INNOVATIONS IN EDUCATION AND CLINICAL PRACTICE

Evaluating the Performance of Inpatient Attending Physicians

A New Instrument for Today's Teaching Hospitals

Christopher A. Smith, MD, Anita B. Varkey, MD, Arthur T. Evans, MD, MPH, Brendan M. Reilly, MD

OBJECTIVE: Instruments available to evaluate attending physicians fail to address their diverse roles and responsibilities in current inpatient practice. We developed a new instrument to evaluate attending physicians on medical inpatient services and tested its reliability and validity.

DESIGN: Analysis of 731 evaluations of 99 attending physicians over a 1-year period.

SETTING: Internal medicine residency program at a university-affiliated public teaching hospital.

PARTICIPANTS: All medical residents ($N = 145$) and internal medicine attending physicians ($N = 99$) on inpatient ward rotations for the study period.

MEASUREMENTS: A 32-item questionnaire assessed attending physician performance in 9 domains: evidence-based medicine, bedside teaching, clinical reasoning, patient-based teaching, teaching sessions, patient care, rounding, professionalism, and feedback. A summary score was calculated by averaging scores on all items.

RESULTS: Eighty-five percent of eligible evaluations were completed and analyzed. Internal consistency among items in the summary score was 0.95 (Cronbach's α). Interrater reliability, using an average of 8 evaluations, was 0.87. The instrument discriminated among attending physicians with statistically significant differences on mean summary score and all 9 domain-specific mean scores (all comparisons, $P < .001$). The summary score predicted winners of faculty teaching awards (odds ratio [OR], 17; 95% confidence interval [CI], 8 to 36) and was strongly correlated with residents' desire to work with the attending again ($r = .79$; 95% CI, 0.74 to 0.83). The single item that best predicted the summary score was how frequently the physician made explicit his or her clinical reasoning in making medical decisions ($r^2 = .90$).

CONCLUSION: The new instrument provides a reliable and valid method to evaluate the performance of inpatient teaching attending physicians.

KEY WORDS: evaluating; teaching; attending physicians.

J GEN INTERN MED 2004;19:766-771.

Profound changes in the practice of medicine and delivery of health care have affected the roles and responsibilities of all physicians, including trainees and their teachers.¹ At most academic hospitals today, attending physicians' responsibilities extend far beyond the teaching of clinical medicine to residents and students. They must also manage patients' day-to-day care efficiently (sicker and quicker); incorporate the best available evidence into their clinical decisions²; exemplify professionalism³; and demonstrate mastery of bedside skills⁴ as well as new technology and treatments. Instruments currently used to evaluate the quality of attending physicians in teaching hospitals do not address these diverse responsibilities. Many measure didactic and professional behaviors more relevant to traditional classroom teaching than modern clinical attending.⁵⁻⁹ Others have not been well validated¹⁰ or provide only a single summary score to rank faculty performance,¹¹ which does not allow specific feedback to foster faculty improvement.¹²⁻¹⁴ For these reasons, no current instrument is valid and valuable, at least at our teaching hospital. Therefore, we developed a new instrument to evaluate inpatient attending physicians and tested its validity and reliability.

METHODS

We developed the instrument (available at www.jgim.org) with the goals of accurately representing the full range of roles and responsibilities of inpatient attending physicians and collecting information that was sufficiently specific to allow targeted efforts at faculty improvement. This project was judged to be exempt from Institutional Review Board approval at our facility.

Setting

Our setting is a university-affiliated public hospital with 145 internal medicine residents and 130 full-time internal medicine clinical faculty. Twelve admitting teams staff the hospital's internal medicine service. Teams consist of 4 residents and 1 attending physician who serves as both the teaching attending and the patients' physician of record. Each team admits approximately 100 patients during a 4-week period. The internal medicine service accepts all medical admissions to the hospital except patients admitted to the coronary care unit, the medical intensive care unit, or

Received from the Department of Medicine (CAS, ABV, ATE, BMR), Cook County Hospital and Rush Medical College, Chicago, Ill.

