–33.4)
Interns	162	18.7† (18.4–19.0)	13.8 (13.5–14.2)	32.5# (32.0–33.1)

Note: \*Significantly different to †; ‡significantly different to # (*p* < 0.05).

**Table 5.** Clinical supervisor evaluations: Correlation of mean grades to student OSCE results.

Clinical supervisor	Evaluation component	Pearson correlation coefficient
Consultants	Clinical ability	0.26
	Total	0.27*
Residents	Clinical ability	0.54†
	Total	0.49†
Interns	Clinical ability	0.54†
	Total	0.49†

Note: \* $p < 0.05$ ; † $p < 0.01$ .

In contrast, means for total evaluation grades awarded by both residents and interns had moderate correlations to individual student OSCE results. Means of subtotal grades for individual clinical ability awarded by both the residents and interns also had moderate correlations with student OSCE results.

## Discussion

The results of this prospective comparison study support our concerns regarding the poor validity of clinical supervisor evaluations completed by consultant surgeons during Year 4 general surgery clerkships. They also demonstrate that surgical residents and interns are capable of providing meaningful performance evaluations for medical students during brief 3-week clerkship attachments. They awarded grades that moderately correlated with the scores from a validated and objective student evaluation of core surgical knowledge and skills.

Few studies have actually demonstrated the reliability and validity of clerkship supervisor evaluations (Rabinowitz & Hojat 1989; Campos-Outcalt et al. 1994, 1999) even though these performance evaluations are used by medical schools to rank students, assist students in appreciating their strength and weaknesses and document achievement of learning outcomes. Problems arise from the subjective nature of these evaluation grades, permitting bias from individual assessor idiosyncrasies regarding clinical focus (competencies they observe) and standards (stringency and leniency) and from insufficient direct observation and supervision by the evaluating senior clinicians which cannot be compensated for by assessor training nor instrument standardisation (Landon et al. 2003; Williams et al. 2003; Hasnain et al. 2004).

Understandably, direct observation of clinical skills by supervising clinicians during an undergraduate clerkship provides an authentic patient-centred teaching environment and improves student history-taking and physical examination skills (Reichsman et al. 1964; Cooper et al. 1983). Direct observation of students by faculty and assessing senior clinicians during clerkships is also critical for accurate evaluation of clinical skills (Hasnain et al. 2004; Holmboe 2004). The reality that clerkship performance ratings are rarely based on direct observations has previously been voiced by the concerned authors (Stillman et al. 1991; Kassebaum & Eaglen 1999). And, although the consultant's lack of opportunity and time to directly observe and supervise students during

undergraduate clerkships is likely an important rationalising factor explaining the findings of this study, there are also other potential contributing factors.

One possible factor is the intrinsic perceptions towards, and expectations of, medical students that vary amongst assessors of different seniorities. De et al. (2004) found that surgical consultants are more likely to base medical student grades on the questions posed by consultants to students, questions asked by students, student performance in the outpatient clinic and quality of the patient presentations. In contrast, surgical residents are more likely to place emphasis on each student's knowledge of ward patients and less likely to be concerned by the student's ability to answer questions posed to them. Stillman (1984) also found a parallel phenomenon when he compared commentary evaluations by attending surgeons and chief residents. Residents are more likely to mention the words 'skills' and 'technique' when providing student performance evaluations and less likely to comment on 'logic', 'judgement' or 'reasoning', words more frequently used by attending surgeons.

Another important potential contributing factor influencing clerkship evaluations is the physical setting in which student supervision takes place. The specific student attitude and skill requirements for different clinical environments might result in different performance evaluation ratings for the same student. During general surgery clerkships, consultant surgeons typically interact with students in the operating theatre or in ambulatory outpatient clinics. In these settings, the well-read student with outstanding interpersonal skills is able to impress consultants with his or her knowledge of anatomy and disease processes.

The same medical student may not, however, be as exceptional in the 'real world' environment of surgical wards or emergency departments where he or she is supervised by surgical residents. In these settings, the resident is also more likely to be assessing basic bedside clinical skills such as history-taking or physical examination. A slow and methodical student who takes thorough patient histories, performs systematic physical examinations and completes tasks and procedures is more likely to be appreciated by the precepting resident or intern. The fact that the study OSCE tests both surgical knowledge (using written short-answer questions) and basic bedside clinical skills (via directly observed demonstrations) should diminish some effects from this influencing factor.

