Clinical review

Observational studies as alternatives to randomized clinical trials in surgical clinical research

Xiaohan Hu, PhD, James G. Wright, MD, MPh,[a] Robin S. McLeod, MD, Al Lossing, MD, *and* **Beverly C. Walters, MD,** *Toronto, Ontario, Canada*

From the Clinical Epidemiology Group, Department of Surgery, University of Toronto, Ontario, Canada

CLINICAL EPIDEMIOLOGY AS A DISCIPLINE has evolved rapidly during the last decade. However, application of epidemiologic principles in surgical clinical research has been limited. In the March 1996 issue of SURGERY we discussed the principles, application, and limitations of randomized clinical trials (RCTs) in surgery. RCTs are often viewed as the gold standard for evaluation of therapeutic efficacy because they are the closest that clinical medicine can come to the controlled environment of laboratory research. However, clinical decisions often have to be made in the absence of data obtained this way. In this issue we will discuss some alternatives that can be used in surgical clinical research.

Alternatives to RCTs fall into the purview of observational epidemiology, which can be grouped into five major categories: case reports, case series, cross-sectional studies, case-control studies, and cohort studies. The former three are often referred to as descriptive studies, and the remaining two are labeled as analytic or comparative studies.

DESCRIPTIVE STUDIES

Descriptive studies differ from analytic studies in that they contain no control or comparison groups and are usually undertaken when little is known about the epidemiology of the disease in question. The analysis often focuses on a description of the disease in the study population according to patient characteristics such as age, gender, ethnic origin, socioeconomic class, occupation, geographic area, time of occurrence, and clinical course. The results of descriptive studies are often used for hypothesis generation, baseline data, rationale for sample size estimation, and, in the absence of treatment, a description of the natural history of the disease. Descriptive studies are also used in the early stages of the evaluation of a new treatment. They are generally considered as the first, but sometimes necessary, step toward subsequent analytic studies or RCTs.

Case report. A case report is the description of clinical events of one or several patients in a narrative form. Its place in clinical research lies in reporting rare and significant complications of disease, or describing the treatment of an unusual disease. The purpose of a case report is to prompt colleagues to look for similar effects in their patients, so that initial isolated observations can be either verified or nullified.

Case series. Case series describe the spectrum of clinical features of a group of patients who are monitored from the same inception point of the clinical course of a disease. Routinely collected data, such as medical records, are commonly used. When a case series is used to describe the response to treatment, patients' clinical courses are usually compared with historical controls because the study design lacks a comparative (control) group. This form of research is prone to bias because the characteristics of the patients or the methods of their assessment may account for differences in outcome rather than the treatment. As discussed in the previous article in the March 1996 issue of SURGERY, case series seldom provide proof of therapeutic efficacy.

Cross-sectional studies. In cross-sectional studies individuals are characterized by hypothesized risk factors and a disease of interest at one specific point in time.

Accepted for publication April 18, 1995.

Reprint requests: James G. Wright, MD, MPh, S107, The Hospital for Sick Children, 555 University Ave., Toronto, Ontario M5G 1X8, Canada.

[a]Recipient of a Medical Research Council Scholarship.

SURGERY 1996;119:473-5.

0039-6060/96/$5.00 + 0 **11/55/70912**

Cross-sectional studies are used to assess the magnitude of a disease in the population and to determine the prevalence of risk factors for disease. One major advantage of a cross-sectional study is that it can be carried out in a timely fashion. Lack of appropriate controls and uncertainty about whether risk factors precede the onset of disease are the two major threats to the validity of results. For example, to explore risk factors for multiple organ failure (MOF) in patients in intensive care unit, a cross-sectional study could be used. All patients with the signs of MOF could have multiple tests including blood cultures. If the results showed that 50% of the patients have positive blood cultures, investigators might conclude that infection is the origin of MOF and recommend prophylactic antibiotics to prevent MOF. Inclusion of a control group without MOF, however, might show that 50% of patients without MOF also have positive blood cultures. Alternatively, infection may be an early manifestation rather than a cause of MOF. Thus cross-sectional studies can suggest an association but do not provide proof of causation.

ANALYTIC STUDIES

Case-control studies. Case-control studies have a control or comparison group, but patients are assembled according to the presence or absence of the outcome. In earlier literature case-control studies have sometimes been referred to as case-referent studies, case-comparison studies, or simply retrospective studies. A group of subjects with a known outcome (cases) and a group of subjects without the outcome (controls) are compared for the proportion of risk factors. For example, in a study of the origin of breast cancer a group of women who are recently diagnosed with breast cancer (cases) and a group of women who are free of breast cancer (controls) are recruited. The two groups are compared with respect to previous oral contraceptive use, estrogen-replacement therapy use, age at first birth, menarche, and menopause, history of benign breast disease or breast feeding, breast cancer in sister or mother, etc. A difference in the frequency of these factors in the two groups shows an association, which may or may not be etiologic.

Case-control studies are particularly valuable in etiologic studies because such design allows comparison of multiple factors as illustrated in the above example. In addition, they generally can be conducted in a short period of time, require small sample sizes, and are the best design for rare events or late outcomes. Case-control studies are superior to descriptive studies because they have a control group but are susceptible to other forms of bias.[1, 2]

The two most prominent types of bias are selection and information bias. Selection bias results from inappropriate selection of the control subjects. Identification of appropriate controls is probably the most difficult and controversial aspect of a case-control study. In principle the control group is intended to provide an estimate of the exposure rate that would be expected to occur in the cases if no association was present between the study disease and exposure.[3] Consequently, patients with conditions known to predispose to or against the exposure under study should be excluded from the control group. For example, patients with lung cancer should be excluded as controls in a study of association between smoking and coronary heart disease. Otherwise, the risk factor (smoking) will be underestimated because subjects with a past smoking history are overrepresented in the control group as a result of the causal association between lung cancer and smoking. Similarly, in a study of the effectiveness of Pap smear tests in reducing mortality from cervical cancer, patients with a hysterectomy should be excluded because they would be less likely to have a Pap smear test, which could reduce the apparent effectiveness of the test (you would expect a much higher percentage of women in control group to have Pap smear test if the test is highly effective). In clinical research, cases are usually selected from a medical institute and the control group is composed of patients with other diseases-conditions from the same institute.

