How to read clinical journals:
VI. To learn about the quality of clinical care*

DEPARTMENT OF CLINICAL
EPIDEMIOLOGY AND BIOSTATISTICS,
MCMASTER UNIVERSITY
HEALTH SCIENCES CENTRE

Clinical journals are publishing increasing numbers of articles about the quality of clinical care, which we have defined as the extent to which clinical acts that do more good than harm are carried out and the extent to which those that do more harm than good are avoided.

Why should busy clinicians read such articles? First, our hospitals and professional organizations are carrying out quality-of-care studies as part of their accreditation system, and if we haven't already been the subjects of such studies we soon shall be. This being so, it's wise to understand more about how we are being judged. Second, in an effort to make quality-of-care audits more accurate, relevant and acceptable we are being asked to help carry them out, as is illustrated in the following presentation:

The head of the clinical section at your hospital has asked you to sit on the audit committee, an offer you cannot refuse. The committee is charged with carrying out at least two audits annually on the quality of care provided to inpatients on your service. Although the medical records staff have agreed to carry out the chart review, you and your colleagues have to decide what they should look for. At your first meeting you are given copies of the previous year's audit re-

ports and a folder full of articles describing quality-of-care studies done elsewhere.

What a thankless job! Still, if it must be done it ought to be done fairly, rationally and in such a way that patients and their clinicians can benefit. Our collective experience over the last decade or so in performing and reading about quality-of-care studies has enabled us to identify some general guides for interpreting such studies. In this part of the series we discuss such a set of guides.

Guides for assessing articles about the quality of clinical care

Six guides will help the busy clinician read about the quality of care (Table I). The first guide serves as a quick screening tool that will help you toss aside many articles early.

1. Did the study focus on what clinicians actually do?

Evaluations of quality of care can be classified into three categories: structure, process and outcome.[1] Clinical-structure studies tally the numbers and qualifications of health professionals and describe the administrative organization and physi-

cal facilities in which they work. Clinical-process studies document health professionals' actions in patient care. These actions are often divided into technical processes (e.g., investigations, physiologic monitoring and drug prescribing) and interpersonal processes (e.g., patient education). Finally, clinical-outcome studies document the results of this care — the "proof of the pudding" — in terms of its effects on the patient's health (survival, symptoms, and physical and psychosocial function).

Articles that focus only on structural descriptions do not deserve time from a busy clinical reader. We all know that impressive diplomas and credentials cannot guarantee clinical competence and that humble, unimpressive physical facilities can be the sites of excellent clinical care. By the same token, the authors of articles that focus only on health outcomes may be fooling themselves if they think that these outcomes are the inevitable consequences of quality of care. There are three ways in which the results of such studies can be misleading:

● When unfavourable outcomes are rare, dangerous practices can appear innocuous when the study sample is small (e.g., in small studies a fatal pulmonary embolism may not occur in patients inappro-

Table I—Guides for assessing articles about the quality of clinical care

1. Did the study focus on what clinicians actually do?
2. Have the clinical actions studied been shown to do more good than harm?
2a. If the clinical actions have not been shown to do more good than harm, is the article about a process-versus-outcome study that assembled an inception cohort, ensured high rates of compliance, included all relevant outcomes and considered prognosis?
3. Are the clinicians, patients and type of practice similar to yours?
4. Were the clinical actions measured in a clinically sensible fashion?
5. Were the clinical actions measured in a scientifically credible fashion?
6. Were both clinical and statistical significance considered?

priately confined to bed after surgery,[2] or aplastic anemia may not develop in patients inappropriately treated with chloramphenicol[3]).

• When unfavourable outcomes are delayed, dangerous practices can appear innocuous when the study period is short (e.g., a brief study would miss the development of retinopathy resulting from treatment with chloroquine[4]).

• When unfavourable outcomes are determined primarily by factors other than the process of care, spurious conclusions will be made (e.g., the risk of death following a severe stroke is influenced far more by the severity of the initial brain damage than by anything clinicians do once the damage has occurred[5]).

As a result, quality-of-care studies that focus only on structure or outcomes are unlikely to provide valid information about the real quality of care being delivered, so you can pass them by. Rather, you should look for articles that focus on the process of clinical care and tell you what clinicians have actually been doing.