Address correspondence and requests for reprints to Dr. Smith: Department of Medicine, Cook County Hospital, Room 1518, 1900 W. Polk Street, Chicago, IL 60612 (e-mail: christopher_smith@rush.edu).

the subspecialty HIV service. Most faculty serve as attending physicians on the medicine inpatient service for two 4-week periods annually, supervising a different group of 4 residents each period. We evaluated the instrument during the academic year from July 2000 to June 2001.

Instrument Development

A 7-member faculty committee, including the department chair, residency program director, and associate program directors reviewed the literature on existing instruments,⁵⁻¹¹ clarified the important attending roles, responsibilities, and expectations within our hospital, and developed questions to evaluate their performance. Because our attending physicians simultaneously serve as the patients' primary inpatient physician and the residents' primary teacher, we included items that assessed, *a priori*, 4 main areas of interest to us—clinical supervision, teaching, feedback, and professionalism. The instrument was revised after review by other faculty and after pre-testing with a sample of residents. The final instrument included 32 items, with most items using 5-point Likert response options (Appendix).

Data Collection

Each resident independently completed an anonymous, scannable questionnaire on the last day of each 4-week rotation during a mandatory 30-minute feedback session. The questionnaire took about 10 minutes to complete, and during the meeting investigators collected the instruments and conducted group interviews with each team of residents to assess faculty performance and

resident experiences that might not have been captured by the questionnaire. Thus, these mandatory sessions offered an opportunity to assess the instrument's clarity and completeness. To maximize candor, we assured resident anonymity by transcribing all handwritten comments and by only presenting attending physicians with their own aggregated results at the end of the year. The confidentiality of faculty was protected by using unique numerical identifiers during data analysis and by not sharing results until their validity and reliability were verified.

Data Analysis

Our analytic goals were to group items into meaningful conceptual domains and then test the instrument's reliability and validity.

Conceptual Domains. After data collection was completed, we evaluated whether there were distinct conceptual domains of performance within the 4 broad areas of clinical supervision, teaching, feedback, and professionalism. The criteria for defining a conceptual domain were that questions had to be conceptually similar as well as statistically related (inter-item correlations greater than 0.4). The conceptual criteria required consensus among the 7 members of the instrument committee. Items that might have been conceptually similar but did not meet our statistical threshold, such as teaching sessions and teaching pertinent to patients, were retained as single-item domains. Two items were dropped because the residents considered them ambiguous. This process allowed us to group 32 items into 9 distinct conceptual domains (Table 1). A summary score was calculated as a mean of all equally weighted items in the 9 domains.

Table 1. Domains Representing the Roles and Responsibilities of Inpatient Attending Physicians

Teaching	Supervision
1. Evidence-based medicine Broad knowledge of medicine (q1) Up to date (q2) Expected questions and literature searches (q17) Modeled questioning and searching (q18) Expected incorporating best evidence (q21) Modeled incorporating best evidence (q22)	6. Patient care Saw patients every day (q3) Independently evaluated each patient (q4) Reviewed daily care plan with team (q5) Contributed information and advice (q6) Helped speak with consultants (q8)
2. Bedside teaching Taught interviewing and communication (q14) Taught physical exam skills (q15)	7. Rounding Effective and efficient postcall rounds (q9)
3. Clinical reasoning Expected differential diagnosis and plan (q7) Expected decisions (q10) Explicit about clinical reasoning (q16) Expected a working diagnosis (q19) Modeled committing to working diagnosis (q20)	8. Professionalism Treated patients with respect (q27) Treated residents with respect (q28) Released residents for conferences (q29) Encouraged residents to call at any time (q30) Sensitive to social and cultural factors (q31)
4. Patient-based teaching Taught about patients' problems (q13)	9. Feedback Provided ongoing feedback (q23) Stated expectations at beginning (q25) [†] Provided midrotation feedback (q26) [†]
5. Teaching sessions Amount of teaching (q24)*	Summary score Mean score for all items

* Scores were adjusted for attending physicians' contact with residents (2 or 4 weeks) and transformed to a 5-point scale.

[†] Dichotomous responses were transformed to either 1 (no) or 5 (yes).