Finally, it is important to appreciate that the circumstances surrounding selection of consultant evaluators were different to that of resident and intern evaluators. Consultants were singled out and nominated by medical students, while the residents and interns were selected by study investigators. Medical students would most likely have selected only those consultants with whom they had built up rapport with, who likely thought highly of them and who they have strived to impress. On the other hand, less effort was almost certainly invested into impressing residents and interns, given that students would not have anticipated evaluations from these junior supervisors. This may partially explain for the significantly higher grades awarded by consultant supervisors.

This study is not the first to investigate associations between assessor seniority and student performance evaluation validity during undergraduate surgical clerkships. Several previous studies have produced results which are conflicting. Carline et al. (1986) correlated the clerkship performance evaluation grades of 163 medical students with individual scores from the surgery component of their written National Board of Medical Examiners (NBME) examination and found that grades awarded by residents consistently correlated while those awarded by consultants did not. Stillman (1984), alternatively, correlated clinical evaluations of 105 students given by faculty members and by a chief resident with the results of a multi-choice written examination and two oral examinations. He found that the chief resident's evaluations were significantly lower in validity compared to those by faculty members. More recently, Awad et al. (2002) retrospectively reviewed the evaluation grades of 354 medical students and found that both consultant surgeons and residents produced subjective performance evaluations which were poorly correlated with results from objective written and oral examinations.

An important limitation to these previous studies is the use of written or oral examinations as comparable evaluations to clerkship performance evaluations. While written and oral examinations objectively and accurately assess clinical knowledge, they are inadequate measures of clinical skills such as history taking and physical examination. Furthermore, many written examinations, such as the NBME, are norm-referenced evaluations. In contrast, clerkship performance evaluations are commonly criterion-referenced. In this study, the investigators therefore, chose a criterion-referenced surgical OSCE as the benchmark measure of basic surgical knowledge and skills.

There are limitations to this study. First, with only a small sample of participating surgical consultants from one single institution, the results are limited in their generalisability. Second, formal imputation methods were not used to address the issue of missing data on incomplete clinical supervisor evaluation forms and blanks were not substituted with best-guess values. Knowledge on how assessing responders and non-responders differed would have allowed an appropriate grade imputation method to be employed (Jones 1996). In addition, the study does not provide insight into the weighting of each causal factor contributing to the findings in this study. An important confounding factor to control for in future investigations is the amount of available direct supervision time amongst clinical supervisors of different seniorities.

Finally, bias was likely introduced in this study by the fact that student proficiency in surgical procedural skills was evaluated as part of the clinical supervisor evaluation when in fact it is not a component in the clerkship OSCE. Furthermore, procedural skills was the evaluation item most commonly left blank by supervising residents and interns who were asked to evaluate medical students. This suggests a lack of confidence in providing students with a grade for this aspect of clerkship learning and the exact reason for this is beyond the scope of this study but indicates an area for future investigation.

Despite these methodology weaknesses, the results of this study have led to changes in the way students are evaluated during Year 4 general surgery clerkships at Middlemore

Hospital. Instead of clinical supervisor evaluations being completed only by consultant surgeons, students are now also able to approach residents for an evaluation of their clerkship performance. One of the two clinical supervisor evaluation forms can now be completed by a resident from either of the two surgical teams a student is attached to during his or her 6-week clerkship.

In conclusion, clinical supervisor evaluation grades provided by surgical residents and interns had significant correlations with results of a validated OSCE. Clinical supervisor evaluations by consultant surgeons, however, lacked notable correlations with results from the same clinical examination. The inclusion of clinical supervisor evaluation grades by surgical residents and interns is likely to enhance overall validity of these common clerkship evaluations and improve their summative and formative assessment value.

## Acknowledgements

The authors acknowledge and thank the study participants from Middlemore Hospital, Counties Manukau District Health Board, New Zealand.

**Declaration of interest:** The authors have no conflicts of interest to declare in regards to this research study including its subject matter, methodologies and funding and to the preparation of this manuscript.

## Notes on contributors

TZU-CHIEH YU, MBChB, is a research fellow in the Department of Surgery, University of Auckland.

BENJAMIN R. L. WHEELER, MBChB, is a General Surgery trainee of the Royal Australasian College of Surgeons.

ANDREW G. HILL, MBChB, MD, FRACS, FACS, is the head of the South Auckland Clinical School and an associate professor in the Department of Surgery, University of Auckland.

