METHODS

Reliability of Chart Examination

Information was abstracted from patients' records by a medical-record librarian. To validate her accuracy two preparatory steps were taken: the librarian made a preliminary chart search for items on an audit list, and then requested that we furnish her with precise definitions of the items that she found to be ambiguous; and, abstracting from an additional set of records, she then found 625 pertinent entries. Reproducibility and reliability were established by comparison of her findings for these 625 items with our independent observations of the same items. When our findings, cross-checked and corrected where there was discrepancy, were used as the standard, the medical-record librarian was found to have made only 2.4 per cent errors.

Acute Appendicitis

For this condition, 150 charts were reviewed of patients seen during a one-year period, 1967-1968, at hospitals A, B and C. At each hospital a group of 50 charts was selected randomly from the total number of patients seen during the period under study with the histopathologically proved diagnosis of acute appendicitis. The charts from hospital A, which has a full house staff, were surveyed to obtain a list of recorded symptoms, signs and laboratory data relevant to the condition. This list was used to audit the records at all three hospitals.

Acute Myocardial Infarction

Fifty randomly selected records were studied of patients with acute myocardial infarction who were discharged alive from Hospital A during the period July 1, 1967 to June 30, 1968. All the data pertaining to the condition that had been recorded in the hospital chart were listed for each patient, and an audit list* assembled that contained the 44 items that had been documented for 75 per cent or more of the patients. This method of constructing the audit list ensured that most of the items on the list had been noted as either present or absent for each patient during the hospital episode. The list contained elements of the history (e.g., chest pain, dyspnea, previous myocardial infarction), physical examination (e.g., heart murmur, neck veins, rales) and a variety of tests (e.g., serum enzymes, electrocardiogram) relevant to the condition. A value of one was assigned to each of these elements. The corresponding outpatient records were studied for the following postinfarction occurrences (outcomes): duration of time before returning to work; angina pectoris, congestive heart failure or arrhythmias that were not recorded in the outpatient chart as present before the heart attack; new myocardial infarction; and death. The average follow-up period was 464 days.

The same 50 records were also examined for some of the items found by Hughes et al.,[1] in their 445 patients, to be related to the outcomes of survival or death after acute myocardial infarction. These items included references to white-cell count, blood pressure, congestive heart failure (for this we required four of five physical signs to have been noted as present or absent — distended neck veins, cardiac gallop rhythm, palpable liver, pulmonary rales and edema of the legs), diabetes (for this we examined the recording of blood sugar), and the histories of previous myocardial infarction and hypertension. Since the blood pressure, the white-cell count and the signs related to the presence or absence of heart failure had been recorded for all our patients, the recording of these aspects of process could not be related to any of the outcomes. No fresh arrhythmia occurred during the follow-up period in any patient, so that this outcome also could not be related to any recorded element of process.

In a further study, also conducted at Hospital A, we requested three cardiologists to compile a traditional audit list for acute myocardial infarction; the cardiologists considered the items that they listed to be minimum requirements for good-quality medical care of this disorder. Using this list, we studied the medical records of a second group of 50 patients with uncomplicated acute myocardial infarction, and compared their audit scores with the scores obtained from a similar audit of records of a third group of 50 patients who died while in the hospital for care of their disease.

RESULTS

Acute Appendicitis

There was a disparity in the recorded data from hospital to hospital (Table 1) that was statistically significant for eight commonly sought symptoms or signs. However, the outcomes, defined as pathologically proved correct or incorrect diagnoses, were the same from hospital to hospital (Table 2).

The percentages in Table 2, which shows the diagnostic outcome at each hospital, are not just from the sample of 50 patients, but are derived from the complete year's experience with removed appendixes, in the pathology department's files at each hospital.

Acute Myocardial Infarction

There was no significant relation between the audit scores and any of the posthospital outcomes for which the records were examined (Table 3). Mean scores for the patients in whom complications developed were similar to those for patients in

*For more detailed information order NAPS Document 01677 from National Auxiliary Publications Service, c/o CCM Information Corporation, 866 3d Ave., New York, N. Y. 10022; remitting $2 for each microfiche-copy reproduction or $5 for each photocopy. Checks or money orders should be made payable to CCM Information Corporation — National Auxiliary Publications Service.

