Creation and Assessment of a Structured Review Course in Physical Diagnosis for Medical Residents

SALVATORE MANGIONE, MD, STEVEN J. PEITZMAN, MD, EDWARD GRACELY, PhD, LINDA Z. NIEMAN, PhD

Objective: To evaluate the effects of a course in physical diagnosis on the knowledge, skills, and attitudes of internal medicine trainees.
Design: A controlled, prospective assignment of housestaff to a year-long curricular program, linked to a set of pre- and posttests. House-officers who could not attend the teaching sessions functioned as control subjects.
Setting: An internal medicine training program at an urban medical school.
Subjects: 56 (86.1%) of 65 eligible internal medicine housestaff (post-graduate years 1 through 3) participated in the intervention and assessment. A comparison group of 14 senior medical students participated in the pretest.
Intervention: 12 monthly lectures emphasizing skills useful in emergencies or validated by the literature.
Measurements: The pre- and posttests included: 1) a multiple-choice questionnaire to assess knowledge; 2) professional standardized patients to assess selected skills; and 3) Likert-type questionnaires to assess self-motivated learning and attitude toward diagnosis not based on technology.
Main results: The residents expressed interest in the program and on a six-point scale rated the usefulness of lectures and standardized patients as 3.5 ± 1.3 and 4.3 ± 1, respectively. For no system tested, however, did they achieve more than 55.2% correct answers (range: 24.2%–55.2%, median = 41.04), and their performance did not differ from that of the fourth-year medical students. There was no significant difference in pre/posttest improvement between the control and intervention groups.
Conclusions: These data confirm the deficiencies of physical diagnostic skills and knowledge among physicians in training. These deficiencies were not corrected by the classroom lecture series. Improvement in these skills may require a more intense experiential program made part of residency requirements.
Key words: education; residents; physical examination; diagnostic skills; performance assessment.
J GEN INTERN MED 1994;9:213–218.

SURVEYS SHOW that internists in practice value physical examination as one of the most useful sets of skills acquired during medical school and internal medicine training.[1, 2] This cost-effective and time-honored art is traditionally learned at the bedside, through experience with patients and supervision by attending physicians and more senior trainees. Structured direct evaluation of competency in physical diagnosis as a prerequisite for board certification is performed usually once during the three years of residency. Although many authors have shown that faculty observation of trainees' clinical skills is often inadequate during this "clinical evaluation exercise,"[3–5] the physical diagnosis proficiency of most residents is assumed to be satisfactory by completion of training.

There are, however, increasing reasons to believe that in our age of high technology this assumption may no longer be valid. Because of the widespread availability of sophisticated (and expensive) diagnostic tools, the time devoted to physical diagnosis is diminishing in most programs. Attending physicians are now spending as little as 16% of ward rounds time at the bedside, with only half of this time dedicated to history gathering and performance of physical examination.[6] As a result, the traditional methods of bedside supervision and teaching have been disrupted, and proficiency in physical diagnosis is suffering. Inaccuracies in physical examination are encountered frequently among trainees.[7–9] In one study, serious errors occurred in as many as two-thirds of the patient examinations reviewed.[7]

Because more than 20% of trainees' errors occur during the cardiovascular examination,[7] we surveyed the time and importance given to cardiac auscultation during training.[10] Only one fourth of internal medicine residencies and one third of cardiology fellowships provide formal teaching in cardiac auscultation, although most program directors consider this "a skill that all practicing physicians should master." As a consequence of limited teaching and lack of bedside supervision, proficiency in physical diagnosis is waning. Internal medicine residents of the mid-Atlantic area, for example, had major difficulty identifying 12 cardiac auscultatory findings, improved little with year of training, and were never better than a group of third-year medical students.[10] Similar results were shown by St. Clair et al.[11] Thus, a generation of physicians is now being trained with inadequate emphasis on bedside clinical skills, and with excessive reliance on technologic diagnosis. This approach to the patient is expensive and potentially dangerous, and threatens our clinical heritage.