Reliability. We measured two types of reliability: the consistency among responses to items within each domain (Cronbach's coefficient α) and the consistency among different residents when they evaluated the same attending physician on the same domain (inter-rater reliability). The inter-rater reliability was analyzed in two ways. First, the inter-rater reliability coefficient was calculated to describe the proportion of the variation among physicians' mean scores that can be attributed to true physician differences. Second, we estimated the precision (twice the standard error, equivalent to a 95% confidence interval [CI]) of a physician's mean domain score under three scenarios—with 1, 4, or 8 resident evaluations. This second measure of inter-rater reliability describes the range of uncertainty around any score and is measured in the units of the original 5-point scale. The two measures of inter-rater reliability provide two complementary perspectives.

Validity. We assessed the instrument's face, content, and construct validity, but the accuracy, or criterion validity, of our instrument could not be assessed, because there is no gold standard. An instrument has face and content validity if its purpose and instructions are clear, the items are specific and unambiguous, the format and length acceptable, the gradations in response options appropriate, and if the instrument is comprehensive (includes all important variables and excludes unimportant ones). These characteristics were evaluated in two ways. First, we solicited feedback from residents through free-text responses at the end of each questionnaire and during the monthly, mandatory, end-of-rotation feedback sessions. Second, at the end of the year, we surveyed attending physicians after they had received their individual aggregate data to determine whether the results were believable and potentially useful.

To evaluate the instrument's construct validity, we conducted 4 tests to verify that it performed as expected. First, we tested whether the instrument could detect important and statistically significant differences between attending physicians on each domain. Second, we tested whether each domain contributed unique information and could thus discriminate among attending physicians even

after adjusting for all other domain scores. Third, we tested whether the instrument's summary score could predict which physicians had won teaching awards over the past 4 years (determined by a vote among residents without input from the formal evaluation process). Fourth, we tested whether the instrument's summary score was correlated with a related variable, how strongly the resident wanted to work with the attending physician again (measured at the end of each questionnaire with a single question and 5 response options, from definitely no to emphatically yes).

Statistical Analysis

Analyses were performed with SPSS statistical software, version 10 (SPSS Inc., Chicago, Ill). Interrater reliability coefficients were calculated using analysis of variance with resident considered a random factor.¹⁵ Calculating the standard error for means of 1, 4, and 8 measurements assessed precision of the domain scores. Tests of construct validity included analysis of variance, multivariable linear regression, logistic regression, and Spearman correlation. Model assumptions were verified for all analyses. All *P* values are two-sided.

RESULTS

Residents completed 731 evaluations of the 99 faculty who served as inpatient attending physicians during the 12-month data collection period (85% of all eligible evaluations), for an average of 7.4 evaluations per attending (range, 3 to 16).

Conceptual Domains

Table 1 describes the items for the 9 conceptual domains.

Reliability

As shown in Table 2, all multi-item domains, except feedback, demonstrated acceptable consistency among items (Cronbach's $\alpha > 0.8$). For interrater reliability, an

Table 2. Reliability of the Domain Scores and Summary Score

Domain	Internal Consistency (Cronbach's α)	Interrater Reliability: 1 Evaluation	Interrater Reliability: Average of 8 Evaluations
Evidence-based medicine	0.93	0.46	0.87
Bedside teaching	0.90	0.35	0.81
Clinical reasoning	0.88	0.36	0.82
Patient-based teaching*	—	0.36	0.82
Teaching sessions*	—	0.32	0.79
Patient care	0.86	0.43	0.86
Rounding*	—	0.33	0.80
Professionalism	0.83	0.29	0.77
Feedback	0.69	0.17	0.62
Summary score	0.95	0.45	0.87

* A Cronbach's α cannot be calculated for single-item domains.