Table 1. Recorded Data for Patients with Histologically Confirmed Acute Appendicitis

Datum	Hospital* A %	B %	C %	Probability†
Anorexia	72	44	64	<0.05
Diarrhea	94	84	82	
Shift of abdominal pain to right lower quadrant	100	100	100	
Nausea or vomiting	100	100	100	
Right-lower-quadrant tenderness	100	100	100	
Rebound tenderness	100	96	94	
Referred rebound	82	26	36	<0.001
Guarding	96	52	88	<0.001
Rigidity	16	16	14	
Psoas sign	48	14	6	<0.001
Obturator sign	32	8	2	<0.001
Bowel sounds	98	94	80	<0.01
Masses	42	62	44	
Organ enlargement	30	26	36	<0.001
Rectum:				
Tender on right	68	50	28	<0.001
Other tenderness	12	6	8	
Normal	20	36	34	
Temperature elevation	100	100	100	
Leukocytosis:				
With left shift	88	90	84	
No shift	10	10	14	
Differential count not done	2		2	

*50 charts from each hospital.

†Results of chi-square analysis of the probability that differences in percentages among the 3 hospitals are due to chance.

whom complications did not develop. No relation was seen between the duration of time away from work and the audit score.

There were no significant differences (Table 4) in 11 of the 12 comparisons, when our patients with myocardial infarct were separated on the basis of (a) the presence or absence in the charts of the items of process found by Hughes et al.[1] to be prognostically important, and (b) the outcomes listed in Table 3.

The overall pattern of the results was the same for the audits that compared the records of 50 patients who survived acute uncomplicated myocardial infarction with those of 50 patients who died (Table 5). The frequency of recording was significantly different for only five of the 26 items on the list, as follows: a family history of arteriosclerotic heart disease or diabetes mellitus (or both) was recorded significantly less often in the charts of the patients who died. Electrocardiograms, before

Table 2. Appendectomy — Diagnosis Outcomes at Hospitals A, B and C, 1967-1968.

Hospital	No. of Patients	Pathologically Proved Acute Appendicitis (%)	No Pathological Findings (%)	Diagnosis Other than Appendicitis (%)
A	466	82.2	13.6	4.1
B	167	83.1	13.7	4.2
C	104	89.4	10.6	0

Table 3. Comparison of Audits of the Process of Care with Outcomes of Myocardial Infarction.*†

Outcome	No. of Patients	Mean Audit Score
New angina present	13	37.1
New angina not present	34	36.8
Data not recorded	3	37.7
New congestive heart failure present	4	39.5
New congestive heart failure not present	30	36.4
Data not recorded	16	37.3
New myocardial infarction present	5	36.6
New myocardial infarction not present	12	37.2
Data not recorded	33	36.8
Death from known or probable myocardial infarction	3	36.7
Patients did not die of myocardial infarction	47	36.9

*Maximum possible audit score = 44.

†Since new arrhythmia did not arise in any patient during the follow-up period, this outcome is not tabulated.

death or before discharge from the hospital, were also recorded significantly less often in the charts of the patients who died, whereas a serologic test for syphilis was recorded more often. Electrocardiograms were absent because the diagnosis was made post mortem in five of the patients, and, of

Table 4. Relation between Outcomes and Recording of Prognostic Indicators as Present or Absent in Myocardial Infarction.*

Outcome	No. of Patients	Patients with Diabetes Mellitus R†	Patients with Diabetes Mellitus NR‡	Patients with Myocardial Infarction R	Patients with Myocardial Infarction NR	Patients with Hypertension R	Patients with Hypertension NR
New angina:							
Present	13	4	9	5	8	10	3
Not present	34	18	16	17	17	22	12
Not recorded	3	1	2	2	1	1	2
New congestive heart failure:							
Present	4	3	1	2	2	2	2
Not present	30	13	17	13	17	20	10
Not recorded	16	7	9	9	7	11	5
New myocardial infarction:							
Present	5	3	2	1§	4	2	3
Not present	12	7	5	11	1	7	5
Not recorded	33	13	20	12	21	24	9
Death from known or probable myocardial infarction	3	1	2	3	0	0¶	3
Patient did not die of myocardial infarction	47	22	25	21	26	33	14

*Results analyzed by t test; unknown outcomes were assumed not to have occurred; differences in outcomes were all not significant ($p > 0.05$) except as indicated¶; results of analyses were unchanged by ignoring unknown outcomes, except box.§

†Recorded as present or absent.