It has been recommended that formal courses on physical diagnosis be organized by program directors in internal medicine for the benefit of their housestaffs.[12] Although no formal guideline has been provided, time constraints in a residency program and the expanding responsibilities of trainees suggest that these courses should probably address only skills so important that all residents would be expected to master them by the end of training. With this premise in mind, we developed, implemented, and field-tested a formal course in se-

Received from the Departments of Medicine and of Community and Preventive Medicine and the Office of Medical Education, the Medical College of Pennsylvania, Philadelphia, Pennsylvania.
Presented in part at the annual meeting of the Society of General Internal Medicine, April 29, 1993, Arlington, Virginia.
Supported in part by a grant from the American Board of Internal Medicine.
Address correspondence and reprint requests to Dr. Mangione: Department of Medicine, the Medical College of Pennsylvania, 3300 Henry Avenue, Philadelphia, PA 19129.

TABLE 1

Design and Content of the Physical Diagnosis Course and Its Evaluation Instrument

Physical Diagnosis Lecture	Number of Hours	Special Teaching Tools	Number of Test Questions: Knowledge-based	Number of Test Questions: Audiovisual-based
The lasting value of physical diagnosis	1	—	—	—
Cardiac examination	3	Heart sounds synthesizer	4	7
Pulmonary examination	1	Lung sounds (recorded)	14	6
Dermatologic examination	1	Slides of skin lesions	1	7
Ophthalmologic examination	2	Slides of eye lesions	1	4
Breast examination	1	—	1	1
Rheumatologic examination	2	Videotape of shoulder examination	10	—
Neurologic examination	1	—	12	1

lected areas of physical diagnosis. Internal medicine residents of the Medical College of Pennsylvania (MCP) were pre- and posttested to detect changes in knowledge, skills, and attitudes as a result of their participation in the course. In this report, we describe our experience over one academic year.

METHODS

Setting and Subjects

The study was conducted as part of the teaching activities of the MCP Department of Medicine. At the beginning of the academic year (1991), all internal medicine residents in the three postgraduate years of training (PGY1–3, $n = 65$) were invited to participate in both the instruction and the assessment program. All medical students serving on medicine and its subspecialties were also invited to attend the teaching sessions.

Instruction and Assessment Programs

Educational Program. The program consisted of a structured curricular review of selected elements of the physical examination, woven into the regular conference schedule of the MCP internal medicine program.

We designed the course around two "themes": 1) the use of physical examination in emergencies or near-emergency situations, when the initial physician should make at least a tentative diagnosis before calling for a

TABLE 2

Material Covered by the Written Multiple-choice Examination (Pre- and Posttest)*

	Knowledge-based Questions	Sound-recognition Questions	Slide-recognition Questions
Cardiac examination	Clinical relevance of S4 gallop Characteristics of mitral stenosis murmur Interpretation of pulsus paradoxus Interpretation of S2 splitting	Pericardial friction rub Aortic stenosis murmur Aortic regurgitation (AR) murmur S3 gallop Mitral valve prolapse click	Proper positioning for AR murmur
Pulmonary examination	Differential diagnosis of crackles Physical findings of pneumothorax Characteristics and interpretation of bronchial breath sounds	Pleural friction rub Bronchial breath sounds Bronchial breath sounds and late-inspiratory crackles	
Dermatologic examination	Drug skin allergy (phenytoin)		Lupus pernio Erythema nodosum Basal cell carcinoma Superficial spreading melanoma Seborrheic keratosis Stevens-Johnson syndrome
Ophthalmologic examination	Characteristics of diabetic retinopathy		Cholesterol emboli of the retina Acute glaucoma Hyphema (anterior chamber bleed) Post-laser-therapy retinal changes
Breast examination	Differential diagnosis of breast discharge		Technique of breast palpation
Rheumatologic examination	Examination of shoulder, elbow, hip, knee, and acromioclavicular joints		
Neurologic examination	Examination of coma Findings of spinal cord compression Localization of cortical lesions Increased intracranial pressure Corneal reflex and its pathways		Decerebrate posturing

*Many topics were assigned more than one question.

technologic aid or consultant (e.g., acute glaucoma, tamponade, pneumothorax, rotator cuff tear, Stevens-Johnson syndrome, assessment of coma, spinal cord compression); and 2) physical examination findings that have been validated in the literature, or by near consensus are accepted as providing information equal to or better than that provided by diagnostic technology (e.g., pleural and pericardial friction rubs, signs of pulmonary consolidation, S3 gallop, mitral valve prolapse click). Breast examination was included in our program because of its particular importance in preventive medicine.