Given that process studies are worth reading, how do we avoid being confronted with "laundry lists" of every conceivable clinical action that the authors think should be carried out on patients with certain illnesses? One solution is for the authors to select processes that are known, from appropriately rigorous studies, to affect outcome. The use of an "indicator condition" is one such process.[6] An indicator condition is an illness, clinical situation or state of health that either evokes clinical actions known to result in more good than harm when correctly applied (e.g., treatment with high doses of acetylsalicylic acid for rheumatoid arthritis[7]) or evokes clinical actions known to result in more harm than good (e.g., treatment with chloramphenicol for self-limited upper respiratory tract infections[3]).

By selecting only articles that study clinical actions in the types of patients seen in your hospital, you will already have eliminated the large number of articles that focus only on structure and outcome or on clinical processes for conditions that are not relevant to your hospital.

2. *Have the clinical actions studied*

been shown to do more good than harm?

This guide draws our attention to part 5 of this series, in which we discussed how to decide whether therapy does more good than harm.[8] The clinical actions of prime importance will be therapeutic: prescribing therapies that help and avoiding those that harm. However, other clinical actions are also important; for example, the actions of diagnosis that establish the need for therapy, those of dose-setting and monitoring, those of detection and management of noncompliance, and those of proper follow-up.

If you are confident that the clinical actions studied in the article have been shown to do more good than harm, you are ready to go on to guide 3. If, however, this is not the case, then guide 2 should be modified, as follows:

2a. *If the clinical actions have not been shown to do more good than harm, is the article about a process-versus-outcome study that assembled an inception cohort, ensured high rates of compliance, included all relevant outcomes and considered prognosis?*

This is a mouthful but is really quite straightforward. It states that you are wasting your reading time unless the article is either *about* clinical actions that do more good than harm or is *establishing whether* the clinical actions (or processes) are linked to clinical outcomes. If the article does neither of these, it should join the mass of articles that merely count patients seen, tests performed and money spent.

The actions that constitute clinical performance and thus the quality of care are only one determinant of the outcome (Fig. 1); without favourable influences from the other components the patient's health will not improve.

Because process-versus-outcome studies provide the rational basis for medical audits and contribute to the assessment of the integration of diagnosis, therapy and patient compliance, you will probably want to read them. If so, there are four special features you should look for.

First, the authors should have assembled an inception cohort of patients (consecutive patients with the target disorder identified at the first visit or at the beginning of therapy). This will ensure that patients who fail to continue attending the clinic or hospital because one of the outcomes of interest has occurred (e.g., death or severe disability) will not be lost to follow-up. Two widely quoted process-versus-outcome studies, one on myocardial infarction[9] and the other on hypertension,[10] did not use inception cohorts; rather, their inclusion criteria were that patients would be available for follow-up or would come back to the clinic. As a result, the authors were likely to miss important outcomes and thus make inaccurate conclusions.

Second, patient compliance must be taken into account in process-versus-outcome studies since it is an essential link between the actions by the health care team and the outcomes of the patients. Studies that have not considered compliance
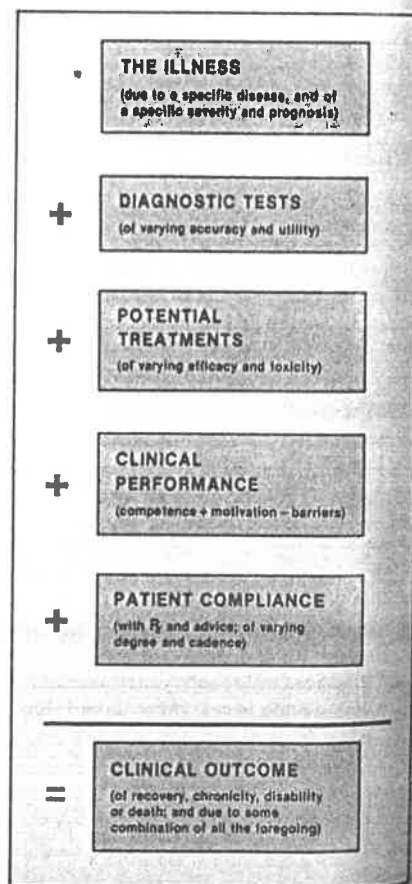

FIG. 1—Place of clinical performance as one of several components that determine clinical outcome.

have failed to demonstrate such an association.[10-14] On the other hand, studies that have considered compliance have documented strong links between the clinical process and the outcome of care in patients with urinary tract infections, hypertension, iron-deficiency anemia and other conditions.[15-17] Before stating that there is no relation between process and outcome the authors should provide evidence that the lack of such an association was not merely due to low rates of compliance.

