17/6/09 & 30/7/09

CLINICAL EPIDEMIOLOGY ROUNDS

How to read clinical journals: III. To learn the clinical course and prognosis of disease

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

This is the third in our consecutive series of Clinical Epidemiology Rounds devoted to efficient yet effective strategies and tactics for reading clinical journals. This round concerns reading these journals to learn more about the clinical course and prognosis of disease. As before, the rules for reading begin with four universal guides for examining the title, the authors, the summary and the site (Fig. 1). These guides then branch, depending on the reader's intent.

Case presentations

An otherwise healthy 32-year-old engineer passes (and discards) a urinary stone. His history, physical exam, routine urinalysis and serum calcium level are normal, and you wonder whether the chances of recurrence are high enough to warrant a "stone work-up".

A robust 12-year-old girl is referred to you by the school nurse, who discovered 10° of scoliosis on a screening exam: you have confirmed this finding. The patient and her family are concerned that she will become a cripple.

A 37-year-old accountant has well controlled ulcerative colitis that began 20 years previously. Although it is confined to the left side of the colon, a colleague has suggested that you consider a prophylactic colectomy to obviate the risk of colorectal cancer.

Reprint requests to: Dr. Peter X. Tugwell, Rm. 2C12, McMaster University Health Sciences Centre, 1200 Main St. W. Hamilton, Ont. L8N 3Z5

Comment

The clinical management of these patients includes, as one central theme, making judgements about the likely time course of their illnesses: Will the urinary stone recur or does it portend serious disease requiring immediate diagnosis? Will the 10° of scoliosis progress to a frank physical or physiologic derangement? Do patients with left-sided colitis face an increased risk of colorectal cancer? These judgements play a central role in deciding whether any intervention is needed by these patients and in planning a thoughtful approach to counselling them and their families.

Each of these patients is experiencing the *natural history* of their disease — that is, the time course of the interaction between the individual, causal factors and the rest

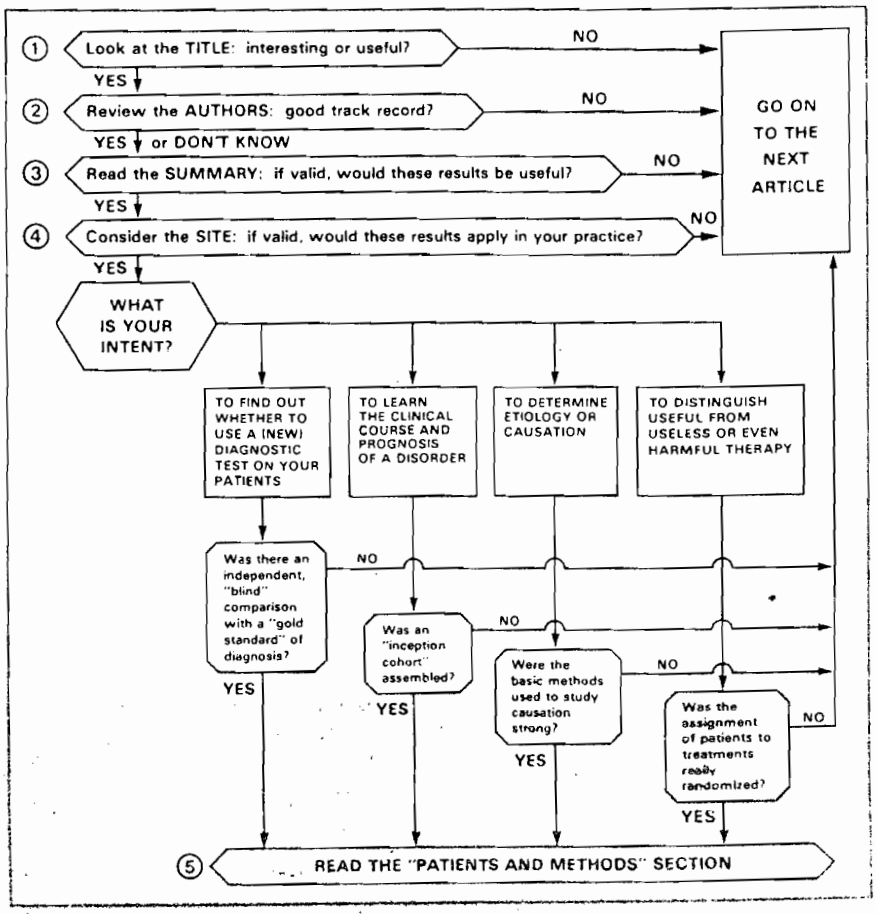

FIG. 1—The first steps in how to read articles in a clinical journal.