Faculty members with a special interest in physical diagnosis and a known reputation for effective teaching styles participated as lecturers. Twelve hour-long teaching sessions, structured around the course's two themes and covering physical examination skills pertinent to both ambulatory and inpatient evaluations, were held once a month (Table 1). Instructors used slides, videotapes, and audiotapes; an infrared, wireless, remote stethophone system and synthesizer (Cardionics Inc., Houston, TX) was used to demonstrate heart and lung sounds. A letter urging (but not requiring) participation in the teaching and assessment activities was signed by the chairman of medicine and sent to all houseofficers prior to the beginning of the course. For each topic, the intervention group included all trainees who completed the pretest, attended the pertinent teaching session(s), and completed the posttest. The control group included those houseofficers who completed both tests but could not attend the teaching session(s), presumably because of assignments outside MCP.

Evaluation. To assess the effect of our program on the residents' knowledge, skills, and attitudes, we designed three tests: 1) a test of knowledge in physical diagnosis; 2) a test of proficiency in selected physical examination skills; and 3) a test of attitude toward physical diagnosis.

The knowledge component of the test contained 69 multiple-choice and short-answer items, selected in coordination with the course lecturers. One third of these questions required identification of audiovisual material, such as prerecorded heart and lung sounds and slides of ophthalmologic and dermatologic abnormalities (Table 2). These audiovisuals were similar, but not identical, to those used during the corresponding teaching sessions. The trainees' performance on the test was subdivided into an Audiovisual Recognition Score (AVRS) and a Cognitive Score (CS), depending on the type of test item. Both the AVRS and the CS were expressed as a percentage of all pertinent questions answered correctly.

The skill component of the test was based on four standardized patient (SP) exercises. As part of these tasks the residents had to demonstrate: 1) auscultation techniques (patient position, area of auscultation) when suspecting aortic regurgitation; 2) chest examination when suspecting pleural effusion; 3) shoulder examination, including special maneuvers, for a patient with shoulder pain; and 4) neurologic examination for a patient with complaints suggesting spinal cord compression. Patient 4 simulated positive neurologic findings. The other patients did not have positive findings: we were assessing mainly the residents' techniques. The SPs, trained by the organizers, functioned as both subjects for the examination and evaluators. The extent to which skills were correctly performed was graded using a checklist of the sort now customary in SP programs. Scores for each houseofficer were expressed as the percentage of all skills/maneuvers performed correctly in each individual exercise.*

The attitude component of the test was based on a 35-item questionnaire designed to assess interest in physical diagnosis and self-motivated learning. In addition, the residents were asked to rate the clinical importance of physical diagnosis for selected areas of practice. This was measured by an Attitude Toward Traditional Diagnosis (ATTD) score, which consisted of 19 Likert-type items. Using a six-point scale the respondents rated various diagnostic modalities, including history and physical diagnosis (traditional tools), computed tomography scanning, roentgenology, ultrasonography, and nuclear medicine (technology-based tools). The perceived importance of traditional versus technologic methods was then expressed as a fraction ranging from 6/1 (traditional very important) to 1/6 (traditional less important) for each physical diagnosis category assessed (pulmonary medicine, cardiology, rheumatology, and neurology). The ATTD score was computed by averaging the four individual scores.

To assess the impact of medical training on attitudes toward physical diagnosis, this questionnaire was administered not only to the participating residents (to compare attitudes of junior and senior residents) but also to a group of 53 MCP freshman students.

