Error Patterns of 3rd-Year Medical Students on the Cardiovascular Physical Examination

Carmen Ortiz-Neu
Division of General Medicine
Columbia University Health Sciences
New York, New York, USA

Joseph Tenenbaum
Department of Medicine
Columbia University Health Sciences
New York, New York, USA

Cheryl A. Walters
Center for Education Research and Evaluation
Columbia University Health Sciences
New York, New York, USA

Jerry A. Colliver
Division of Statistics and Research Consulting
Southern Illinois University School of Medicine
Springfield, Illinois, USA

Hilary J. Schmidt
Center for Education Research and Evaluation
Columbia University Health Sciences
New York, New York, USA

Background: *Recent research documents widespread deficits in the physical examination skills of practicing physicians.*
Purpose: *This study explored physical examination skills of 3rd-year medical students after completion of a course in physical diagnosis.*
Methods: *Standardized patient physical examination checklist data were analyzed for a cohort of 2,038 medical students for a patient presenting with classic signs and symptoms of an acute myocardial infarction. A follow-up paper case and survey explored reasons underlying omissions.*
Results: *Students systematically omitted 3 of 10 component maneuvers critical to the evaluation of a patient with shortness of breath and chest pain. The same pattern of omissions was observed across 8 medical schools and over 2 successive years. The paper case follow-up study ruled out time constraints and performance anxiety as the cause. Survey data revealed that students may omit a maneuver due to inability to recall pertinence (blood pressure in both arms) or difficulty discriminating findings (heart sounds at different locations), or because of inadequate technical mastery (percussion of the lungs).*
Conclusions: *These data suggest fundamental inadequacies in the current paradigm for teaching physical examination skills. Standardized patient checklist data can provide an informative window into the efficacy of teaching practices.*

Teaching and Learning in Medicine, *13*(3), 161–166

Cognitive psychologists have long recognized that many human errors are quite systematic in character, and that the study of human errors may provide a window into the organization of knowledge in memory and constraints on information processing and performance. The study of errors in medicine likewise reveals that "mistakes" are frequently not random, but rather reflect systematic problems within the healthcare system.[1] The *Journal of the American Medical Association*[2] recently published an article recommending that the profession analyze medical errors and regard them as opportunities to understand and correct deficiencies in medical practice. Many errors in medicine are the result of systems failure, fatigue, and other psychological factors. When errors occur because of lack of knowledge or skills, the clinician–educator's attention is focused on how medical training may be

This work was supported in part by Grant P116B60417, The Comprehensive Program, Fund for the Improvement of Postsecondary Education, U.S. Department of Education.

This article was presented in part at the annual session of the American Educational Research Association, San Diego, CA, April 17, 1998.

We thank Mark H. Swartz and the staff of the Morchand Center for Clinical Competence at the Mount Sinai School of Medicine, Ethel Crawford and Flora Atkins for assistance with data collection, and Daniel Schwartz for data presentation.

Correspondence may be sent to Carmen Ortiz-Neu, Division of General Medicine, Columbia University College of Physicians and Surgeon, 630 West 168th Street, 9th Floor, Room 105N, New York, NY 10032, USA. E-mail: co6@columbia.edu

improved to ensure more efficient, predictable, and safe outcomes.

High rates of errors, ranging from 13% to 80%, in physical examination skills have been documented for medical school graduates.[3–8] Converging evidence suggests that errors in physical examination skills may be traced to inadequate training in the undergraduate medical curriculum. Nationwide and program surveys reveal a lack of structured teaching of physical examination skills in the clinical clerkships, residencies, and fellowships, suggesting the introductory physical diagnosis course in the 1st or 2nd year of the undergraduate medical curriculum is the only formalized instruction in physical examination skills most physicians ever receive.[9–16] Rates of errors are consistent from the 3rd-year medical student to the attending physician level.[3–5] The Canadian national health services database reveals a strong association between some typical errors in medical practice and the physician's medical school.[17]

