

Users' guide to the surgical literature

Case-control studies in surgical journals

Alexandra Mihailovic, MD;*§ Chaim M. Bell, MD, PhD;†§ David R. Urbach, MD, MSc*‡§

Introduction: Some articles in surgical journals identify themselves as case-control studies, but their methods differ substantially from conventional epidemiologic case-control study (ECC) designs. Most of these studies appear instead to be retrospective cohort studies or comparisons of case series. **Methods:** We identified all self-identified "case-control" studies published between 1995 and 2000 in 6 surgical journals, to determine the proportion that were true ECCs and to identify study characteristics associated with being true ECCs. **Results:** Only 19 out of 55 articles (35%) described true ECCs. More likely to be ECCs were those articles that reported "odds ratios" (ORs) (the OR for being an ECC if a study reported "ORs" compared with those reporting no "ORs" 15.3; 95% confidence interval [CI] 2.8–82.6) and whose methods included logistic regression analysis (OR 3.6, CI 1.0–12.9). Studies that focused on the evaluation of a surgical procedure were less likely to be ECCs (OR 0.2, CI 0.1–0.7) than other types of studies, such as those focusing on risk factors for disease. **Conclusions:** The term "case-control study" is frequently misused in the surgical literature.

Introduction : Certains articles dans les journaux chirurgicaux se présentent comme des études «cas-témoin», mais leurs méthodes diffèrent considérablement des modèles traditionnels d'études épidémiologiques cas-témoin (EECT). Il semble que la plupart de ces études soient plutôt des études rétrospectives de cohortes ou encore des comparaisons de séries de cas. **Méthodes :** Nous avons recensé toutes les études s'affichant comme des études «cas-témoin» publiées entre 1995 et 2000 dans six journaux chirurgicaux afin de déterminer la proportion de celles qui étaient vraiment des EECT et de définir les caractéristiques de l'étude qui en faisaient de véritables EECT. **Résultats :** Seulement 19 des 55 articles (35 %) décrivaient de véritables EECT. Les articles les plus susceptibles d'être des EECT étaient ceux qui faisaient état de «coefficients de probabilité» (coefficient de probabilité [CP] que l'article corresponde à la définition d'une véritable EECT si des «coefficients de probabilité» étaient signalés dans l'étude, comparativement à l'absence de «coefficient de probabilité» dans l'étude, 15,3; intervalle de confiance [IC] à 95 %, 2,8–82,6) et ceux dont les méthodes prévoyaient des analyses de régression logistique (CP 3,6; IC 1,0–12,9). Les études qui portaient avant tout sur l'évaluation d'une intervention chirurgicale étaient moins susceptibles de correspondre à la définition d'une EECT (CP 0,2; IC 0,1–0,7) que les autres types d'études, comme par exemple celles s'articulant autour des facteurs de risque d'apparition d'une affection. **Conclusion :** L'expression «étude cas-témoin» est souvent employée improprement dans les écrits chirurgicaux.

Case-control studies are efficient epidemiologic research designs used primarily for studying risk factors for disease. In a typical case-control study, the distribution of 1 or more exposures among a sample

of persons with the disease of interest (the *cases*) is compared with the distribution of the exposures among a sample of persons selected to represent the source population of the cases (the *controls*). Well-designed

case-control studies can provide valid and important evidence relating to many health research questions, primarily risk factors for rare diseases and diseases occurring long after exposure to a potential risk factor.¹

From the *Departments of Surgery and Health Policy, Management and Evaluation, the †Department of Medicine, the ‡Division of Clinical Decision Making and Health Care, Toronto General Hospital Research Institute, the §Clinical Epidemiology and Health Care Research Program, and the Institute for Clinical Evaluative Sciences, University of Toronto, Toronto, Ont.