‡Not recorded.

§When unknown outcomes were ignored, $p < 0.01$.

¶$p < 0.05$.

Table 5. Audit Scores in Myocardial Infarction — Comparison of Results in 50 Patients Discharged Alive with Those in 50 Patients Who Died.

Datum	Audit Score (%)*		Probability
	Patients discharged alive	Patients who died	
History:			
Documentation of presenting symptoms	100	100	
Previous symptoms of coronary heart disease	100	100	
Predisposing conditions:			
Hypertension	54	54	
Diabetes mellitus	38	40	
Family history:			
Arteriosclerotic heart disease	56	30	<0.01
Hypertension	16	14	
Diabetes mellitus	66	42	<0.05
Physical examination:			
Blood pressure	100	100	
Pulse rate	100	100	
Rhythm	100	98	
Gallop (S-3)	62	44	
P2	28	20	
Pericardial friction rub	72	54	
Pulmonary rales	100	96	
Distended neck veins	88	92	
Palpation of pedal pulses	70	60	
Laboratory:			
Electrocardiographic or enzyme study	100	88†	<0.05
Hematocrit	100	100	
Urinalysis	92	90	
Venereal Disease Research Laboratory Test	8	26	<0.01
Chest x-ray study during hospital stay	84	80	
Electrocardiogram before discharge	84	0	<0.001
Management:			
3 or more days in coronary-care unit	88	90	
Continuous intravenous drip, 1st 3 days	88	96	
Record of intake & output	90	94	
Indication of intent to see patient within 4 weeks of discharge	100	0	<0.001

*Percentages refer to items recorded as present or absent.

†For those without electrocardiogram, diagnosis of myocardial infarction was made post mortem.

course, no electrocardiogram or indication of intent to follow up in the outpatient clinic was possible before discharge in the group that died.

Discussion

Past attempts to measure quality of care have relied heavily upon examination of the recorded process of medical care. Lembcke's[2] studies were almost entirely devoted to process. Donabedian,[3] who emphasized the need to separate process from outcome, considered that recording is itself a legitimate dimension of the quality of practice as well as the medium for evaluation of most other dimensions. Nevertheless, he held that outcomes remain the ultimate validators of the effectiveness and quality of medical care.

There is a common assumption that if the documentation of events is in accord with current practice and teaching, the outcome is most likely to be optimal. This assumption may not always be correct; medical practice is seldom based upon certain knowledge of outcomes. This has led to the adoption of standards of care that are in fact standards of the process of care. In turn, this has contributed to the development of medical audits that are designed to measure the quality of care rendered but measure only the process. These audits, moreover, depend on chart documentation of that process, and yet it is self-evident that the mere act of writing cannot improve a patient's outcome. Furthermore, the only practical way to examine the process of care is to read the record; yet daily observation shows that outstanding clinicians may keep inadequate records, whereas others less competent may write profusely. The theoretical limitation inherent in auditing only the recorded process of medical care in an attempt to measure quality is obvious: if the records of all patients with a given condition contained references to every conceivable element of process, every record would be scored "perfect" in the traditional audit. Therefore, since different patients with a given condition experience different outcomes, the total audit scores could have no reproducible relation to those outcomes.

Our findings clearly show that neither quantity nor quality of recorded data was related to the outcomes of either acute appendicitis or myocardial infarct. For acute appendicitis the recording of diagnostically useful information was surprisingly variable among three hospitals, with statistically significant differences between the hospitals in the frequency of recording of eight commonly sought symptoms or signs. Despite this, each hospital achieved similar percentages of histologically proved correct diagnoses — i.e., this outcome of the clinical diagnosis of appendicitis was the same. For myocardial infarction we were unable to show any correspondence between the data recorded during the acute hospital episode and outcomes after discharge from the hospital such as newly occurring angina or congestive heart failure, recurrent myocardial infarction or death.