of the environment, beginning with the biologic onset of disease and ending with recovery, death or some other physical, social or emotional state. Of special interest to clinicians is the portion of this natural history that begins with the first unambiguous signs or symptoms of illness: the *clinical course*. More specifically, we are called upon to make judgements about *prognosis* — the probability of one or another of the outcomes developing in the natural history and clinical course of an illness.

We find out about the clinical course and prognosis of disease by reading about them ourselves or by asking someone who has read about them. But can we accept the published reports at face value? The reported rates of stone recurrence vary from 40%[1] (in which case we might forgo a work-up following a first stone) to nearly 100%[2] (which suggests that we might as well get on with it). The published spontaneous recovery rates for scoliosis vary sevenfold,[3] and those for cancer risk in patients with ulcerative colitis vary threefold, from 3%[4] to nearly 10%.[5] With this much inconsistency in clinical journals, how are we to use them sensibly in making clinical decisions about the clinical course and prognosis of illness?

As it happens, most of the inconsistencies among the published studies of clinical course and prognosis are due to the different ways (many of them wrong) their authors selected and followed the patients. None the less, the requirements for the proper study of clinical course and prognosis are relatively straightforward, involve the sort of "applied common sense" we've called upon earlier in this series of rounds, and can be translated into a brief set of standards or readers' guides that can be applied quickly to published articles.[6]

Guides for reading articles about the clinical course and prognosis of disease

The six guides for reading such articles can be posed as questions, which are listed in Table I. Because they focus on how the clinical study was carried out they are applied to the "Patients and methods" section of the articles.

1. *Was an "inception cohort" assembled?*

Patients should be identified at an early and uniform point ("inception") in the course of their disease (such as when unambiguous symptoms first develop or when the patients receive their first definitive therapy), so that those who succumb or completely recover are included with those whose disease persists.

Many studies of prognosis are done backwards. For example, several studies of the risk of stone recurrence ask *currently* symptomatic patients if they've had stones *previously*, failing to realize that recurrent stone-formers (with positive past histories) have multiple chances to be included in such studies; however, patients without recurrences (with negative past histories) have only one chance of being included. No wonder that recurrence rates vary all over the map! Similarly, in one recent study of cancer risk in patients with ulcerative colitis the stimulus for including several patients in the study was the development of cancer.[5] Clearly they had not been followed since the inception of the ulcerative colitis, and since colitis patients who remained cancer-free would not be entered through such a mechanism the apparent risk of cancer in patients with ulcerative colitis would become spuriously elevated.[6]

The failure to start a study of clinical course and prognosis with an inception cohort has an unpredictable effect on its results. In the cases that opened this round, the effect would be to make the prognosis appear gloomier than it really is. The opposite mistake can also occur, however. Suppose we wanted to learn more about the prognosis of patients who suffer a myocardial infarction and we therefore read an article about a collection of such patients that had been assembled in a coronary care unit. One major problem with such a study is that it would fail to include patients with myocardial infarction who died before they could get to the coronary care unit, yet these patients would account for more than half of the deaths within the first year following the myocardial infarction. As a result, and in contrast to the earlier examples, a falsely rosy prognosis would be concluded from such a study.

Thus, the failure to assemble a proper inception cohort of patients who are at an early and uniform point in the course of their disease usually constitutes a fatal flaw in studies of prognosis. Moreover, the application of this guide to an article can increase your efficiency in reading the clinical literature: if the authors failed to assemble an inception cohort, discard the article early and go on to something else.