Accepted for publication Apr. 14, 2004

Correspondence to: Dr. David R. Urbach, Division of Clinical Decision Making and Health Care, Toronto General Hospital, 200 Elizabeth St., Eaton Wing, 9th Fl., Rm. EN 9-236A, Toronto ON M5G 2C4; fax 416 340-4211; david.urbach@uhn.on.ca

We have noticed that some articles in journals targeted to surgical audiences have identified themselves as “case-control” studies but differ substantially from conventional epidemiologic case-control studies (ECCs). Instead of comparing the distribution of exposures among diseased and non-diseased subjects to make inferences regarding *causation*, these studies typically compare the clinical outcomes of groups of subjects treated with different interventions to make inferences about *effectiveness*. These study designs are better classified as retrospective cohort studies or comparisons of case series, and are notoriously prone to inappropriately concluding that one treatment is superior to another.²⁻⁴ The differences between cohort studies and case-control studies are illustrated in Figure 1.

We identified all “case-control” studies published over a 6-year period in 6 surgical journals, to (1) determine the proportion of studies that were true ECCs, and (2) identify study characteristics associated with true ECCs.

Methods

We conducted an electronic search of the Medline database using the text words “case” and “control,” with any suffix allowed after the term “control” (such as “controlled”). We

limited the search to articles published from January 1995 through December 2000, in 6 surgical journals: the *American Journal of Surgery*, *Annals of Surgery*, *Archives of Surgery*, *British Journal of Surgery*, *Journal of the American College of Surgeons* and *Surgery*. These were selected because they are general surgical journals with a broad readership, and their content and article types reflect material commonly encountered in the surgical literature. Titles and abstracts of the articles brought forth by this search strategy were scanned to identify papers with descriptions of the principal method as being “case control,” “case-control,” or “case-controlled.”

In addition to noting the journal of publication and the number of authors, we evaluated each article for whether evaluation of a surgical procedure or disease causation was its focus, logistic regression (LR) analysis was used, and associations were expressed as odds ratios (ORs). Reporting of ORs and use of LR analysis are common features of conventional case-control studies. We selected these explanatory variables *a priori*, because they were factors we thought might be associated with being a true ECC. We had initially planned to examine the effects of funding from peer review agencies and having an investigator with training in clinical epidemiology or

biostatistics on whether studies were appropriately attributed as having “case-control designs.” However, it was impossible to reliably measure these variables in published papers, because the journals varied in how they reported study sponsorship and author affiliations and credentials.

The investigator who abstracted these data did not participate in the determination of whether a study was a true ECC. We used the following definition of a case-control study, by Breslow and Day:

A case-control study (case-referent study, case-compeer study or retrospective study) is an investigation into the extent to which persons selected because they have a specific disease (the cases) and comparable persons who do not have the disease (the controls) have been exposed to the disease’s possible risk factors in order to evaluate the hypothesis that one or more of these is a cause of the disease.⁵

The abstract and methods sections of each article were reprinted in a generic format to blind reviewers to the journal. Studies were classified as “case-control studies” or “non-case-control studies” according to the consensus of the 2 reviewers.

Associations between the characteristics of the reports and whether the report was a true ECC were estimated with LR modeling and expressed as ORs and 95% confidence intervals (CIs). Analyses were done with SAS Release 8.02 for Windows (SAS Institute, Cary, NC).

Results

Our search strategy identified 56 citations; 1 was a meta-analysis of published case-control studies, which was excluded from analysis. Of the 55 remaining, only 19 articles (35%) met our definition of a case-control study (Table 1). Around half of these focused on evaluation of a surgical procedure. Reporting of ORs and use of LR analysis occurred in only a minority. We found that only 35% of the studies in surgical journals that identified themselves as case-control

