A Method for Evaluating Patient Care and Auditing Skills of Family Practice Residents

Robert L. Kane, MD, Elbert H. Leigh, MD, David W. Feigal, MD, David N. Sundwall, MD, and Martha A. Farney, RN
Salt Lake City, Utah

A system of chart review is presented which develops both auditing skills and patient care skills. The model suggested is a flexible one, applicable at various levels of training in both inpatient and outpatient settings. Its application to family practice resident training is shown by comparing the performances of residents at the beginning and end of a demonstration period.

The pressures to develop meaningful measures of quality of care and to implement these in an effective system of peer review are already apparent.[1] The need to insure the quality of health care delivered implies a responsibility to teach health professional students the skills of medical auditing. However, such a learning experience can have an even wider effect. The audit may be a valuable learning tool both for measuring performance and for teaching patient care techniques at the same time. The audit experience represents a potentially nonthreatening and effective means to allow a student to work through problem-solving experiences by reviewing the decisions made by another practitioner.

The schema to be described permits the separation of knowledge of how to care for the patient (as represented by auditing ability) and actual patient care behavior (as represented by the patient care score). This audit program is patterned after a previously described operational audit system which utilizes the Weed problem-oriented medical record.[2] The audit relies on peer judgments of the degree to which each of the four basic steps in the medical care process (I. data base collection, II. problem identification, III. plan formulation, and IV. actions taken) has been accomplished.

By having various individuals audit the same chart, different types of information can be obtained. The situation diagrammed in Figure 1 illustrates the various applications possible. If we assume that student S_1 has provided the care recorded in the chart and that this chart is then audited by student S_2 and the instructor, the following types of information are potentially available from the experience:

1. The patient care skills of student S_1 could be shown by comparing the patient care rendered by that student with the results of an instructor's audit.
2. The auditing skills of student S_2 can be obtained by comparing the audit by student S_2 with that of the same instructor. The instructor audit is considered to be a measure of truth.

This system can be utilized in the training of residents by presenting data about both their patient care and auditing skills. In order to be effective in modifying the resident's behavior, adequate feedback on his performance must be provided. Further criteria of a useful system were seen to be (1) minimum time requirements for auditing, (2) no increased burden for those providing care, (3) separate assessment of charting and patient care performance, and (4) applicability to ambulatory patients still receiving care.

Methods

With these concepts in mind, the audit program for the Family Practice Center at the University of Utah was instituted in fall, 1973. The results described here represent the first six months of resident audit experience from November, 1973, through May, 1974. An audit form* was designed to assess three different areas of resident performance: charting skills, patient care skills, and overall patient care. This form is completed independently by both the instructor and the resident who audits the chart.

The first section deals with charting skills, including the presence of a problem list, signatures and dates for each visit, problem-oriented progress notes, and a complete data base, ie, past history, review of systems, family history, social history, physical examination data, and lab results filed in order.

The second section is concerned with items reflecting patient care skills. These include a problem-appropriate data base, recognition of all problems presented, adequacy or excess of laboratory studies performed, recognition of abnormal lab study results, specificity and appropriateness of medications ordered, patient instruction provided, description of procedures, and adequacy of consultation and follow-up.

Finally, for the third section, a number score from zero for "unacceptable" to four for "excellent" is independently assigned by the auditor after he has reviewed the overall care the patient received.

The first two sections are scored by checking one of three options — yes, the item is present and adequate; no, the item is not present; or the item is