Purposes

The specific aims of this study were to use a standardized assessment of clinical skills performance to investigate whether systematic patterns of errors on the cardiovascular physical examination are manifest at the level of the 3rd year of undergraduate medical training and to begin to explore the bases for errors. We chose to study the cardiovascular physical examination because the cardiovascular component maneuvers have the highest error rates in published reports, and students rank listening to heart sounds as the second most challenging maneuver of the comprehensive physical examination.[5,7,9–13]

Methods

We analyzed the inclusion or omission of pertinent maneuvers on the cardiovascular physical examination as assessed by a Standardized Patient Examination checklist for 2,038 medical students from eight schools: Albert Einstein College of Medicine, Columbia University College of Physicians and Surgeons, Cornell University Medical School, Mount Sinai School of Medicine, New York Medical College, New York University School of Medicine, State University of New York—Health Sciences Center at Brooklyn, and State University of New York—Health Sciences Center at Stony Brook. Medical students completing the 3rd year in 1995 or 1996 at these schools participated in a full-day Standardized Patient Examination (SPE) at the Morchand Center for Clinical Competence at the Mount Sinai School of Medicine.[18,19] The SPE is designed to assess several aspects of clinical competence including the ability to take a focused history, perform a focused physical examination, and communicate effectively with patients. The examination includes seven cases developed by an expert panel of clinical faculty of the participating schools to portray a representative mix of problems commonly encountered in internal medicine, pediatrics, surgery, gynecology, and psychiatry. Twenty minutes are allotted for each student–standardized patient (SP) encounter. In the postencounter period, the SP completes checklists, scoring the student's performance of the focused history, the focused physical, and communication behaviors. The student–SP encounter also is videotaped to support a one-on-one critical observation and feedback session of the student with a clinician–educator.

As part of this assessment, each student examines a slightly obese, elderly male SP who presents with a chief complaint of chest pain, and associated symptoms and risk factors consistent with an acute myocardial infarction. The physical examination checklist consists of a set of 10 pertinent component maneuvers of the cardiovascular physical examination. Specifically, these maneuvers are considered by an expert panel of clinical faculty of the participating medical schools to be salient to the diagnostic assessment of important precipitants of chest pain (e.g., abnormal heart rate or rhythm, hypotension, hypertension, obstructive lung disease, pneumonia, pulmonary embolism, aortic dissection, or critical valvular heart disease) and important prognostic indicators that may alter treatment decisions (e.g., heart rate and rhythm, adequacy of blood pressure control, adequacy of peripheral perfusion, presence of congestive heart failure, critical valvular heart disease, or new heart murmur). The developers of the case and the larger medical faculty agree that by the end of the 3rd year of medical school, all students should know the pertinence (the "why") and the technique (the "how") of applying each of these component maneuvers—especially in the context of the classic presentation of an acute myocardial infarction.

Because the checklist data only provide information about the inclusion or omission of the component maneuvers, we performed further sampling from the main cohort and a cohort of 4th-year students to test subhypotheses on the reasons for omissions. We reviewed the history checklist scores for the cues for the most frequently omitted maneuvers to ascertain whether missed historical cues could explain the omissions. We critically observed 32 videotapes, a 25% sample randomly selected from the total of 126 students at Columbia University College of Physicians and Surgeons who took the examination in 1995, to determine whether time constraint of the examination or performance anxiety might have prevented students from completing the physical examination. To evaluate other contextual variables (e.g., omission of maneu-

vers for which students believed there would be no physical findings in an "actor"–patient), we devised a paper case with the identical presentation of the SP case. We randomly selected a 20% sample (n = 31) of medical students in the spring semester of their 4th year at Columbia in 1998, and asked them to assess the case and list up to 12 maneuvers that would be important to perform in the focused physical examination of the patient. In addition, we asked the students to explain the pertinence or utility of each maneuver in detecting physical findings and discriminating clinically important differences. To investigate hypothesized limitations in our teaching methods as bases for the observed error patterns, after the same group of students had handed in the analysis of the paper case, we asked them to consider each of the 10 maneuvers in Table 1 and rate their individual confidence in their ability to (a) technically perform the steps and sequencing of the maneuver, (b) differentiate normal from abnormal physical findings detected with the maneuver, and (c) discriminate different types of abnormal findings detected with the maneuver.