The residents' knowledge and attitudes were assessed before the first teaching session and again at the end of the academic year. The residents' skills were evaluated by SPs prior to and again within four weeks of the pertinent teaching session. The posttests were identical to the pretests in content, administration, and scoring. At the close of the academic year all the participating residents were asked to anonymously rate the usefulness of lectures and SP exercises, as well as the effect of the course on their own interest in physical diagnosis.

Statistical Analysis

Analyses consisted of analyses of variance (ANOVAs) with Tukey honestly significant difference (HSD) multiple comparisons follow-ups, t-tests, and Pearson correlations. A significance level of 0.01 was used to control type I error. Specific analyses are described in more detail in the results section.

*Checklists for each of the four standardized patient exercises are available from the authors upon request.

RESULTS

Because of the differences in our evaluation instruments, we review separately the results pertaining to knowledge, skills, and attitude.

Knowledge in Physical Diagnosis

Pretest Analyses. The analysis of the pretest knowledge scores involved the 56/65 (86.1%) eligible residents who completed both pre- and posttesting. Comparisons among the years of residency consisted of one-factor ANOVAs. Although the mean cumulative test score and the mean ophthalmology CS were close to significance ($p = 0.012$ for both), only the mean cardiology CS and the mean ophthalmology AVRS attained $p \leq 0.01$ ($p = 0.0099$ and $p = 0.006$, respectively). Multiple comparisons indicated that PGY3s were more knowledgeable than PGY1s for the cardiology section ($p < 0.05$) and more accurate than PGY2s for the slide-recognition section of the ophthalmology examination.

Because of the interest in possible pretest differences between the residents and medical students, for each area of physical diagnosis we compared the residents' performance to that of a group of 14 senior medical students serving on medicine and its subspecialties at the time of pretesting. Comparisons consisted of unpaired t-tests. There was no significant difference between the residents and the medical students on any of the scores tested. Although the residents' means were superior in eight of the 11 scores, except for the cardiology examination this superiority never exceeded an average of 6 percentage points. The two most interesting results were in opposite directions: the residents scored higher on the sound-recognition portion of the cardiac examination ($p = 0.044$) and lower on the knowledge section of the breast examination ($p = 0.045$). Again, these differences did not meet our significance level of 0.01. Overall, the residents never achieved more than 55.2% correct answers in any of the systems tested (range: 24.2%–55.2%, median = 41.04) and were never significantly better than the senior medical students.

Pre/Posttest Comparison. The comparison of pre- and posttesting knowledge scores involved only the 56 residents (and no student) who completed both tests. Analysis was carried out according to year of training, attendance or not at the pertinent lecture(s), and time of testing (pre and post) in three-factor ANOVAs. The great majority of the effects analyzed were not significant even at the 0.05 level, let alone the 0.01 criterion used for significance. There was a clear effect of year of training for the ophthalmology CS ($p = 0.001$), mainly reflecting the superior scores of PGY3s, in particular five PGY3s who did not attend the ophthalmology lectures (year-by-attendance-to-lectures interaction, $p = 0.025$). The same pattern emerged for the slide-recognition section of the ophthalmology examination, except that both the year of training ($p < 0.001$) and the year-by-attendance-to-lectures interaction ($p = 0.003$) attained our criterion of significance. For the slide-recognition portion of the dermatology examination, a year-by-time-of-testing interaction ($p = 0.01$) emerged, suggesting greater pre- to posttest changes for some years of training than for others. For the section on breast examination, a year-

TABLE 3

Percentage of Correct Answers by the Internal Medicine Residents and the Senior Medical Students on the Written Multiple-choice Examination*