Third, a process-versus-outcome study must examine all clinically relevant outcomes. These might include death, improvement or deterioration in physical, emotional or social function, or even in the physiologic control of a risk factor for a disease (when it is justified to substitute alterations in such risk factors for death or disability, as in hypertension[18]), patient satisfaction and the side effects of therapy. The methods for assessing these outcomes should be accurate and, to minimize bias, should have been used by an assessor who was blind to the details of the patients' management.

Finally, to isolate the effects of the clinical processes from those of other determinants of outcome, process-versus-outcome studies must control for the other determinants. Because patients with more severe disorders usually need more extensive assessment, monitoring and intervention but will tend to have poor outcomes irrespective of the quality of care provided, a study that failed to control for the severity of illness would conclude that the greater the number of processes carried out, the worse the outcome!

Once you are satisfied that there is indeed a link between the clinical process and outcome, you are ready to apply the next guide.

3. Are the clinicians, patients and type of practice similar to yours?

Information on the clinicians' training, certification, years of experience and workload, on the source and types of patients, and on the type of practice (primary, secondary or tertiary care) in an article will help you decide whether the results

of the study will apply to you and your patients. Moreover, the methods used to select the clinicians and patients should be described in the article so that you can judge whether the individuals are likely to be representative of a particular population. For example, if only the clinicians who volunteered for the study were included, and they accounted for less than 80% of clinicians who were eligible for the study, it would not be wise to apply the results to those who did not volunteer or to clinicians in general.

Finally, the rules for what constitutes a "case" should be realistic. For example, eligibility for inclusion in a study of an indicator condition should be based on the presence of an illness or complaint (e.g., sleep disturbance) rather than on the presence of disease (e.g., depression), since the latter assumes a correct diagnostic process and therefore excludes patients in whom the indicator condition was present but not diagnosed. Similarly, except for discrete emergencies, the definition of a case should allow the assessment of episodes of care, including follow-up visits when appropriate.

Since it is unlikely that any study setting will exactly replicate your practice, you should ask yourself Are the clinicians and their practices *so different* from those in the area in which I am interested that I could *not* apply the study results?

4. Were the clinical actions measured in a clinically sensible fashion?

The rules for measuring the clinical actions must be explicit, comprehensive and flexible. First, the measurement criteria should be explicitly stated so that you can see if they make sense. This means that the actual criteria used in the study must be available or, ideally, mentioned in the article.

Second, the process criteria should be comprehensive. That is, they should include all the clinically important aspects of care (e.g., history-taking, physical examination, laboratory tests, therapy, patient education, consultation etc.), and they should have been applied to all the patients (e.g., in different age groups and different socioeconomic groups). You cannot assume that a

clinician who complied with the defined standards for history-taking, physical examination and laboratory investigation also complied with the standards for drug therapy, patient education and follow-up. Moreover, because of conflicting evidence as to whether a clinician's actions for one type of problem are representative of his or her actions for other types of problems,[19,20] the authors should have evaluated all the conditions for which they drew conclusions. However, a sizeable proportion of most practices can be represented by a small selection of health problems. Therefore, by covering just these, you can obtain a reasonable overview of a clinician's quality of care.[21]

Finally, it is important to consider whether the process criteria were flexible enough to allow for appropriate changes in management according to variations in diagnosis, severity of disease and complications. In many situations this is not necessary, so it would be acceptable to apply identical criteria to all the patients (e.g., a criterion for all children to have been fully immunized by a certain age would be appropriate). On the other hand, you should demand more sophisticated assessment methods when patients present with complex problems (e.g., pleuritic chest pain). Similarly, flexible process criteria are appropriate for assessing the management of patients with different complications or degrees of severity of disease (e.g., those with chest pain due to acute myocardial infarction). All too often, long and invariant "laundry" lists are applied to heterogeneous groups of patients. On the other hand, Sibley and colleagues' indicator-condition approach,[6] which stratifies patients according to severity of disease, and Greenfield and associates' criteria-mapping method,[22] which uses branching criteria to reflect sequential judgements, are two ways of addressing this issue; authors should therefore have incorporated such methods when they were needed.