Information bias occurs as a results of flaws in the methods used to collect the data. For example, patients with breast cancer might recall exposures more thoroughly than their controls. To prevent information bias it is imperative that efforts to ascertain exposure are comparable in the two groups. Whenever possible, for example, interviewers should be unaware or blind to disease status of study subjects, and patients or medical record abstractors should not know the main purpose of the study.

Case-control studies can also be used to evaluate the effectiveness of treatments but are rarely used in surgical clinical research.[4] For example, Nissen and Toupet fundoplication procedures are concurrently practiced in our hospital as alternatives for children with gastroesophageal reflux (GER) for whom medical management fails. At present the choice between the two procedures is based on surgeon preference. Assuming that surgical skills are comparable among the surgeons, a case-control study could be conducted to determine whether one procedure is superior to the other. Patients in whom recurrence of reflux has occurred (cases) would be compared with patients in whom resolution of reflux has been observed (controls). The type of the procedure received by subjects, along with characteristics known to influence GER, could be compared. In the absence of systematic bias the outcome as estimated by odds ratio would reflect the magnitude of the relative therapeutic effectiveness between the two procedures.

Cohort study. Cohort studies start by assembling two groups of subjects. One group of patients receives a treatment (or has a risk factor), and the other group does not have the treatment (or the risk factor). Both groups are followed up over a period of time, and then the incidence of events in each group is observed and compared. For example, a group of patients with high cholesterol level and a group of control subjects with normal cholesterol level are followed up over a certain period. At the end of study the incidence of coronary heart disease in each group is tabulated and the rates are compared. In another example, to compare the effectiveness of Nissen and Toupet fundoplication procedures for patients with intractable GER, patients are grouped by the procedure they received. Patients are evaluated 2 years after operation, and the outcomes in the two groups are compared.

Compared with case-control studies, the major advantages of the cohort study are that multiple outcomes can be evaluated, rates of outcomes can be directly calculated, and exposure always precedes the outcome. Moreover, the cohort study is intuitively appealing because its temporal sequence represents the natural causal pathway. Cohort studies are useful for evaluation for risk factors because patients cannot be randomly assigned to risk factor or not to receive the risk factor. Cohort studies like RCT, however, are not feasible when the incidence of the event under study is low, such as the risk of the human immunodeficiency virus infection after an infected sharp injury among surgeons, or the time from exposure to the presentation of the event is protracted, such as the association of smoking with emphysema.

One strategy often used to overcome the disadvantage of delayed outcome is to assemble a historical cohort, in which a group of individuals is identified on the basis of their exposure-intervention status in the past and then their subsequent disease experience is determined up to a point: in the more recent past, the present time, or even the future. This method is termed a historical cohort study or retrospective cohort study. A retrospective cohort study shares the advantages of a cohort study. In addition, because the study can be completed in a much shorter time period, it is therefore considerably less expensive. However, such a study is only suitable when comprehensive information on baseline status, exposure, or intervention is available.

The major source of bias in cohort studies, called confounding, derives from imbalance of known or unknown prognostic factors between compared groups. For example, in a comparison of mortality rates for women with breast cancer who are treated with either local excision or radical mastectomy, tumor size and differential levels should be considered as potential confounding factors because women undergoing radical mastectomy may have a larger size of tumor or the tumor is less differentiated, which could bias results in favor of local excision. To control this confounding factor we could (1) confine the study to subjects with a narrow range of tumor size and the same differential level, (2) select subjects in such a way that the distribution of tumor size and differential levels are comparable in two groups, or (3) adjust for tumor size and differential levels in the statistical analysis. These strategies, however, cannot adjust for unknown prognostic factors, which is the rationale for randomized trials.

Analytic studies have been a major thrust in population epidemiology. However, their role in clinical research has been somewhat dampened probably because of the stance taken by some clinical epidemiologists that only RCTs are valuable,[5] even though a contrary view has been presented by others.[6] The fact that many clinical situations can only be evaluated by using observational studies, albeit more susceptible to bias and errors, underlines the need for surgeons to become more familiar with the alternatives to RCTs. Understanding principles of observational epidemiology will not only help us to exercise critical appraisal of research reports but also guide us to design and execute studies that are destined to be scrutinized by our peers.

REFERENCES

1. Feinstein AR. Clinical epidemiology, the architecture of clinical research. Philadelphia: WB Saunders Company, 1985.
2. Sackett DL, Haynes RB, Tugwell P. Clinical epidemiology: a basic science for clinical medicine. Boston: Little, Brown, 1985.
3. Schlesselman JJ. Case-control studies. Oxford, United Kingdom: Oxford University Press, 1982.
4. Solomon MJ, McLeod RS. Clinical studies in surgical journals—have we improved? Dis Colon Rectum 1993;36:43-8.
5. Sacks H, Chalmers TC, Smith H. Randomized versus historical controls for clinical trials. Am J Med 1982;72:233-9.
6. Van der Linden W. Pitfalls in randomized surgical trials. Surgery 1980;87:258-62.