Results

The proportion of students who performed each of the 10 pertinent component maneuvers is presented in Table 1. It can be seen that three of these maneuvers were neglected quite reliably: 95% of the students failed to take the blood pressure in both arms, 63% failed to listen to the heart at the base and apex while the patient was sitting and leaning forward, and 75% failed to percuss the back from side to side at two levels at least. Other maneuvers were performed with a much higher likelihood: Students did take the blood pressure in one arm (68%), did listen to the heart at the base and apex while the patient was lying down (76%), took the pulse in both feet and ankles (62%), palpated the right upper quadrant of the abdomen (62%), and palpated for pretibial edema (65%).

These patterns of errors, from maneuvers most to least likely to be included, for Columbia students are virtually identical to those of the entire cohort of students from the consortium of eight New York medical schools (Table 1). Segmenting the data by school reveals that in all eight schools, the maneuver most likely to be omitted is taking the blood pressure in both arms (range of inclusion = 2–8%); in all eight schools, the next most likely maneuver to be omitted is percussing the lung at two levels (20–44%); and in seven of the eight schools, the third most likely maneuver to be omitted is listening to the heart while the patient is sitting and leaning forward (43–67%). Additional analyses show the same essential patterns of omissions and inclusions for the cohorts of students who participated in 1995 and 1996.

The history checklist scores for the cues for the three most frequently omitted maneuvers are consistently high for the Columbia students (79–100%) and the eight-school cohort (62–100%), suggesting that missed historical cues cannot explain the omissions. On critical observation of a 25% sample of videotapes of the students at Columbia who were tested in 1995, we see that students consistently closed this SP encounter well in advance of a buzzer indicating 5 min remaining. Students also appeared reasonably focused and not unduly anxious. As seen in Figure 1, virtually the same patterns of inclusions and omissions are obtained on the paper case as on the SP case, providing further evidence that context specificity (i.e., a constraint of the SPE assessment method) is not the reason for the observed deficiencies in performance.

Students' confidence ratings are summarized in Figure 2. Although students report being very confident in their ability to perform the maneuvers (range = 3.58–3.90, grand mean = 3.68), they are somewhat less confident in being able to tell normal from abnormal findings (range = 2.74–3.52, grand mean = 3.17), and even less confident in being able to discriminate different types of abnormal findings (range = 2.16–3.2, grand mean = 2.80). Across maneuvers, there is little

Table 1. *Inclusion of Pertinent Component Maneuvers of the Cardiovascular Physical Examination*

Checklist Cardiovascular Physical Examination Items	% All Students Performing Maneuver[a]	% Columbia Students Performing Maneuver[b]
1. Took blood pressure in other arm.	5	4
2. Percussed back from side to side (at two levels at least).	25	25
3. Listened to heart at apex and base while sitting and leaning forward.	37	36
4. Palpated abdomen in right upper quadrant.	62	73
5. Took pulse in both feet and ankles.	62	81
6. Took blood pressure in one arm.	65	78
7. Palpated for pretibial edema.	65	81
8. Listened to lungs in back (side to side, at least two levels).	67	67
9. Took pulse at one wrist.	68	77
10. Listened to heart at apex and base while supine.	76	80

[a]n = 2,038. [b]n = 269.