The reader may wonder whether a large number of possibly irrelevant audit items or failure to "weight" the scores in Table 3 might obscure a potential relation between outcome and the recording of process of care in myocardial infarction. We believe this is not the case for three reasons, the first being that, independently of whether any of the elements of the medical process are truly related to outcome, the mere recording of whether the elements were present or absent is not sufficient evidence that they influenced outcome. Secondly, when we examined separately the items of process found by Hughes et al.[1] to be related to survival or death after acute myocardial infarction, we found no difference in the outcomes according to whether or

not these items of process had been recorded (Table 4). Finally, we compared the audit scores of 50 patients who survived acute uncomplicated myocardial infarction with those of 50 patients who died from it in the hospital (Table 5). The overall patterns of the results of these two audits were similar, any differences being either expected from the circumstances or irrelevant. It appears from these findings that recording of the items on an audit list prepared by experts had little to do with whether the patients lived or died from the condition under consideration. The results also suggest that the items audited represented minimal criteria for professional performance rather than for outcome.

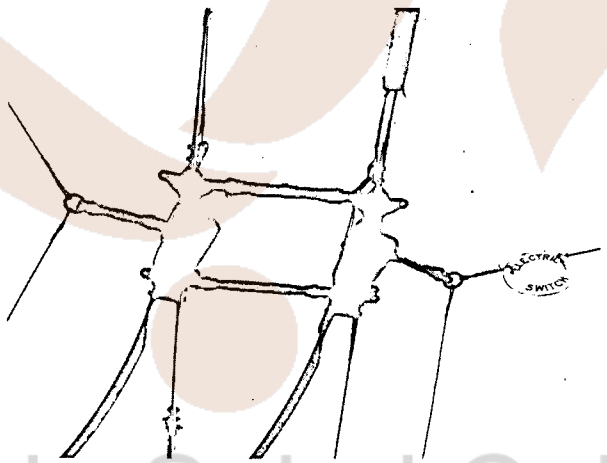
Thus, orthodox methods of audit, based upon chart review of the process of medical care, may contain serious methodologic flaws. Although we have reservations about relying upon chart review of the medical process to assess quality of care, we neither imply a condemnation of the medical record, the clinical value of which is unquestioned, nor question the clinical importance of any particular element of process and its relation to standards of medical practice.

We are firmly convinced, nevertheless, of the need to conduct audit procedures to monitor quality of medical care, and suggest that a valid medical audit should also include measures of actual outcomes of the patient's illness. We offer the following considerations for conducting such measurements: except when outcomes are very unusual or statistically improbable, one cannot evaluate an individual patient's record, but can only compare groups of records of patients having the same condition, of similar initial severity, treated by different doctors or by different methods, or perhaps at different institutions. The groups of patients being compared must initially be as similar as possible in measurable factors that are thought to influence the outcomes under consideration. Examples of outcomes that may be evaluated in many common conditions include the following: the number of days spent in the hospital; histologic confirmation of the asserted diagnosis; numbers and types of postoperative complications in the hospital; complications after leaving the hospital; later recurrences of conditions thought to have been treated definitively; long-term and short-term survival rates; elapsed time before returning to work; and improvement in functional status in cases of chronic disease (e.g., rheumatoid arthritis). Finally, some measure of the patient's satisfaction with his management and a comparison of the associated costs are necessary.

We are indebted to numerous colleagues, particularly to Drs. Morris F. Collen and John G. Smillie, for helpful discussion and criticism, to Mrs. J. S. Ilkin, R.R.L., who made the studies on the records, and to A. B. Siegelaub, M.S., who made the statistical analyses.

References

1. Hughes WL, Kalbfleisch JM, Brandt EN Jr, et al: Myocardial infarction prognosis by discriminant analysis. Arch Intern Med 111:338-345, 1963
2. Lembcke PA: Medical auditing by scientific methods: illustrated by major female pelvic surgery. JAMA 162:646-655, 1956
3. Donabedian A: Evaluating the quality of medical care. Milbank Mem Fund Quart 44 (3):166-203, Part 2, 1966



Alan Brush