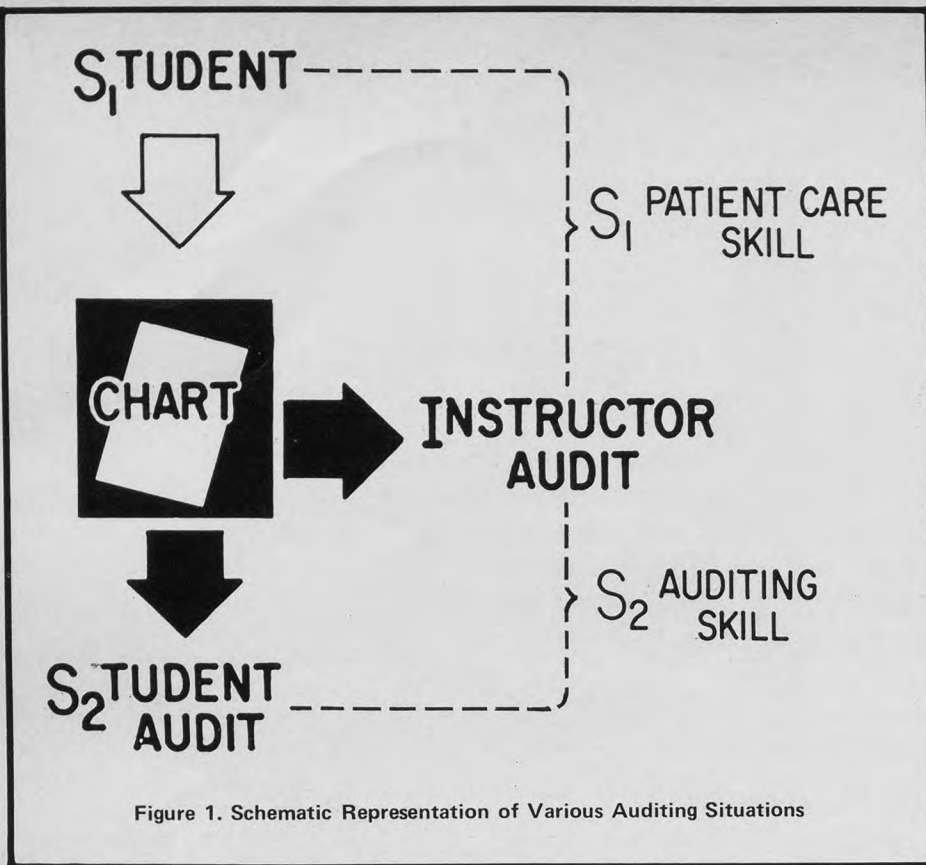*Copies of these forms are available from the authors.

Figure 1. Schematic Representation of Various Auditing Situations

not applicable. Space is provided for auditor comment after each item and for general comments at the end.

The auditing procedure was as follows: a sample of three charts was pulled each day from patients seen approximately one week earlier. The delay allowed for the transcription of clinic notes and the insertion of laboratory data. An attempt was made to pull one chart for each resident who worked in the clinic that day. However, since not every resident at work on any one day will see patients, this was not always possible.

The chart, together with an audit form, was first given to a resident who did not provide the patient care, but who was assigned to the clinic on the day the chart was pulled. He was requested to complete the audit and return the form within 24 hours. Following this, the second audit form and the same chart were given to a faculty member for his independent audit. Each resident and faculty member audited approximately two charts per week.

The completed audit forms were then duplicated with one copy going to the resident providing the care, one to data processing, and a third to the program's audit committee. Monthly, members of this committee summa-

rized* the records of each resident giving care and entered these data in his personal evaluation folder. Where instances of unsatisfactory work were noted, the chart numbers were recorded and the resident was instructed to review his handling of the case. The system thus provided two levels of feedback to the residents about their patient care skills as assessed by both faculty and peers.

The data processing copy of the audit form was coded and key punched for computer analysis. This analysis provides individual performance profiles for each resident in terms of both patient care and auditing. Patient care skills are measured on the basis of the assessments of the faculty auditors. Auditing scores are, in effect, measures of concordance between resident and faculty auditors on a case-by-case basis.

In order to assess the effects of this program on resident performance, the first 325 audits performed, covering a five-month span, were split into halves and compared for the resident group as a whole. Unfortunately, no control group was available to distinguish possible secular trends from effects of feedback; statistical signifi-

---

*Copies of these forms are available from the authors.

cance was not assessed under these conditions.

Results

Figure 2 shows the profile of patient care scores for the residents divided into first and second halves of 161 and 164 cases, respectively. The items shown are those appearing on the audit form described above. The scores are presented in one section representing patient management skills and a second representing charting skills. The results shown here are derived from the faculty audit. The points plotted on the graph represent the percentage of charts judged adequate.