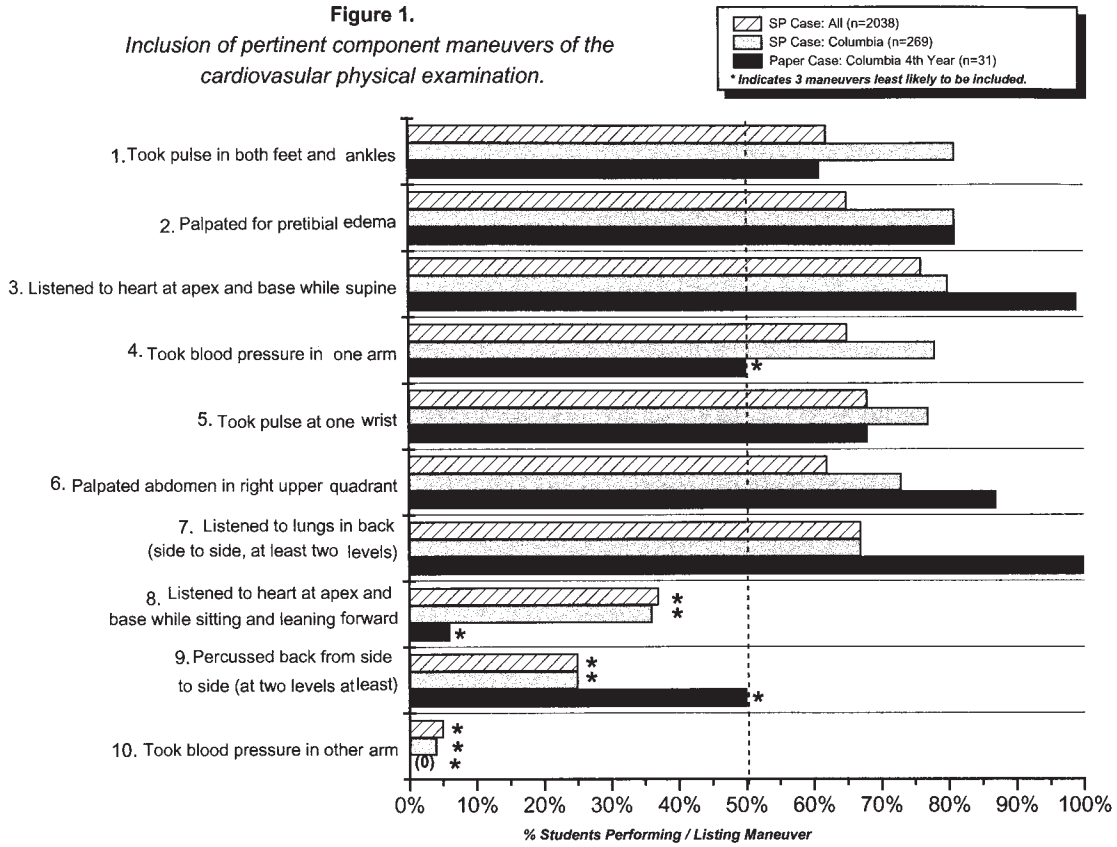

Figure 1. *Inclusion of pertinent component maneuvers of the cardiovascular physical examination.*

variation in confidence in the ability to perform the maneuver, greater variation in confidence in the ability to differentiate normal from abnormal findings, and greatest variation in confidence in the ability to discriminate different types of abnormal findings.

Conclusions

This methodological approach of studying performance error patterns based on SPE aggregate item data provides an informative window into systematic strengths and weaknesses in our undergraduate medical education programs for the teaching of physical diagnosis. The SPE data reveal a remarkable consistency of error patterns on three core components of the cardiovascular physical examination—across schools, years, and testing contexts—suggesting a fundamental inadequacy in our undergraduate clinical training programs. All medical schools provide rigorous formal instruction in physical diagnosis typically involving demonstrations, practice opportunities, text and video references, and checklists to follow. These methods are clearly insufficient. These findings call for a careful analysis of the bases for these errors. This resounding consistency of findings supports the conclusion that these omissions in the cardiovascular physical examination are not a mere artifact of the SPE contextual variables such as time constraint, "actor"–patient, particular case content, or performance anxiety. We encourage continued use of this approach to explore and analyze systematic deficits in our teaching of other parts of the physical examination. The exposure of such systematic yet unanticipated patterns of errors calls for reexamination of our teaching practices, provides a focus for our teaching efforts, and can motivate research to help elucidate the underlying causes for such widespread errors.