	Residents (Nonattendees)		Residents (Attendees)		Senior Students
	Pretest	Posttest	Pretest	Posttest	Pretest
Cardiac examination					
CS	53.2 (20)	57.3 (20)	55.05 (36)	59.09 (36)	50.7 (14)
AVRS	37.5 (20)	41.7 (20)	38.4 (36)	41.7 (36)	27.4 (14)
Pulmonary examination					
CS	40.2 (26)	42.3 (26)	42.8 (30)	42.8 (30)	41.4 (14)
AVRS	46.8 (26)	46.2 (26)	50 (30)	45.6 (30)	51.2 (14)
Dermatologic examination					
CS	42.1 (43)	43.3 (43)	36.5 (13)	49 (13)	40.2 (14)
AVRS	39.1 (43)	41.4 (43)	34.6 (13)	46.1 (13)	38.1 (14)
Ophthalmologic examination					
CS	24 (35)	32 (35)	24.8 (21)	28.6 (21)	19.6 (14)
AVRS	24.3 (35)	33.6 (35)	23.8 (21)	29.8 (21)	20 (14)
Breast examination (CS)	53.8 (39)	52.6 (39)	52.9 (17)	50 (17)	67.9 (14)
Rheumatologic examination (CS)	35.4 (22)	39.5 (22)	42.6 (34)	44.4 (34)	41.4 (14)
Neurologic examination (CS)	53.8 (30)	45.9 (30)	49.4 (26)	52.4 (26)	46.2 (14)

*When applicable, scores for each section are presented as both Cognitive Scores (CSs) and Audiovisual Recognition Scores (AVRSs). The residents' pre- and posttest scores are subdivided based on attendance at the corresponding lectures. The students' scores refer only to pretest. (*n*) indicates the number of residents or students in each group. There was no significant pre/post change for the residents on any measure, nor any indication of greater change among the attendees than among the nonattendees ($p > 0.05$ for all).

by-lecture interaction ($p = 0.003$) was noticed, indicating greater attendee/nonattendee differences for some years of training than for others.

For eight of the 11 scores there was an improvement from pre- to posttesting, albeit small and nonsignificant. For the remaining three scores there was a decline. Thus, the residents' deficiencies in physical diagnosis were not corrected over the course of one year, either as a result of participation in our program or as a result of the customary training (Table 3).

Skill in Physical Diagnosis

Only a subgroup of the residents (and no student) completed the SP exercise pre- and posttests (Table 4). None of the effects from the three-way ANOVA were significant for cardiology ($p > 0.10$ for all), pulmonary medicine ($p > 0.15$ for all), or rheumatology ($p > 0.09$ for all). For neurology there was an overall pre/post improvement only for those residents who had attended the lecture (attendance-to-lecture-by-time-of-testing interaction, $p = 0.014$), although this did not reach our level of significance, set at 0.01.

Attitudes toward Physical Diagnosis

The ATTD scores were analyzed for the 45/56 residents who completed the questionnaire and also for 53 freshmen medical students. Comparisons between the residents and the students consisted of unpaired t-tests. Only the rheumatology scale even approached statistical significance ($p = 0.062$, with the students attributing greater importance than the residents). Overall, the residents' ATTD scores ranged at pretest between 0.88 and 1.66 (median = 1.15, with 6 and 1/6 indicating, respectively, the highest and lowest importance attributed to traditional diagnosis). Attendance at lecture, year of training, and time of testing (pre or post) had no effect on the residents' scores ($p < 0.10$ for all by three-factor ANOVAs). There was no significant correlation between ATTD score at posttest and number of lectures attended ($r = 0.0838$, $p = 0.283$).

The ratings of the program and of its ability to stimulate interest in physical diagnosis were analyzed for all 56 residents. The trainees expressed good general interest in the program, rating usefulness of lectures and SPs as 3.5 ± 1.3 and 4.3 ± 1, respectively (mean $\pm$ SD, with 1 indicating "useless" and 6 indicating "very useful"). They also felt that the program had increased their interest in physical diagnosis (mean $\pm$ SD: 3.9 ± 1.2 on a 1–6 scale, with 1 indicating "not increased" and 6 indicating "much increased"). The scores of objective interest in physical diagnosis, however, and the scores of self-motivated learning did not change significantly between the pre- and posttests. They also did not correlate with the number of lectures attended, suggesting, again, that better attendance in the course did not improve the trainees' attitudes toward physical diagnosis.