Some improvement in performance from the first to the second time period is displayed for most items involving patient management and charting skills. The exceptions are items concerned with consultation and follow-up. Greater performance improvement appears to have occurred in charting behavior than in patient care skills. When the item data are collapsed, analysis reveals that the average charting skills score rose from 77 percent to 87 percent adequate. Average patient care scores, on the other hand, rose only from 89 percent to 90 percent adequate.

Auditing scores are similarly displayed and represent the percentage of agreement between each pair of resident and faculty auditors for the first and second periods. Results again indicate greater improvement in the assessment of charting as compared to patient care skills. Scores for auditor concordance in judging charting skills originally averaged 70 percent and rose to an average of 77 percent. Scores for auditor agreement on patient management skills remained almost unchanged: 65 percent during the first period and 66 percent during the second (Figure 3).

Table 1 displays the comparison of overall patient care assessment scores given by faculty and resident auditors during both periods. These tables, constructed by the computer, compare the percentage of agreement between resident and faculty auditors for each of the five assessment categories: "unacceptable" through "excellent." The center diagonal line gives the percentage of exact agreement between the two. The diagonal columns above and
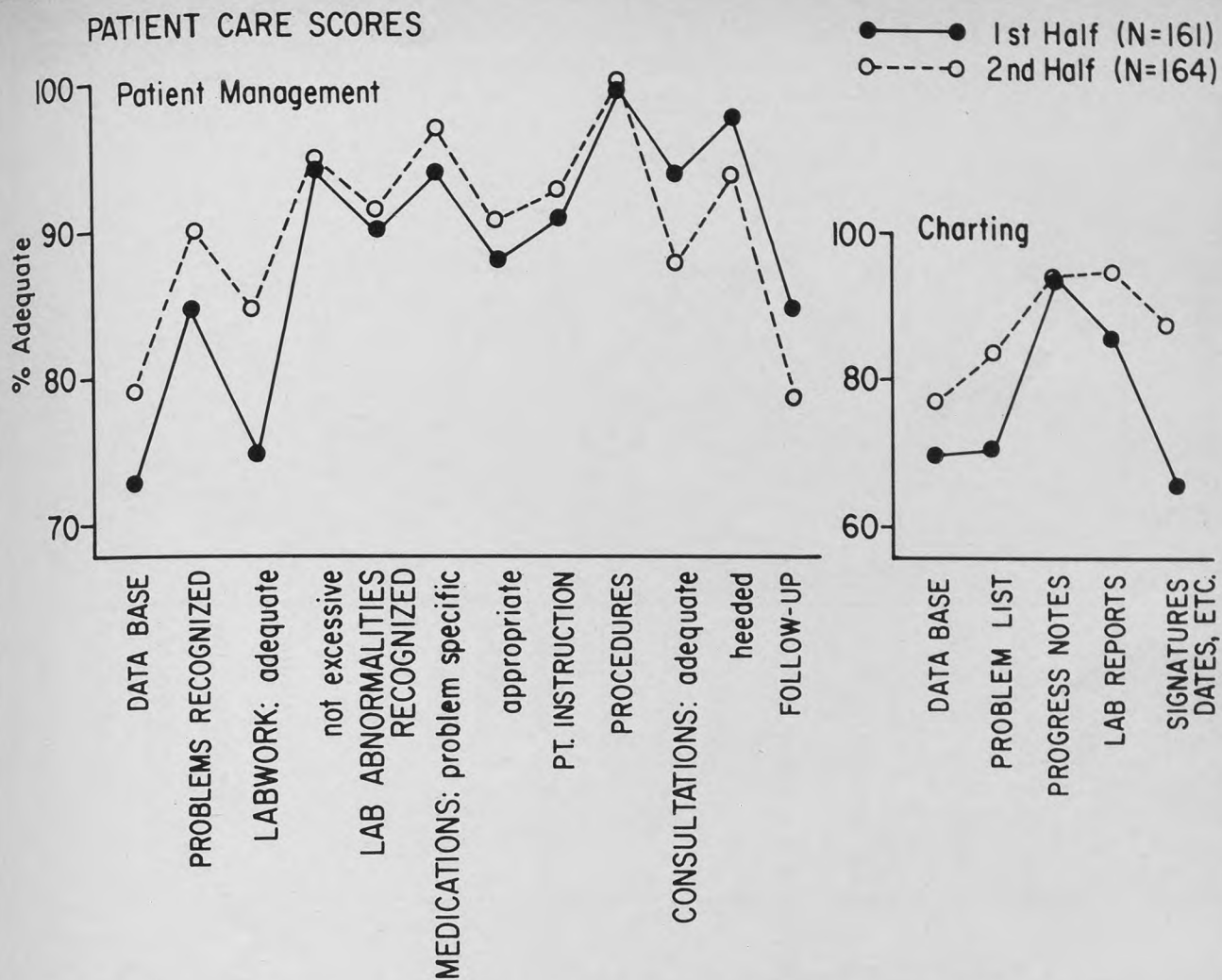
Figure 2. Change in Patient Care Scores Comparing the First 161 Audits with the Second Set of 164 Audits