Our data demonstrate that missed historical cues are not the bases for the errors observed on the cardiovascular physical examination. We suggest that the reason for omitting a pertinent maneuver may differ as a function of the maneuver. Our current practices for teaching and reinforcing physical examination skills may benefit from a careful task analysis that delineates four basic learning tasks associated with each maneuver. Each maneuver requires (a) knowledge and recall of the pertinence of the maneuver, (b) mastery of basic technical (sensorimotor) skills, (c) the ability to perceptually discriminate (detect and interpret) normal from abnormal physical findings, and (d) the ability to differentiate different types of abnormal findings. Our student perception survey suggests that the majority of students can adequately execute the technique (the

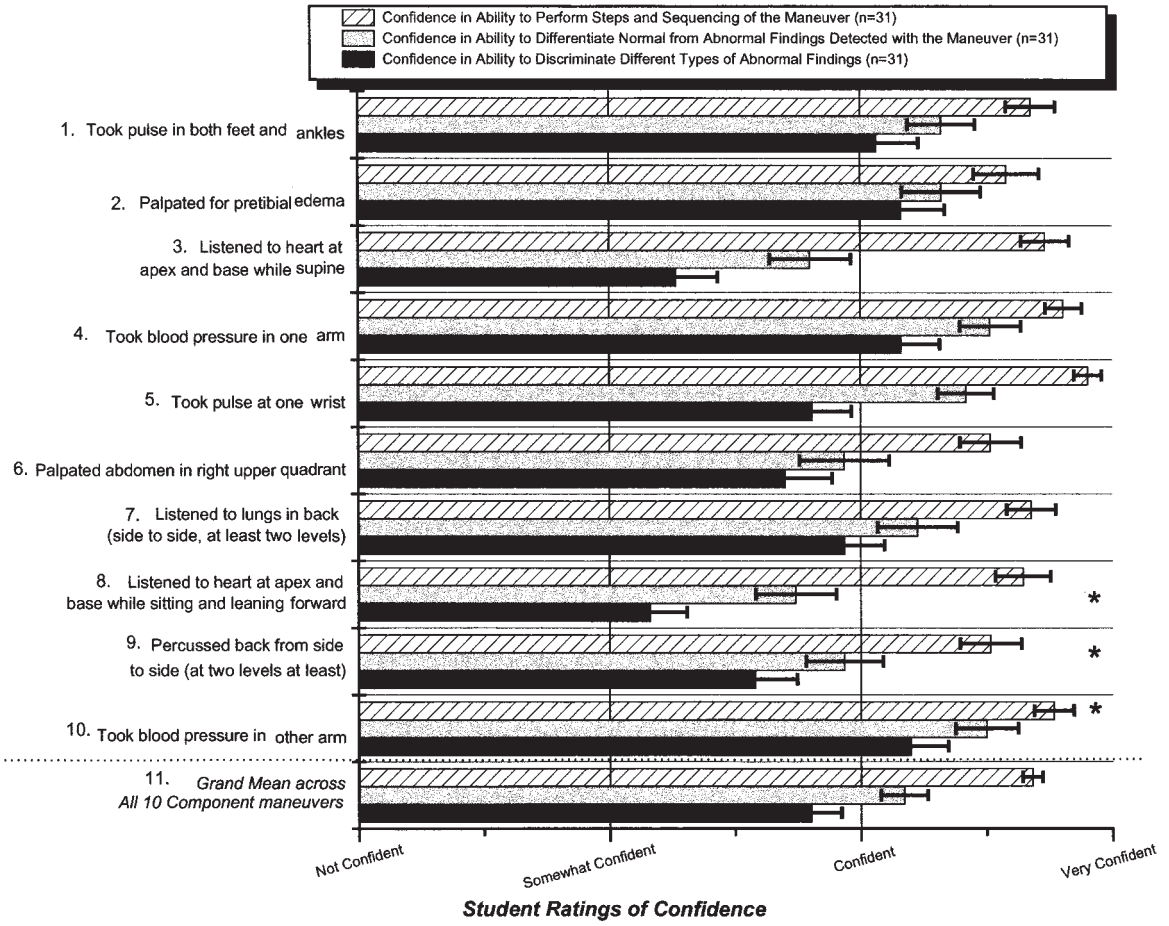

Figure 2. *Confidence in ability to perform and extract clinical information with maneuvers of the cardiovascular physical examination.*

"how") of most of the component maneuvers but have difficulty discriminating normal from abnormal findings, and even greater difficulty discriminating different types of abnormal heart sounds. Our paper case analysis and discussions with students provide additional evidence that other maneuvers are omitted because students fail to understand the pertinence or salience of the maneuver to the differential diagnosis or prognosis (the "why").

In the physical examination, any maneuver may be omitted because the student has missed a medical historical cue or not learned how to perform the skill. Alternatively, a maneuver may be omitted because of a failure to understand the significance of the potential findings, or because of a lack of confidence in one's ability to interpret findings based on the maneuver. In this study, it appears that students omit taking the blood pressure in both arms because, in this situation, they fail to consider the pertinence of the maneuver to the discrimination of aortic dissection from other causes in the differential diagnosis of chest pain (inert knowledge); they take the blood pressure with technical finesse in one arm, and are confident in their ability to detect and interpret abnormal blood pressures. Students also may omit listening to the heart with the patient in different positions, because they fail to understand the relevance to the discrimination of a new mitral regurgitation murmur from critical aortic stenosis. However, our data suggest that a confounding reason may be related to their relatively low confidence in discriminating normal from abnormal heart sounds, and even lower confidence in discriminating different types of