Table 3. Precision of Domain Scores for 1, 4, or 8 Evaluations per Attending

Physician*	Precision: $\pm 2 \times \text{SE}$ (equivalent to 95% CI) [†]		
	1 Evaluation	4 Evaluations	8 Evaluations
Evidence-based medicine	± 0.6	± 0.3	± 0.2
Bedside teaching	± 0.9	± 0.4	± 0.3
Clinical reasoning	± 0.5	± 0.2	± 0.2
Patient-based teaching	± 0.6	± 0.3	± 0.1
Teaching sessions	± 1.0	± 0.5	± 0.3
Patient care	± 0.5	± 0.2	± 0.2
Rounding	± 0.7	± 0.4	± 0.2
Professionalism	± 0.3	± 0.2	± 0.1
Feedback	± 1.2	± 0.6	± 0.4
Summary score	± 0.4	± 0.2	± 0.1

* All scores are on a scale from 1 to 5.

[†] SE = standard error of the mean.

evaluation by a single resident provided an unreliable assessment of an attending physician's performance. However, if 8 residents were averaged (the number of residents an attending physician would supervise during two 4-week rotations at our hospital), the interrater reliability coefficients were high (>0.8) for 7 of the 9 domains (Table 2). For 2 domains—feedback and professionalism—the values for the reliability coefficient were relatively low (0.62 and 0.77). However, when interrater reliability was assessed using the CI (Table 3), professionalism was measured with good precision with 8 resident evaluations (95% CI ± 0.1), but feedback was measured with the least precision (CI ± 0.4).

Validity

Face and Content Validity. Residents confirmed that the instrument was acceptable, sensible, and comprehensive during the 13 feedback sessions. Attending physicians verified that the results were believable and potentially useful after they had reviewed a summary of their personal scores at the end of the year.

Construct Validity. The instrument is able to discriminate among physicians, as there were statistically significant differences among attending physicians within each domain (ANOVA, $P < .001$ for each domain). The greatest variation among physicians was in bedside clinical teaching (mean scores ranging from 1.1 to 5) and in teaching sessions (mean scores ranging from 1 to 5), where individual physician's mean scores spanned the full range of potential values (Fig. 1). The least variation was in professionalism (mean scores ranged from 3.8 to 5). On average, our attending physicians were rated highest on professionalism and lowest on teaching sessions (Fig. 1).

Each domain contributed unique information about attending physicians, because there were statistically significant differences among physicians within each domain, even after accounting for their performance in all other domains using multivariable linear regression (for each

FIGURE 1. Boxplots of the 99 attending physicians' average scores by domain. The gray box represents the distribution of the middle half of physicians' average scores; the vertical black line transecting the box represents the median value; the whiskers extend to the minimum average score and maximum average score.

domain, $P < .005$). Thus, the instrument can discriminate among the physicians' different roles and responsibilities represented by the 9 domains. The instrument's summary score was a strong predictor of those physicians who had recently won teaching awards; for every increase of 1 point in the summary score, the odds of having won a teaching award increased 17-fold (95% CI, 8 to 36).

Finally, there was a strong correlation between the instrument's summary score and the resident's desire to work with the attending physician again ($r = .79$; 95% CI, 0.74 to 0.83).

Summary Score and Individual Items

Among the individual items, the strongest predictor of an attending physician's summary score was the frequency that attending physicians explicitly described their clinical reasoning when discussing clinical decisions ($r^2 = .90$). This item of explicit clinical reasoning was also the best

predictor of a resident wanting to work with the attending physician again ($r^2 = .79$).

DISCUSSION

These findings strongly support the validity and reliability of our new evaluation instrument. Unlike others, this instrument evaluates attending physicians' educational and supervisory roles when caring for medical inpatients, measuring performance in several separate domains relevant to each of those responsibilities. Such comprehensive, detailed evaluations are needed now, given current expectations that teaching hospitals set new standards for evidence-based care, patient safety, and the training of our nation's future physicians.

Assessment of the instrument's validity yielded impressive results. Each of its 9 conceptual domains discriminated among attending physicians—evaluation of different physicians varied significantly within each domain—and each domain contributed specific, unique information. Its construct validity was supported by its powerful prediction of physicians who won teaching awards and residents' desire to work again with individual physicians. Although there is no gold standard to assess its criterion validity (accuracy), both residents and attending physicians attested to the instrument's face and content validity.