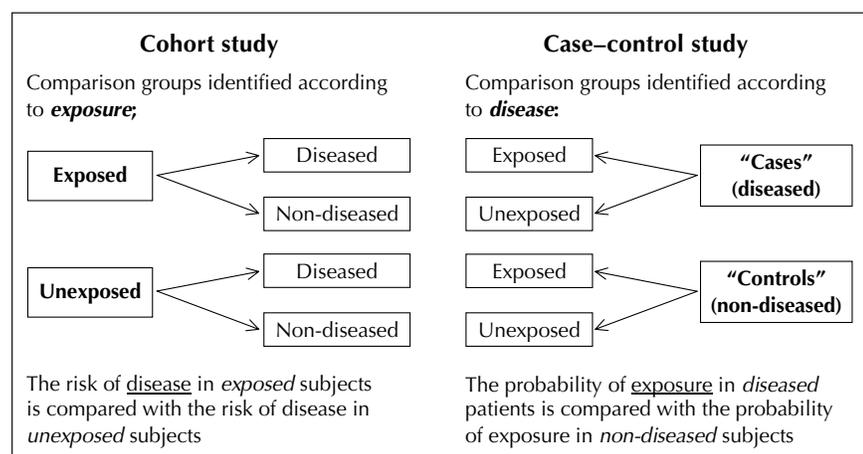


FIG. 1. Comparison of cohort and case-control study designs.

studies were ECCs, whereas 65% were not.

Variation among the journals with respect to whether a study identified as a case-control study was an ECC was not statistically significant; likewise, number of authors. Studies that reported odds ratios (OR 15.3 — that is, the OR for being an ECC if the study “reported ORs” compared with “no ORs reported”; CI 2.8–82.6) and whose methods included LR analysis (OR 3.6, CI 1.0–12.9) were significantly more likely to be ECCs. In contrast, studies that focused on the evaluation of a surgical procedure were less likely to be ECCs (OR 0.2, CI 0.1–0.7) than other types of studies, such as those focusing on risk factors for disease. An example of an article misclassified as a “case-control study” whose focus was a surgical procedure was an article that compared the outcomes of groups of patients undergoing lap-

aroscopic versus open splenectomy.

Discussion

We found that 65% of studies in surgical journals that identified themselves as case-control studies were not true ECCs. True ECC designs were most identifiable by their use of LR analysis and reporting of odds ratios, and by focusing on a problem other than evaluation or comparison of a surgical procedure.

Our findings suggest a persistent lack of appreciation of research methods in the surgical community, despite recent efforts to promote evidence-based surgery. The case-control characteristics we identified provide a framework for researchers, journal editors, peer reviewers and end users of surgical research to predict whether a “case-control study” is properly labelled. However, only a detailed examination of a study’s

methods can provide assurance that its research design is represented appropriately.

Use of LR analysis is a common feature of ECCs, in which controlling for potentially confounding exposures is often an integral part of the data analysis. Consequently, odds ratios, which are easily estimated by LR models, are usually reported as the measures of effect in ECCs. Although use of LR analysis and reporting of ORs may be appropriate in a variety of different study designs, we believe there are several reasons why these factors were associated with ECCs in our study. First, comparisons of case series often try to control for differences between treatment groups by matching, instead of by adjustment in a statistical analysis. Second, investigators using regression analysis in their studies may have had additional training in epidemiology and quantitative methods, and would be more likely to use an appropriate description of a study design. Third and finally, although logistic regression analysis is used in other epidemiologic designs such as cohort studies, we did not find that well-designed cohort studies in surgical journals were misclassified as case-control studies.

Case-control studies are usually employed to study risk factors for rare diseases. While it is possible to assess treatment efficacy using case-control designs, methodologic challenges associated with observational methods such as selection bias, measurement bias and confounding factors make it difficult to use this study design to evaluate therapeutic efficacy. Of the 55 studies we assessed, only 5 of the 28 studies whose primary focus was a surgical procedure used an ECC design.

Since certain research design hierarchies⁶ consider cohort studies to represent a higher level of evidence than case-control studies, some may argue that misclassification of retrospective cohort studies as case-control studies is not a major problem. How-

Table 1

Characteristics of self-identified case-control studies (CCs) in 6 surgical journals, 1995–2000, and factors associated with their being verified as such