abnormal heart sounds (Why bother if you cannot interpret your findings?). Percussing the lung at two levels on the back of the patient may be omitted because of a lack of understanding of the pertinence to the discrimination of normal lung from a pleural effusion at the lung bases (a sign of a clinically important prognostic indicator—congestive heart failure). Our data suggest percussion also may be omitted because students have relatively low confidence in their ability to perform this challenging maneuver (Why bother if your motor skills are not adequately developed to elicit a finding?).

It is also possible that the medical profession's increased reliance on diagnostic technology and poor role modeling is undermining the teaching and learning of basic physical diagnosis skills, and undermining students' belief in the primacy of their bedside diagnostic findings. This possibility is concerning, because the results of even a highly sensitive diagnostic test are

best interpreted in a clinical context. The diagnostic reasoning process begins with the history and physical examination. Each pertinent item of the history or physical examination contributes to the interpretation of the test result by increasing or decreasing the estimate of the probability of the disease before the test is run. In the absence of reliable history and physical examination findings on which to base estimates of the pretest probability of disease, a medical student or physician has greater diagnostic uncertainty. The best he or she can do is to assume the highest and lowest possible pretest probabilities, and consider if this range of estimates will change clinical decision making.

Physical examination skills are fundamental to the diagnostic reasoning process and to the selection and interpretation of diagnostic tests. It is therefore important to more completely characterize the learning tasks of the physical examination to inform the instructional design of our teaching of cardiovascular physical examination skills. A starting point might be to adopt approaches that more effectively and fully integrate the teaching of physical examination skills, basic and clinical anatomy, and physiology.[22] By strengthening these links, students may be more likely to understand and recall the appropriate pertinence, sequencing, and positioning of each maneuver. Moreover, we should carefully study the extent to which clinical role models and bedside teaching reinforce or undermine basic physical examinations skills. The study of performance error patterns based on SPE aggregate checklist data has the potential to identify similar systematic deficiencies in the teaching of other components of the physical examination.