below the center line represent agreement within one category.

In the first three-month period, exact agreement between faculty and resident auditors occurred for 34.7 percent of the charts assessed. Comparing auditor agreement across contiguous categories, such as "poor" to "satisfactory," or "satisfactory" to "good," reveals 75.7 percent agreement within neighboring categories. During the second three-month period, exact agreement dropped to 32.8 percent of the audits, while agreement within contiguous categories rose to 78.3 percent.

These data were analyzed alternatively in terms of α and β errors. An α error would be the number of cases which the faculty rated as "unacceptable" or "poor," but which the auditing resident rated "satisfactory" or better. Similarly, a β error would be the number of cases judged by the fac-

ulty as "good" or "excellent" but rated by the resident as "satisfactory" or poorer. The α error rates were .956 before and .771 after; the β errors fell from .281 to .134. This would suggest that the residents were improving in their ability to detect unacceptable work but still were missing large numbers of cases. Conversely, they correctly identified most of the good care and continued to improve.

Discussion

In reviewing alternative methods for assessing quality of care, Brook[3] divides techniques for process assessment into those which use explicit criteria and those using an implicit approach in which a competent judge rates the care received by the patient. The advantage of the criteria method has generally been a higher level of inter-rater reliability. However, Brook

concludes that either method provides reasonable reliability for 100 or more patients with a given disease.

The system described here was designed to permit the maximum flexibility required in reviewing cases as diverse as those found in family practice. With this purpose in mind, the method chosen uses elements of both the implicit and explicit approach. We would term it "structured implicit." By this is meant a controlled choice of items to be assessed, but with the ultimate judgment of the adequacy of care left up to the reviewer who considers the total context of the patient in making his judgment.

The results of this study suggest that a resident's charting is more responsive to change than is his patient care or his auditing. The relationship between charting and patient care remains cloudy. Although academic