2. *Was the referral pattern described?*

The pathways by which patients entered the study sample should be described. It must be possible for you, the reader, to be able to tell whether the results apply to patients in your practice. Did they come from a primary care centre? Were all hospitals in a defined region scoured for cases? Were the patients assembled in a tertiary care centre that attracted the hopeless, wealthy or bizarre? It is in the assembly of patients that studies of the course and prognosis of disease often flounder, for it is here that four types of bias are most pervasive.[6,7]

Because a major clinical centre's reputation results in part from its particular expertise in a specialized area of clinical medicine, it will be referred problem cases likely to benefit from this expertise (the *centripetal bias*), and its experts may preferentially admit and keep track of these cases over other less challenging or less interesting ones (the *popularity bias*). In any event,

Table I—Guides for reading articles about the clinical course and prognosis of disease

1. Was an "inception cohort" assembled?
2. Was the referral pattern described?
3. Was complete follow-up achieved?
4. Were objective outcome criteria developed and used?
5. Was the outcome assessment "blind"?
6. Was adjustment for extraneous prognostic factors carried out?

the selection that occurs at each stage of the referral process can generate patient samples at tertiary care centres that are much different from those found in the general population (the *referral filter bias*). Finally, patients differ in their financial and geographic access to the clinical technology that identifies them as eligible for studies of the course and prognosis of disease. If this degree of access is linked to the risk of a poor outcome (such as would occur, for example, if patients whose headaches were due to brain tumours had greater access to computer-assisted tomography scanners than patients with benign causes for their headaches), the resulting *diagnostic access bias* will distort the conclusion of the study.

Thus, sampling biases can distort both the timing and the rates of important prognostic outcomes. Despite this serious drawback, the study of inception cohorts at tertiary care centres yields useful information to other clinicians who work in such settings. Moreover, studies in tertiary care centres can provide useful information on the potential importance of prognostic subgroups as long as sampling biases affect each of the subgroups equally; however, this equality may be difficult to show and risky to assume.

These sampling biases are largely responsible for the chaos that characterizes most discussions of the course and prognosis of disease; such pitfalls are not casually avoided. Short of the Framingham type of study, in which a large population of individuals is assembled and closely followed for decades, the sampling approach worth looking for is the one that has systematically gathered eligible cases from *all* the clinical facilities in a given catchment area through the careful review of old clinical records or, better still, through continuing surveillance for new cases.

In summary, an understanding of how the patients were assembled for the study, plus additional information such as age, sex, severity of disease and coexisting disorders, will help you decide whether they resemble your own patients enough for you to apply the results of their follow-up in your own practice.

3. *Was complete follow-up achieved?*

All members of the inception cohort should be accounted for at the end of the follow-up period, and their clinical status should be known. Patients don't disappear from a study for trivial reasons; rather, they leave the study because they refuse therapy, recover, die, retire to the Sunbelt or simply grow tired of being followed. All of these reasons are linked to important prognostic outcomes and, if you are to use the results of the article in making prognostic judgements about your own patients, you deserve to know how *all* the members of the inception cohort fared.

Of course, it's difficult for the authors to achieve perfection; they're bound to lose a few members of their inception cohort. There are, however, some rough rules of thumb that you can apply here. The loss to follow-up of more than 10% of the original inception cohort is cause for concern. If more than 20% are lost the results of the study are probably not worth reading. Thus, this rule provides you with yet another tactic for increasing your efficiency: if less than 80% of the inception cohort are accounted for in a study of clinical course and prognosis discard the article.

4. *Were objective outcome criteria developed and used?*

The prognostic outcomes should be stated in explicit and objective terms so that you, as the reader of the subsequent report, will be able to relate them to your own practice. Suppose that you came upon an article about the prognosis of patients with transient ischemic attacks. If the article describes the risk of "subsequent stroke" without presenting explicit, objective criteria for what constitutes a "stroke" you are in a quandary. Are the strokes limited to severe derangements of sensation or motor power such that their victims require assistance in dressing, feeding and toileting? Or are most of them merely transient or trivial changes in sensation or in deep or superficial reflexes? The implications of these different definitions for counselling patients or initiating therapy are whopping.

Not only should an article describe explicit and objective outcome criteria, but also it should provide reassurance that these criteria were applied in a consistent manner. As you will recall from an earlier set of Clinical Epidemiology Rounds on clinical disagreement, even experienced clinicians will disagree among themselves about key manifestations of a disease.[8] As a result, what is a stroke to one clinician is merely a normal variant to another; a patient's apparent prognosis will be determined not by biology but by the luck of the draw in who is selected to perform the final examination.

All of this brings us to the next guide.

5. *Was the outcome assessment "blind"?*

The examination for important prognostic events should be carried out by clinicians who are "blind" to the other features of the patients. This is essential if two additional sources of bias are to be avoided.[6,7] First, the clinician who knows that a patient has a prognostic factor of presumed importance may carry out more frequent or more detailed searches for the relevant prognostic event (the *diagnostic suspicion bias*). Second, pathologists and others who interpret diagnostic specimens can have their judgements dramatically influenced by prior knowledge of the clinical features of the case (the *expectation bias*).