5. Were the clinical actions measured in a scientifically credible fashion?

The method used to decide whether specific clinical actions were carried out must be both accurate and

precise. Accuracy means that the method should have revealed the clinically important deficits or differences in the clinical process, and precision means that different examiners applying the same method to the same cases should have come to the same conclusions about whether specific clinical processes were carried out.

A major limitation to the accuracy of measuring clinical processes by methods that rely on medical records is the inconsistency with which different physicians record their clinical actions.[23] As a result it may be difficult or impossible to distinguish clinical actions that were not performed from those that were performed but not recorded. On the other hand, it can be argued that good clinical care necessitates good record keeping; important actions have to be recorded for recall, for communication with others and for minimizing the clinical errors that result from recording only the conclusions and management decisions and not the observations that led to them. Surprisingly, despite the faults of clinical records, several studies have shown that they can provide accurate, adequate data for a wide range of clinical processes; however, this accuracy is usually achieved when explicit criteria have been restricted to the essentials.[6,22] When one relies only on implicit clinical judgements about quality there is often a lack of consistency, even among "experts".[11] To achieve precision, therefore, the criteria for acceptable quality of care must be explicit. Moreover, there should be a high degree of interobserver agreement in the application of the criteria.

Finally, the information should have been collected in a blind fashion to avoid bias due to the expectations of the abstractor, who knows the patient, the clinician or the outcome. When logistics have made this difficult, a sample of the observations should have been assessed in a blind fashion to obtain a measure of the bias.

6. *Were both clinical and statistical significance considered?*

The analysis must be clinically sensible, and its results must be applicable in routine clinical care.

Beware of "scores" and analyses that have simply added up the number of "right" actions, for such analyses can generate ludicrous conclusions. For example, a clinician caring for a patient who is vomiting blood could take a thorough history (worth lots of points), flawlessly execute a series of diagnostic tests (worth lots more points) and end up with a reasonably high score while the patient continues to lose blood for want of proper therapy. In fact, actions related to correct management constituted only 1% to 16% of the scores assigned in widely quoted quality-of-care studies.[9,10,24]

Appropriate analyses are of two sorts. The first simply comprises a table that lists each action and the proportion of cases in which it was carried out. This descriptive analysis permits you to draw your own conclusions as well as to study those of the authors. The second establishes a minimum set of actions, *all* of which must be carried out for a case to have been handled with adequate quality of care.

Finally, any conclusions about the quality of care provided in such articles should be backed up by tests of statistical significance. These were discussed in part V of this series; to recap, you should ask yourself the following questions:

● If the difference in quality of care is statistically significant, is it also clinically significant?

● If the difference is not statistically significant, was the study sample large enough to enable detection of clinically significant differences in the quality of care if there were any?

Conclusion

The readers' guides in this paper can serve many purposes. The first guide (Did the study focus on what clinicians actually do?) can help you choose the few studies that are most likely to be both valid and useful. Moreover, the use of all the guides can help you not only to understand the studies done by others but also to design your own studies. The application of previously described guides on how to select accurate diagnostic tests[25] and efficacious treatments[8] will help you decide which clinical processes you should include in your studies. The results

of your study can then become the basis for a grand rounds or the starting point for clinical research.