AUDITING SCORES

Patient Management

Charting

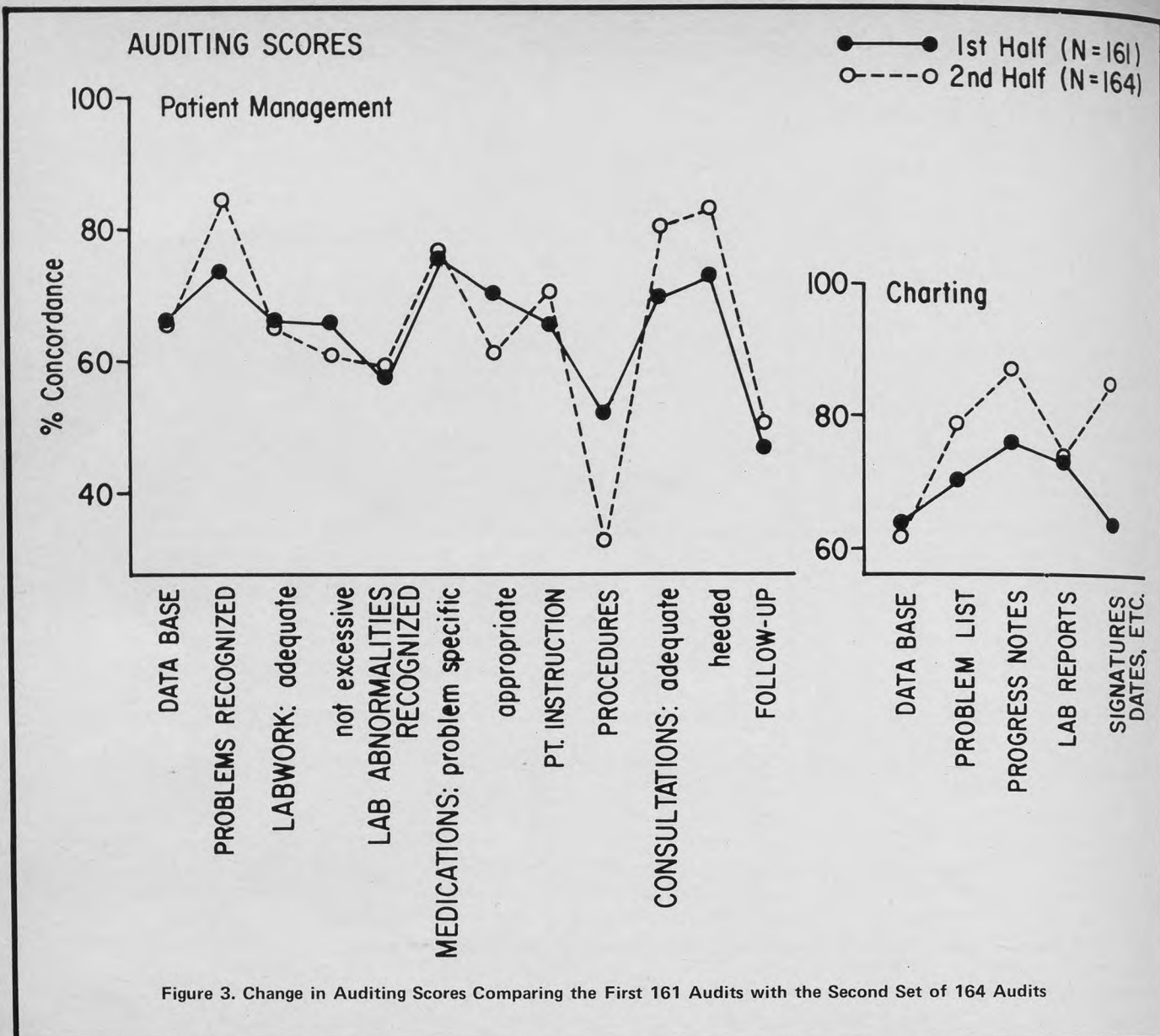1st Half (N=161)
2nd Half (N=164)

Figure 3. Change in Auditing Scores Comparing the First 161 Audits with the Second Set of 164 Audits

training values the chart highly as a patient care tool, studies are not consistent in demonstrating a positive relationship between care and charting. Fessel and Van Brunt[4] failed to find a correlation between charting and outcomes of care for myocardial infarctions or appendicitis. However, using techniques of process assessment, Lyons and Payne[5] suggest "that measures of good medical recording performance and good medical care performance are related, that the overall relationship is not at all perfect, and that the reliabilities of the measures may attenuate the degree of obtained relationships." The differences in these two sets of findings may lie in the lack of relationship frequently found between the process and outcome of care.

Whether or not better charting pro-

duces better care, several aspects of the data presented suggest that feedback to the resident produces beneficial changes in his behavior. The changes in patient care scores were more dramatic than changes in auditing scores. The former were the subject of feedback while the latter were not. The changes in charting were greater than those in patient management. Informal observations confirm the impression that the residents became quickly aware that their charts were being audited and set about to improve their recording. This type of Hawthorne effect, if it produces the desired results, may in itself be a worthwhile byproduct in a training program.

To a greater extent than expected, the residents complied with the audit program, completing their audits on

time. The audits for each resident provided clear documentation of improving, or occasionally worsening, performance which proved extremely useful in counseling the resident.