The expectation bias should be familiar to readers who follow these Clinical Epidemiology Rounds, for it appeared in our initial series on clinical disagreement (remember the tonsils?).[8] In order to reduce expectation bias, many clinical centres recognize the need, even in routine practice, for an initial "blind" assessment of diagnostic tests such as electrocardiograms and roentgenograms.[9] The diagnostic suspicion bias can be avoided by subjecting *all* patients to the *same* diagnostic studies, perhaps at prescribed intervals and, by all means, at the end of the study.

What about the outcome of death? Surely blindness is not a prerequisite for assessing so "rigorous" an outcome. Although reporting the *fact* of death is an unambiguous task (and need not, therefore, be blind), assigning a *cause* of death is subject to both diagnostic suspicion bias and expectation bias and ought, therefore, to be done blindly.

6. Was adjustment for extraneous prognostic factors carried out?

Suppose you wanted to know whether the duration of your patient's ulcerative colitis was an important determinant of the risk of cancer. To get a clear-cut answer to this question you would like to be sure that there would be no interference from other factors that might affect both disease duration and prognosis (such as earlier age at onset, more courses of a wider range of therapy etc.). The failure to meet this standard could result in assigning causal roles to factors that are merely "markers" for other factors of real importance.

Of course, it's usually impossible for the clinical reader to become sufficiently familiar with the mathematical tap-dancing used to adjust for extraneous prognostic factors to be able to tell whether, for example, the authors should have used a discriminant function analysis rather than a stepwise multiple logistic regression. This is one of those situations in which you must (and usually can) rely on the journal to have had the article previously reviewed and approved by a good biostatistician. However, the article is much less likely to have undergone sophisticated methodologic review if no adjustment for extraneous prognostic factors was attempted; the reader is therefore left with the responsibility to decide whether an adjustment procedure was required (but not how well it was executed).

Use of these guides to reading

This round has presented six guides that a busy clinician can apply to an article on the clinical course and prognosis of disease. Their application should have two results: first, you can discard early many, if not most, of the prognosis articles you come across, thereby increasing the efficiency with which you spend your precious reading time. This will be particularly so if you rigorously apply the first guide (Was an inception cohort assembled?), because most of the faulty studies of clinical course and prognosis fail this initial test. Second, the articles that do pass muster will provide you with prognostic information that is increasingly valid, consistent and applicable in your own clinical practice.

The next in this series of Clinical Epidemiology Rounds will consider how to read clinical journals to learn about the causes of human illness.

References

1. ALMBY B, MEIRIK O, SCHÖNEBECK J: Incidence, morbidity and complications of renal and ureteral calculi in a well defined geographical area. *Scand J Urol Nephrol* 1975; 9: 249–253
2. COE FL, KLCK J, NORTON ER: The natural history of calcium urolithiasis. *JAMA* 1977; 238: 1519–1523
3. ROGALA EJ, DRUMMOND DS, GURR J: Scoliosis: incidence and natural history. A prospective epidemiological study. *J Bone Joint Surg [Am]* 1978; 60: 173–176
4. MOTTET NK: Neoplastic sequelae. In *Histopathologic Spectrum of Regional Enteritis and Ulcerative Colitis*. Saunders, Philadelphia, 1959: 217–235
5. GREENSTEIN AJ, SACHAR DB, SMITH H, PUCILLO A, PAPATESTAS AE, KREEL I, GELLER SA, JANOWITZ HD, AUFSES AH JR: Cancer in universal and left-sided ulcerative colitis: factors determining risk. *Gastroenterology* 1979; 77: 290–294
6. SACKETT DL, WHELAN G: Cancer risk in ulcerative colitis: scientific requirements for the study of prognosis (E). *Gastroenterology* 1980; 78: 1632–1635
7. SACKETT DL: Bias in analytic research. *J Chronic Dis* 1979; 32: 51–63
8. Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ont.: Clinical disagreement: I. How often it occurs and why. *Can Med Assoc J* 1980; 123: 449–504
9. SPODICK DH: On experts and expertise: the effect of variability in observer performance. *Am J Cardiol* 1975; 36: 592–596