The instrument's reliability was also impressive. All domains except one (feedback) demonstrated good interrater reliability—either reliability coefficient >0.8 or CI $<\pm 0.4$. These findings highlight the importance of three issues. First, it is important to obtain multiple independent evaluations before judging an attending physician's performance. Second, measuring interrater reliability in more than one way provides a more comprehensive and clearer perspective of the results (for example, professionalism). Third, they demonstrate the need to improve reliability in evaluating feedback, a domain also rated lower in several previous studies.^{6,7,10,16} At least in part, this problem may be perceptual: some learners seem not to "recognize" feedback even when it has been well-documented.¹⁷

After analyzing our data we recognized the limitations of some of the items and revised them. The reliability for assessing feedback was unacceptably low and there was no evaluation of end-of-rotation feedback. Our revised questionnaire now assesses whether attending physicians are providing continuous feedback and we plan to test the reliability of these revisions. End-of-rotation feedback is important and captured in another instrument when the attending physician completes the written resident evaluation and discusses it with the trainee. Other revisions to the instrument have included an assessment of the quality, as well as quantity, of the teaching sessions. We also plan to improve and test the specificity of the bedside teaching items to ensure that these measures actually assess teaching time at the bedside.

It will be important to test whether our findings are generalizable to other institutions where attendings teach

and supervise inpatient care simultaneously. If so, use of the instrument would allow credible evaluation of clinical faculty in all such hospitals, help to standardize criteria for academic promotion of clinician-educators, and facilitate multicenter trials to improve attending performance. Table 3, for example, provides precise estimates of the upper and lower confidence limits for *individual* faculty members' scores in each domain. Such data can provide the basis for attending-specific or department-wide improvement efforts directed at any or all of the 9 domains. These metrics also can be used to study many unresolved questions about modern inpatient teaching services: are hospitalists the answer?^{18,19} Do generalists outperform specialists?²⁰ Can bedside teaching be resurrected?^{21,22} Can the inpatient practice of "evidence-based medicine" be promoted and improved?²³⁻²⁵

We note with interest domains whose ratings exhibited marked variation among attending physicians (Fig. 1). For example, evaluation of teaching sessions and bedside teaching had wide interquartile ranges, with individual physician scores spanning the full range of potential values. Although previous research illuminates some of the characteristics and practices of successful clinical teachers^{17,26-28} and role models,²⁹⁻³¹ little is known about how some physicians manage to teach more than others when working on a busy inpatient service. We hope to learn more about this from our own attending physicians who received the highest ratings in these domains.

We note also the pivotal importance of the query about explicit clinical reasoning. Intended to evaluate the attending physicians' conscious effort to make transparent to trainees the process of clinical problem solving, this one item was the dominant predictor of both the instrument's summary score and residents' desire to work with the attending again. It may not be surprising that learners want teachers to "think out loud," but we are unaware of previously published data about the importance of this issue to physicians-in-training. More research is needed to understand better its meaning and implications. Although we found this single item to be highly predictive of the summary score, this does not obviate the need for the other items in the questionnaire. For an instrument to provide useful feedback to academic ward attending physicians, one still requires specific questions that address each of their major roles and responsibilities.

Finally, it is important to study whether feedback of evaluations to attending physicians improves their performance, overall or in specific domains, individually and department-wide. If simple feedback were unproductive, targeted interventions in faculty development would be appropriate next steps. In fact, this same sequential strategy for evaluation and improvement makes sense in other clinical teaching venues: intensive care units, subspecialty consultation services, emergency departments, and outpatient clinics. For use in these venues, our evaluation instrument may need some site- and discipline-specific modifications. Nevertheless, in its current form,

the instrument provides a valid, reliable tool to begin to understand better the complex, critical responsibilities of attending faculty in teaching hospitals today.

The authors thank Drs. Maurice Lemon, Peter Clarke, Krishna Das, and Avery Hart for their valuable contributions to the design of the instrument and the medicine residents and attending physicians of Cook County Hospital who participated in the study.

Financial support came from the Department of Medicine, Cook County Hospital.