The lack of a control group has already been noted. While it seems unlikely that a secular trend could account for the full extent of the change found, one cannot exclude the possibility of a maturation effect. However, the majority of residents participating were in their second year of training at a point where their patient care skills should have developed.

The assessment of auditing skills in this model depends on the assumption of the faculty's ability to consistently judge the quality of care. To test this, a subsample of 50 charts was independently reviewed by a second faculty auditor. The level of absolute agree-

Table 1. Congruence of Auditing Scores between Faculty and Residents

FIRST HALF

FACULTY RATING

RESIDENT RATING	Unacceptable	Poor	Satisfactory	Good	Excellent	
Unacceptable	0	0	1.9	2.5	0	
Poor	0	.6	1.2	1.2	.6	
Satisfactory	1.9	1.2	4.3	10.6	1.9	
Good	1.9	7.5	14.9	26.7	7.5	19.3%
Excellent	.6	0	4.3	5.6	3.1	34.7%
						21.7%
TOTALS	4.4%	9.3%	26.6%	46.6%	13.1%	

SECOND HALF

FACULTY RATING

RESIDENT RATING	Unacceptable	Poor	Satisfactory	Good	Excellent	
Unacceptable	1.2	.6	0	1.2	0	
Poor	1.2	.6	1.8	3.7	0	
Satisfactory	1.8	2.4	2.4	9.1	3.0	
Good	1.2	4.9	7.9	26.2	7.9	19.4%
Excellent	.6	1.2	3.7	14.6	2.4	32.8%
						26.1%
TOTALS	6.0%	9.7%	15.8%	54.8%	13.3%	

$$S_2 \text{ audit vs. I audit} = S_2 \text{ auditing skill}$$
$$S_1 \text{ care vs. I audit} = S_1 \text{ patient care skill}$$
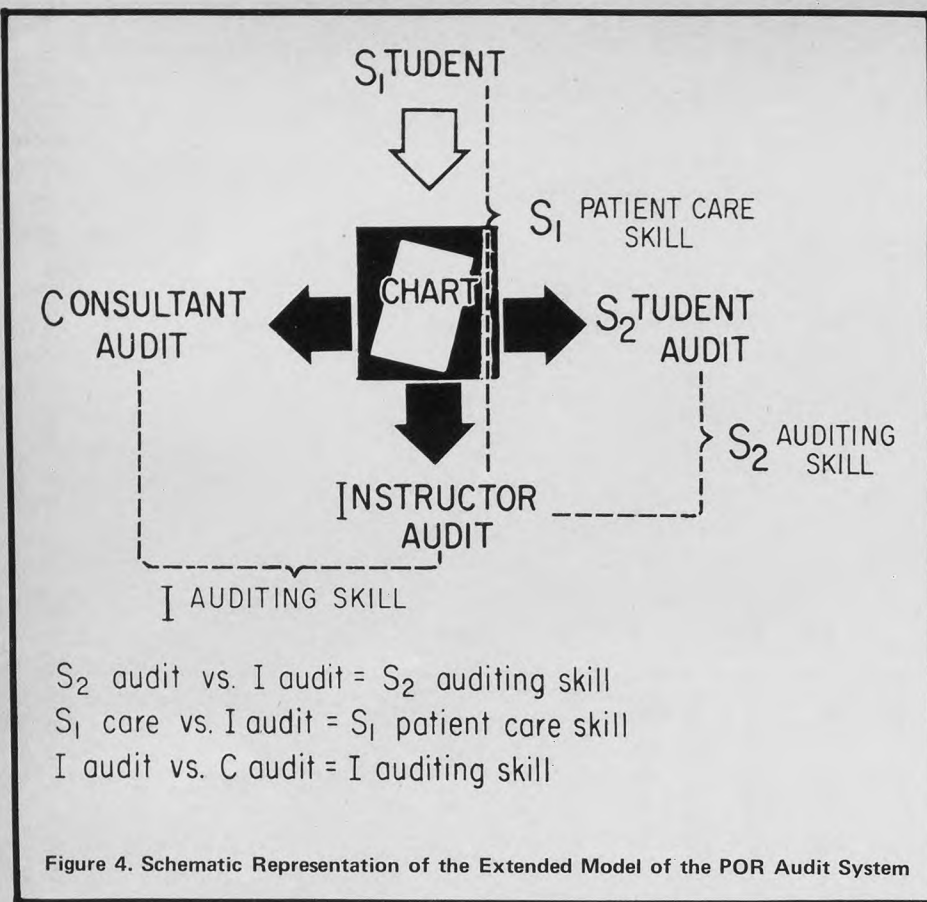$$I \text{ audit vs. C audit} = I \text{ auditing skill}$$