References

1. Bordage G. Why did I miss the diagnosis? Some cognitive explanations and educational implications. *Academic Medicine* 1999;74:S138–43.
2. Leape LL. Error in medicine. *Journal of the American Medical Association* 1994;272:1851–7.
3. Wray NP, Friedland JA. Detection and correction of house staff error in physical diagnosis. *Journal of the American Medical Association* 1983;249:1035–7.
4. Goetzl EJ, Cohen P, Downing E, Erat K, Jessiman AG. Quality of diagnostic examinations in a university hospital outpatient clinic. *Annals of Internal Medicine* 1973;78:481–9.
5. Mangione S, Nieman LZ. Cardiac auscultatory skills of internal medicine and family practice trainees: A comparison of diagnostic proficiency. *Journal of the American Medical Association* 1997;278:717–22.
6. Mangione S, Peitzman SJ. Physical diagnosis in the 1990s: Art or artifact? *JGIM* 1996;11:490–3.
7. Johnson JE, Carpenter JL. Medical house staff performance in physical examination. *Archives of Internal Medicine* 1986;146:937–41.
8. St. Clair EW, Oddone EZ, Waugh RA, Corey GR, Fuessner JR. Assessing housestaff diagnostic skills using a cardiology patient simulator. *Annals of Internal Medicine* 1992;117:751–6.
9. Walters CA. The need for office-based precepting. *Hospital Practice* 1993;48:7–13.
10. Walters CA. *New directions in professional training.* Workshop presented to the Fund for Improvement of Post-Secondary Education XXV Project Directors' Meeting, Washington, DC, November 1, 1997.
11. Walters CA, Schmidt HJS. *New directions in clinical skills teaching.* Grand Rounds, NY: Albert Einstein College of Medicine, Department of Pediatrics, 1997.
12. Walters CA, Ortiz-Neu CI. *New directions in teaching physical diagnosis skills: A faculty development workshop.* Workshop held at the 1998 Annual Session of the Society of General Internal Medicine, Chicago, May 1998.
13. Walters CA, Ortiz-Neu CI. New directions in teaching physical diagnosis skills: A faculty development workshop. Workshop held at the 1999 annual session of the Society of General Internal Medicine, San Francisco, May 1999.
14. Dunnington G. Teaching physical diagnosis in the surgical clerkship. *Teaching and Learning in Medicine* 1992;4:110–4.
15. Ortiz-Neu CI. *Structured teaching of the Physical Examination in Clinical Clerkships: A survey of clerkship directors at the College of Physicians and Surgeons.* Unpublished raw data, Spring 1998.
16. Mangione S, Nieman LZ, Gracely E, Kaye D. The teaching and practice of cardiac auscultation during internal medicine and cardiology training: A nationwide survey. *Annals of Internal Medicine* 1993;119:47–57.
17. Tamblyn R. Outcomes in medical education. *Advances in Health Sciences Education* 1999;4:9–25.
18. Swartz MH, Colliver JA. Using standardized patients for assessing clinical performance: An overview. *The Mount Sinai Journal of Medicine* 1996;63:241–9.
19. Swartz MH, Colliver JA, Bardes CL, Charon R, Fried ED, Moroff S. The validity of standardized patient assessment using faculty global ratings as the gold standard criterion. *Academic Medicine* 1997;72:619–26.
20. Jaeschke R, Gordon H, Guyatt G, Sackett DL (for the Evidence-Based Medicine Working Group). III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? *Journal of the American Medical Association* 1994;271:703–7.
21. Katz MA. A probability graph describing the predictive value of a highly sensitive diagnostic test. *New England Journal of Medicine* 1974;291:1115–6.
22. Schmidt H. Integrating the teaching of basic sciences, clinical sciences, and biopsychosocial issues. *Academic Medicine* 1998;73:S24–31.

Received 23 November 1998
Final revision received 22 January 2001