References

1. DONABEDIAN A: Evaluating the quality of medical care. *Milbank Mem Fund Q* 1966; 44 (suppl): 166–206
2. Prevention of fatal postoperative pulmonary embolism by low doses of heparin. An international multiple centre trial. *Lancet* 1975; 2: 45–51
3. WALLERSTEIN RO, CONDIT PK, KASPER CK, BROWN JW, MORRISON FR: Statewide study of chloramphenicol therapy and fatal aplastic anemia. *JAMA* 1969; 208: 2045–2050
4. HOBBS HE, SORSBY A, FREEDMAN A: Retinopathy following chloroquine therapy. *Lancet* 1959; 2: 478–480
5. HUTCHINSON EC, ACHESON EJ: *Strokes: Natural History, Pathology and Surgical Treatment*, Saunders, Toronto, 1975: 138–153
6. SIBLEY JC, SPITZER WO, RUDNICK KV, BELL JD, BETHUNE RD, SACKETT DL, WRIGHT K: Quality-of-care appraisal in primary care: a quantitative method. *Ann Intern Med* 1975; 83: 46–52
7. MAINLAND D, SUTCLIFFE MI: Aspirin in rheumatoid arthritis, a seven-day, double-blind trial — preliminary report. *Bull Rheum Dis* 1965; 16: 388–391
8. Department of clinical epidemiology and biostatistics, McMaster University Health Sciences Centre: How to read clinical journals: V. To distinguish useful from useless or even harmful therapy. *Can Med Assoc J* 1981; 124: 1156–1162
9. FESSEL WJ, VAN BRUNT EE: Assessing quality of care from the medical record. *N Engl J Med* 1972; 286: 134–138
10. NOBREGA FT, MORROW GW, SMOLDT RK, OFFORD KP: Quality assessment in hypertension: analysis of process and outcome methods. *N Engl J Med* 1977; 296: 145–148
11. BROOK RH: *Quality of Care Assessment; a Comparison of Five Methods of Peer Review* (DHEW publ no HRA 74-3100), US Health Resources Administration, Bureau of Health Services Research and Evaluation, Rockville, Md, 1973
12. LINDSAY MI JR, HERMANS PE, NOBREGA FT, ILSTRUP DM: Quality-of-care assessment. I. Outpatient management of acute bacterial cystitis as the model. *Mayo Clin Proc* 1976; 51: 307–312
13. LINDSAY MI JR, NOBREGA FT, OFFORD KP, CARTER ET, RUTHERFORD BD, KENNEL AJ, MANKIN HT: Quality-of-care assessment. II. Outpatient medical care following hospital dismissal after myocardial infarction. *Mayo Clin Proc* 1977; 52: 220–227
14. ROMM RJ, HULKA BS, MAYO F: Correlates of outcomes in patients with congestive cardiac failure. *Med Care* 1976; 14: 765–776
15. STARFIELD B, SCHEFF D: Effectiveness of pediatric care: the relationship between processes and outcome. *Pediatrics* 1972; 49: 547–552

16. PAYNE BC, LYONS TF: Method for estimating and improving ambulatory medical care. In *Final Report to National Center for Health Services Research*, National Center for Health Services Research, Rockville, Md, Nov 1978

17. Hypertension Detection and Follow-up Program Cooperative Group: Five year findings of the Hypertension Detection and Follow-up Program [two parts]. *JAMA* 1979; 242: 2562–2577

18. HAYNES RB, GIBSON ES, TAYLOR DW, BERNHOLZ CD, SACKETT DL: Process versus outcome in hypertension: a positive result. *Circulation* 1982; 65: 28–33

19. LYONS TF, PAYNE BC: Interdiagnosis relationships of physician performance measures in hospitals. *Med Care* 1977; 15: 475–481

20. OSBORNE CE: Interdiagnosis relationships of physician recording in ambulatory child health care. Ibid: 465–474

21. HAMATY D: ASPERF research progress report. In *American Society of Internal Medicine Assessment of Performance. Proceedings of Assessing Physician Performance, San Francisco, June 18–19, 1976* (publ 328), Am Soc Intern Med, San Francisco, 1976

22. GREENFIELD S, LEWIS CE, KAPLAN SH, DAVIDSON MB: Peer review by criteria mapping: criteria for diabetes mellitus. The use of decision-making in chart audit. *Ann Intern Med* 1975; 83: 761–770

23. OSBORNE CE, THOMPSON HC: Criteria for evaluation of ambulatory child health care by chart audit: development and testing of a methodology. Final report of the joint committee on quality assurance of ambulatory health care for children and youth. *Pediatrics* 1975; 56: 625–692

24. FERNOW LC, MACKIE C, McGOLL I, RENDALL M: The effect of problem-oriented medical records on clinical management controlled for patient risks. *Med Care* 1978; 16: 476–487

25. Department of clinical epidemiology and biostatistics, McMaster University Health Sciences Centre: How to read clinical journals: II. To learn about a diagnostic test. *Can Med Assoc J* 1981; 124: 703–710

COPING
continued from page 374

among medical students and physicians. *Univ Toronto Med J* 1982; 60: 3–6

24. RIPPE J, BENNET N, TAYLOR H, WARNER CK: Learning from each other: reflections on a medical student discussion group. *J Med Educ* 1978; 53: 686–688

25. BERG JK, GARRARD J: Psychosocial support in residency training programs. *J Med Educ* 1980; 55: 851–857

26. SIEGEL B, DONNELLY JC: Enriching personal and professional development: the experience of a support group for internes. *J Med Educ* 1978; 53: 908–914