Figure 4. Schematic Representation of the Extended Model of the POR Audit System

ment was 34.7 percent, with another 36.7 percent agreeing within one category. This total of 71.4 percent is actually less than that between faculty and resident auditors. To treat this problem of low inter-rater reliability, the POR audit system can be expanded further by adding a consultant audit as shown in Figure 4. This would permit an evaluation of the instructor's auditing skills by comparing his audit to that performed by the consultant expert, and should result in improved inter-rater reliability.

Such a system has the potential for both student and faculty training. As noted in the diagram, one can examine the behavior of instructors as well as students. In terms of resident training, one can analyze a year's activities by type of case to generate information about behavior (as reflected in the patient care score) and about knowledge (as reflected in the auditing score). These categories can be still further subdivided by the scores for each of the four steps in the medical care process (discussed on the first page of this article) as well as an overall score. This information could be displayed as follows:

By type of case

BEHAVIOR
(Patient care score)

| I | II | III | IV | Total |

KNOWLEDGE
(Auditing Score)

| I | II | III | IV | Total |

The discrepancy between a student's behavioral scores and knowledge scores represents the gap between what he knows and what he does. This discrepancy could be fed back to the student to encourage him to reassess his performance and search for potential changes that might improve it. Another variation permitted by the system would be to encourage the student to audit his own charts and thus even more directly measure the difference between his knowledge and his behavior.

The approach described here represents a model which may be generalized and applied at different levels of training. For example, the student role can be filled by undergraduate medical students and the instructor role by housestaff. Or, as described here, the housestaff could serve as the

students with the faculty playing the instructors. This latter design fits very neatly into already developing trends in residency training for ongoing certification of competence by the individual programs.[6,7]

Such an approach to student training and evaluation is particularly compatible with the current surge in family practice training programs. The two levels of supervisory responsibility offer a meaningful role for both family practitioners and specialists in the education of primary care physicians. Finally, this technique seems highly compatible with the future of medical education as outlined by the National Board of Medical Examiners.[8]

This technique of chart auditing provides both practice experience toward developing peer review skills that will become increasingly important in medical practice and a learning opportunity to review the management of a wider variety of cases than otherwise might fall within the range of the trainee's patient contacts. In this sense, the chart review system described here represents a relatively inexpensive simulation exercise with feedback from instructors. Finally, it is a flexible model which can be applied to both inpatient and outpatient care.

References
1. Welch CE: Professional standards review organizations — problems and prospects. N Engl J Med 289:291-295, 1973
2. Putsch RW III, Humphrey J, Schuster JS: Quality care problem orientation medical audit. In Hurst JW, Walker HK (eds): The Problem Oriented System. New York, Medcom Press, 1972
3. Brook RH: Critical issues in the assessment of quality of care and the relationship to HMO's. J Med Educ 48 (suppl): 114-134, 1973
4. Fessel WJ, Van Brunt EE: Quality of care and the medical record. N Engl J Med 283:904-907, 1972
5. Lyons TF, Payne BC: The relationship of physicians' medical recording performance to their medical care performance. Med Care 12:463-469, 1974
6. Petersdorf RG, Beck JC: The new procedure for evaluating the clinical competence of candidates to be certified by the American Board of Internal Medicine. Ann Intern Med 76:491-496, 1972
7. Burg FD, Wright FH: Evaluation of pediatric residents and their training programs. J Pediatr 80: 183-189, 1972
8. Evaluation in the continuum of medical education. Report of the Committee on Goals and Priorities of the National Board of Medical Examiners. Philadelphia, 1973