TABLE 4

Percentage of Skills/Maneuvers Performed Correctly by the Internal Medicine Residents on the Four Standardized Patient Exercises*

	Nonattendees		Attendees	
	Pretest	Posttest	Pretest	Posttest
Cardiac examination (aortic regurgitation)	67.8 (9)	61.1 (9)	78 (15)	79.3 (15)
Pulmonary examination (pleural effusion)	56 (10)	62.5 (10)	65.3 (14)	62.1 (14)
Rheumatologic examination (shoulder)	36 (24)	38 (24)	37.6 (10)	46 (10)
Neurologic examination (spinal cord compression)	57.7 (18)	61.8 (18)	60 (10)	77.4 (10)

*Pre- and posttest performance is shown on the basis of the residents' attendance at the corresponding lectures. (*n*) indicates the number of residents in each group who participated in the exercises. There was no significant pre/post change for the residents on any measure, nor any indication of greater change among the attendees than among the nonattendees, with the only exception of neurology, where there was a greater (but not significant) pre/post improvement among the attendees ($p = 0.014$).

DISCUSSION

We undertook the creation of a physical examination course for medical residents with optimism and enthusiasm. The project seemed consistent with calls for "curriculum" in residency training and with a fostering of greater reliance on bedside skills. We found ourselves after one year disappointed but not dismayed, sobered but still supportive of physical diagnosis and its teaching. Even in our age of expanding technology, this time-honored art remains a valuable tool for all practicing physicians because it provides valuable information for a correct diagnosis,[13] guides selection of costly diagnostic technology, gives a baseline for serial observations, and preserves the therapeutic value of the physical contact with the patient. Even more important, it can restore the intellectual pleasure of a bedside diagnosis made only with our physical senses.

And yet, we found very little benefit for participation in our review course. The best support for the value of our program would have been a greater improvement in the attendees' scores compared with those of the nonattendees. This, however, never attained a p-value of 0.05, let alone the 0.01 criterion set for significance. Moreover, there was never a significant change between pre- and posttesting. That is, neither customary training over one year, nor our course, produced improvement.

Why was there essentially no detectable gain in knowledge or performance? What were the possible problems with the course, the learners, or the evaluation instruments? Clearly, our study had two methodologic limitations. First, we used illustrative slides (eye, skin) or electronic sounds (heart, lungs) for several of the teaching sessions and pertinent questions in the pre- and posttests. Though these are conventional teaching aids,

there is little experimental evidence that accuracy in recognizing images or electronic sounds correlates with accuracy at the bedside. Because the teaching sessions and the evaluation instruments used the same tools, and because the sounds and images demonstrated were intentionally selected as clear and distinct representations of "classic" findings, this probably represented only a minor limitation to our study.

Second, we allowed just a few knowledge questions for each topic (heart, lungs, joints, etc.) to avoid excessively long pre- and posttests. Power analysis indicated that only quite large effects, such as a mean difference between groups equal to one standard deviation, would have had 80% power given our sample size. Thus, our instrument could detect only relatively large gains in knowledge or skill for each tested area. Although this represents a limitation of our study, small gains would probably not have justified the trouble of conducting the instructional program.

Methodologic considerations aside, the major problem of our study was attendance at the series, which was variable, not easy to enforce, and disappointingly low. While many residents seemed enthusiastic about the program (and indeed promoted it as an asset when meeting with senior students interviewing for the residency), others apparently saw physical diagnosis as a less compelling subject than the more customary fare for housestaff conferences. Those who attended seemed outwardly interested in the presentations and rated them positively for the most part. It remains unclear, however, what a houseofficer (or anyone else) learns and retains from any single didactic "conference." Listening, albeit attentively, may not be enough.