REFERENCES

1. Ende J. What if Osler were one of us? Inpatient teaching today. *J Gen Intern Med.* 1997;12(S2):S41-S48.
2. Evidence-based Medicine Working Group. Evidence-based medicine. A new approach to teaching the practice of medicine. *JAMA.* 1992;268:2420-5.
3. Barondess JA. Medicine and professionalism. *Arch Intern Med.* 2003;163:145-9.
4. Ahmed Mel-B K. What is happening to bedside clinical teaching? *Med Educ.* 2002;36:1185-8.
5. Irby DM. Clinical teacher effectiveness in medicine. *J Med Educ.* 1978;53:808-15.
6. Irby DM, Rakestraw P. Evaluating clinical teaching in medicine. *J Med Educ.* 1981;56:181-6.
7. Irby DM, Gilmore GM, Ramsey PG. Factors affecting ratings of clinical teachers by medical students and residents. *J Med Educ.* 1987;62:1-7.
8. Litzelman DK, Sratos GA, Marriott DJ, Skeff KM. Factorial validation of a widely disseminated educational framework for evaluating clinical teachers. *Acad Med.* 1998;73:688-95.
9. Copeland HL, Hewson MG. Developing and testing an instrument to measure the effectiveness of clinical teaching in an academic medical center. *Acad Med.* 2000;75:161-6.
10. Guyatt GH, Nishikawa J, Willan A, et al. A measurement process for evaluating clinical teachers in internal medicine. *CMAJ.* 1993;149:1097-102.
11. Williams BC, Litzelman DK, Babbott SF, Lubitz RM, Hofer TP. Validation of a global measure of faculty's clinical teaching performance. *Acad Med.* 2002;77:177-80.
12. Arreola RA. Issues in developing a faculty evaluation system. *Am J Occup Ther.* 1999;53:56-63.
13. Green ME, Ellis CL, Fremont P, Batty H. Faculty evaluation in departments of family medicine: do our universities measure up? *Med Educ.* 1998;32:597-606.
14. Bland CJ, Wersal L, VanLoy W, Jacott W. Evaluating faculty performance: a systematically designed and assessed approach. *Acad Med.* 2002;77:15-30.
15. Fleiss JL. Reliability of measurement. In: *The Design and Analysis of Clinical Experiments.* New York: John Wiley & Sons; 1986:1-32.
16. Ramsey PG, Gillmore GM, Irby DM. Evaluating clinical teaching in the medicine clerkship. *J Gen Intern Med.* 1988;3:351-5.
17. Irby DM. What clinical teachers in medicine need to know. *Acad Med.* 1994;69:333-42.
18. Auerbach AD, Wachter RM, Katz P, Showstack J, Baron RB, Goldman L. Implementation of a voluntary hospitalist service at a community hospital: improved clinical efficiency and patient outcomes. *Ann Intern Med.* 2002;137:859-69.
19. Meltzer D, Manning WG, Morrison J, et al. Effects of physician experience on costs and outcomes on an academic general medicine service: results of a trial of hospitalists. *Ann Intern Med.* 2002;137:866-74.
20. Weingarten SR, Lloyd L, Chiou C, Braunstein GD. Do subspecialists working outside of their specialty provide less efficient and lower quality care to hospitalized patients than do primary care physicians? *Arch Intern Med.* 2002;162:527-32.
21. LaCombe MA. On bedside teaching. *Ann Intern Med.* 1997;126:217-20.
22. Kroenke K, Omori DM, Landry FJ, Lucey CR. Bedside teaching. *South Med J.* 1997;90:1069-74.
23. Ellis J, Rowe J, Sackett DL. Inpatient general medicine is evidence based. *Lancet.* 1995;346:407-10.
24. Michaud G, McGowan JL, van der Jagt R, Wells G, Tugwell P. Are therapeutic decisions supported by evidence from health care research? *Arch Intern Med.* 1998;158:1665-8.
25. Nordin-Johansson A, Asplund K. Randomized controlled trials and consensus as a basis for interventions in internal medicine. *J Intern Med.* 2000;247:94-104.
26. Pinsky LE, Irby DM. "If at first you don't succeed": using failure to improve teaching. *Acad Med.* 1997;72:973-6.
27. Pinsky LE, Monson D, Irby DM. How excellent teachers are made: reflecting on success to improve teaching. *Adv Health Sci Educ Theory Pract.* 1998;3:207-15.
28. O'Malley PG, Kroenke K, Ritter J, Dy N, Pangaro L. What learners and teachers value most in ambulatory educational encounters: a prospective qualitative study. *Acad Med.* 1999;74:186-91.
29. Wright S. Examining what residents look for in their role models. *Acad Med.* 1996;71:290-2.
30. Wright SM, Kern DE, Kolodner K, Howard DM, Brancati FL. Attributes of excellent attending-physician role models. *N Engl J Med.* 1998;339:1986-93.
31. Wright SM, Carrere JA. Excellence in role modeling: insights and perspectives from the pros. *CMAJ.* 2002;167:638-43.