## References

- Awad SS, Liscum KR, Aoki N, Awad SH, Berger DH. 2002. Does the subjective evaluation of medical student surgical knowledge correlate with written and oral exam performance? *J Surg Res* 104:36–39.
- Busari JO, Prince KJ, Scherpier AJ, Van der Vleuten CP, Essed GG. 2002. How residents perceive their teaching role in the clinical setting: A qualitative study. *Med Teach* 24:57–61.
- Campos-Outcalt D, Watkins A, Fulginiti J, Kutob R, Gordon P. 1999. Correlations of family medicine clerkship evaluations and Objective Structured Clinical Examination scores and residency directors' ratings [Erratum appears in *Fam Med* 31: 308]. *Fam Med* 31:90–94.
- Campos-Outcalt D, Witzke DB, Fulginiti JV. 1994. Correlations of family medicine clerkship evaluations with scores on standard measures of academic achievement. *Fam Med* 26:85–88.
- Carline JD, Cook CE, Lennard ES, Siever M, Coluccio GM, Norman NL. 1986. Resident and faculty differences in student evaluations: Implications for changes in a clerkship grading system. *Surgery* 100:89–94.
- Cooper D, Beswick W, Whelan G. 1983. Intensive bedside teaching of physical examination to medical undergraduates: Evaluation including the effect of group size. *Med Educ* 17:311–315.
- De SK, Henke PK, Ailawadi G, Dimick JB, Colletti LM. 2004. Attending, house officer, and medical student perceptions about teaching in the third-year medical school general surgery clerkship. *J Am Coll Surg* 199:932–942.

- Duffield KE, Spencer JA. 2002. A survey of medical students' views about the purposes and fairness of assessment. *Med Educ* 36:879–886.
- Griffith 3rd CH, Wilson JF. 2008. The association of student examination performance with faculty and resident ratings using a modified RIME process. *J Gen Intern Med* 23:1020–1023.
- Hasnain M, Connell KJ, Downing SM, Olthoff A, Yudkowsky R. 2004. Toward meaningful evaluation of clinical competence: The role of direct observation in clerkship ratings. *Acad Med* 79:S21–S24.
- Holmboe ES. 2004. Faculty and the observation of trainees' clinical skills: Problems and opportunities. *Acad Med* 79:16–22.
- Howley LD, Wilson WG. 2004. Direct observation of students during clerkship rotations: A multiyear descriptive study. *Acad Med* 79:276–280.
- Jones J. 1996. The effects of non-response on statistical inference. *J Health Soc Pol* 8:49–62.
- Kassebaum DG, Eaglen RH. 1999. Shortcomings in the evaluation of students' clinical skills and behaviors in medical school. *Acad Med* 74:842–849.
- Landon BE, Normand SL, Blumenthal D, Daley J. 2003. Physician clinical performance assessment: Prospects and barriers. *JAMA* 290:1183–1189.
- Lowry SF. 1976. The role of house staff in undergraduate surgical education. *Surgery* 80:624–628.
- Minor S, Poenaru D. 2002. The in-house education of clinical clerks in surgery and the role of housestaff. *Am J Surg* 184:471–475.
- Neufeld VR, Norman GR. 1985. *Assessing Clinical Competence*. Springer Series on Medical Education, Vol. 7. New York, NY: Springer.
- Pelletier M, Belliveau P. 1999. Role of surgical residents in undergraduate surgical education. *Can J Surg* 42:451–456.
- Rabinowitz HK, Hojat M. 1989. A comparison of the modified essay question and multiple choice question formats: Their relationship to clinical performance. *Fam Med* 21:364–367.
- Reichsman F, Browning FE, Hinshaw JR. 1964. Observations of undergraduate clinical teaching in action. *J Med Educ* 39:147–163.
- Risucci DA, Lutsky L, Rosati RJ, Tortolani AJ. 1992. Reliability and accuracy of resident evaluations of surgical faculty. *Eval Health Prof* 15:313–324.
- Stillman RM. 1984. Pitfalls in evaluating the surgical student. *Surgery* 96:92–96.
- Stillman P, Swanson D, Regan MB, Philbin MM, Nelson V, Ebert T, Ley B, Parrino T, Shorey J, Stillman A, et al. 1991. Assessment of clinical skills of residents utilizing standardized patients. A follow-up study and recommendations for application. *Ann Intern Med* 114:393–401.
- Whittaker Jr LD, Estes NC, Ash J, Meyer LE. 2006. The value of resident teaching to improve student perceptions of surgery clerkships and surgical career choices. *Am J Surg* 191:320–324.
- Wilkerson L, Lesky L, Medio FJ. 1986. The resident as teacher during work rounds. *J Med Educ* 61:823–829.
- Williams RG, Klamen DA, McGaghie WC. 2003. Cognitive, social and environmental sources of bias in clinical performance ratings. *Teach Learn Med* 15:270–292.
- Yu TCW, Wheeler BR, Hill AG. 2011. Effectiveness of standardised clerkship teaching across multiple sites. *J Surg Res* 168:e17–e23.