Traditional "noon conference" topics for internal medicine programs, such as metabolic acidosis or respiratory failure, are also taught by day in the clinical setting, as well as actively "studied" at night to pass certification examinations. This is not so for physical diagnosis. Teaching rounds dwell too often in the conference room, and residents are observed too infrequently at the bedside. Moreover, independent study at night usually neglects this area, which is not extensively tested during written certification examinations and is assumed to be already "mastered" by most residents. Because neither attendance at our lectures nor achieving a certain score in our tests was in any way a criterion for progressing in the residency program, there was little "encouragement" toward independent study. It is disturbing to notice that even the control group did not improve with just one year of customary training, suggesting that residency per se may not correct the physical examination deficiencies observed in graduating medical students.[14]

In conclusion, our study further validates the mounting evidence that physicians in training are not proficient in physical diagnosis. Our trainees improved little with year of residency and were not better than a group of senior medical students. Thus, physical diagnosis should gain more attention both in medical schools and residency programs. As we could show little benefit from a classroom review, innovative ways to teach and reinforce physical examination skills need to be further explored. Particular attention should be given to those methods and signs most useful in our technologic era. Perhaps attendance at bedside teaching activities should be mandatory and residents required to demonstrate proficiency as they progress through their training. Because physical examination is best learned with patients and experienced clinicians, "teaching should return to the bedside, rather than take place in conference rooms."[15]

The authors express their thanks to Drs. Nathan Blank, Edward Bondi, Marvin Greenbaum, Mark Lopatin, Ajit Sachdeva, and Alex Sokil for their enthusiastic help in organizing the course lectures.

REFERENCES

1. Kern DC, Parrino TA, Korst DR. The lasting value of clinical skills. JAMA. 1985;254:70–6.
2. Mandel JH, Rich EC, Luxenberg MG, Spillane MT, Kern DC, Parrino TA. Preparation for practice in internal medicine. A study of ten years of residency graduates. Arch Intern Med. 1988;148:853–6.
3. Herbers JE Jr, Noel GL, Cooper GS, Harvey J, Pangaro LN, Weaver MJ. How accurate are faculty evaluations of clinical competence? J Gen Intern Med. 1989;4:202–8.
4. Noel GL, Herbers JE Jr, Caplow MP, Cooper GS, Pangaro LN, Harvey J. How well do internal medicine faculty members evaluate the clinical skills of residents? Ann Intern Med. 1992;117: 757–65.
5. Norcini JJ, Webster GD, Grosso LJ, Blank LL, Benson JA. Ratings of residents' clinical competence and performance on certification examination. J Med Educ. 1987;62:457–62.
6. Collins GF, Cassie JM, Daggett CJ. The role of the attending physician in clinical training. J Med Educ. 1978;53:429–31.
7. Wray NP, Friedland JA. Detection and correction of house staff error in physical diagnosis. JAMA. 1983;249:1035–7.
8. Aloia JF, Jonas E. Skills in history-taking and physical examination. J Med Educ. 1976;51:410–5.
9. Wiener S, Nathanson M. Physical examination. Frequently observed errors. JAMA. 1976;236:852–5.
10. Mangione S, Nieman LZ, Gracely E, Kaye D. The teaching and practice of cardiac auscultation during internal medicine and cardiology training. A nationwide survey. Ann Intern Med. 1993; 119:47–54.
11. St. Clair EW, Oddone EZ, Waugh RA, Corey GR, Feussner JR. Assessing housestaff diagnostic skills using a cardiology patient simulator. Ann Intern Med. 1992;117:751–6.
12. Hurst JW. Osler as visiting professor: house pupils plus six skills. Ann Intern Med. 1984;101:546–9.
13. Peterson MC, Holbrook JH, Von Hales D, Smith NL, Staker LV. Contributions of the history, physical examination and laboratory investigation in making clinical diagnoses. West Med J. 1992; 156:163–5.
14. Sox HC, Morgan WL, Neufeld VR, Sheldon GF, Tonesk X. Subgroup report on clinical skills. J Med Educ. 1984;59(11 pt 2):139–47.
15. Fitzgerald FT. The case for internal medicine. N Engl J Med. 1993;328:654–6.