Attending Physician Evaluation Form

Clinical Supervision

<i>The attending</i>	Never					Always				
1) ... demonstrated a broad knowledge of medicine.	1	2	3	4	5					
2) ... was up to date.	1	2	3	4	5					
3) ... saw all patients every day.	1	2	3	4	5					
4) ... independently evaluated each patient.	1	2	3	4	5					
5) ... reviewed the care plan of each patient with the team every day.	1	2	3	4	5					
6) ... contributed additional clinical information or advice when needed.	1	2	3	4	5					
7) ... expected me to develop a thorough differential diagnosis and management plan for each active problem.	1	2	3	4	5					
8) ... helped speak with consultants and helped arrange tests in order to provide the best care for the patients.	1	2	3	4	5					
9) ... conducted post-call rounds in an effective and efficient manner.	1	2	3	4	5					
10) ... required me to be an active decision-maker in patient care, rather than always following the attending's lead.	1	2	3	4	5					
11) Aside from post-call rounds, did the attending participate in bedside work rounds (7:30-9 am) for the first week on service and at least twice a week thereafter?	1 No					5 Yes				
12) Did work rounds always take place at the bedside?	1 No					5 Yes				

Teaching Activities and Feedback

<i>The attending</i>	Never					Always				
13) ... made sure the teaching sessions pertained to patient problems.	1	2	3	4	5					
14) ... effectively taught interviewing and communication skills.	1	2	3	4	5					
15) ... effectively taught physical examination skills.	1	2	3	4	5					
16) ... was always explicit about his or her reasoning when discussing clinical decisions.	1	2	3	4	5					
17) ... expected me to be an active learner, by requiring me to ask focused questions, finding the best literature, and sharing my findings with the team.	1	2	3	4	5					

	Never			Always		
18) ... personally modeled active, continuous learning by asking questions, searching the literature, and sharing his or her findings.	1	2	3	4	5	
19) ... expected me to commit to a working diagnosis.	1	2	3	4	5	
20) ... personally modeled committing to a working diagnosis.	1	2	3	4	5	
21) ... expected me to incorporate the best evidence from the literature with the patient's unique circumstances and preferences.	1	2	3	4	5	
22) ... personally modeled incorporating the best evidence from the literature with the patient's unique circumstances and preferences.	1	2	3	4	5	
23) ... provided ongoing feedback on my history-taking skills, physical exam skills, written documentation, oral presentations, and clinical reasoning.	1	2	3	4	5	
24) What's your best estimate for the number of teaching sessions with the team and the attending? (Consider a "teaching session" to mean a minimum of 45 minutes devoted to education.)						
25) ... clearly stated his or her expectations for my performance this rotation?	1	2	3	4	5	
26) ... provided specific, detailed mid-rotation feedback?	1	2	3	4	5	

Professional Values

The attending

	Never			Always		
27) ... treated the patients with respect.	1	2	3	4	5	
28) ... treated me with respect.	1	2	3	4	5	
29) ... released me for all scheduled conferences.	1	2	3	4	5	
30) ... encouraged me to call him or her at any time for any reason.	1	2	3	4	5	
31) ... was sensitive to the emotional, economic, social, and cultural aspects of the patients' illnesses.	1	2	3	4	5	
	Absolutely No			Enthusiastically Yes		
32) Would you like to work with this attending again?	1	2	3	4	5	

Please write comments below to help us better evaluate the attending's performance.

What are this attending's top three strengths?

How could this attending be more effective?