Variable	Self-ID'd CCs, no. (& %)	Verified as true epidemiologic CCs*			
		No. (& %)	OR	95% CI	p value
Journal					
<i>American Journal of Surgery</i>	4 (7)	1 (25)	0.3	0–3.8	
<i>Annals of Surgery</i>	4 (7)	2 (50)	0.8	0.1–8.8	
<i>Archives of Surgery</i>	30 (55)	9 (30)	0.3	0.1–1.7	
<i>British Journal of Surgery</i> [†]	7 (13)	4 (57)	1.0	—	
<i>Journal of the ACS</i>	6 (11)	2 (33)	0.4	0–3.6	
<i>Surgery</i>	4 (7)	1 (25)	0.3	0–3.8	
Total case-control studies	55 (100)	19 (35)			
Number of authors					
					0.88
2 or 3 [†]	13 (24)	4 (31)	1.0	—	
4 or 5	18 (33)	7 (39)	1.4	0.3–6.5	
6 or more	24 (44)	8 (33)	1.1	0.3–4.8	
Focus on a surgical procedure					
28 (51)	5 (18)	0.2	0.1–0.7	0.011	
Not on a surgical procedure [†]	27 (49)	14 (52)	1.0	—	
Logistic regression analysis used					
14 (25)	8 (57)	3.6	0.1–12.9	0.045	
Not used [†]	41 (75)	11 (27)	1.0	—	
Odds ratios reported					
11 (20)	9 (82)	15.3	2.8–82.6	0.002	
Not reported [†]	44 (80)	10 (23)	1.0	—	

* Percentages are of self-identified case-control studies; the odds ratio and 95% confidence interval (CI) were estimated for each variable by logistic regression; p values are for the Wald test of the null hypothesis that none of the categories of the variable was associated with being a true case-control study.

[†] The referent category for the regression model.

ACS = American College of Surgeons; CCs = case-control studies; ID'd = identified; OR = odds ratio. Percentages may not add to 100 because of rounding.

ever, cohort studies that are poorly designed or analyzed inappropriately are subject to biases that seriously affect their results, which we found to be the case in many of those reviewed in our study. Many papers compared the outcomes of 2 or more surgical procedures among patients who were selected to have 1 procedure or the other. This selection bias creates a scenario where patients in the comparison groups differ substantially with respect to measurable characteristics such as age, gender and comorbid illnesses, in addition to unmeasured characteristics such as functional and socioeconomic status.⁷ At a minimum, these studies should be analyzed with methods that attempt to adjust for these characteristics.

Does it matter that some important principles of evidence-based medicine and “levels of evidence” have not diffused effectively to surgeons and surgical journals? There are several reasons why we believe it is important for surgeons to understand concepts of research design. First, surgeons involved in health research should have a good understanding of the appropriate research design for addressing study questions, to produce the highest-quality research. Second, study design is often used to assign “levels of evidence” when evidence-based guidelines are produced.⁸ Users of health research

may assume that the results of study designs that are higher on the “design hierarchy” (meta-analyses of randomized controlled trials, randomized controlled trials, cohort studies and case-control studies) are more likely to be valid than those of study designs that are further down (case series and case reports). We have seen examples of the misclassification of nonrandomized comparative studies as “case-control” studies being used to support the validity of studies with relatively weak designs.⁹

Authors and editors of surgical research should be more explicit about how study designs are identified. Descriptors such as “case-control” should be reserved for studies with a conventional epidemiologic case-control design.

Acknowledgments: Dr. Urbach is a Career Scientist of the Ontario Ministry of Health and Long-Term Care, Health Research Personnel Development Program. Dr. Bell holds a Clinician-Scientist Award from the Canadian Institutes of Health Research. This research was funded, in part, by the physicians of Ontario through the Physicians’ Services Incorporated Foundation.

Competing interests: None declared.

References

1. Concato J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. *N Engl J Med* 2000;342(25):1887-92.
2. Sacks H, Chalmers TC, Smith H Jr. Randomized versus historical controls for clinical trials. *Am J Med* 1982;72(2):233-40.
3. Sacks HS, Chalmers TC, Smith H Jr. Sensitivity and specificity of clinical trials: randomized v. historical controls. *Arch Intern Med* 1983(4);143:753-5.
4. Chalmers TC, Celano P, Sacks HS, Smith H Jr. Bias in treatment assignment in controlled clinical trials. *N Engl J Med* 1983; 309(22):1358-61.
5. Breslow NE, Day NE. *Statistical methods in cancer research. Volume 1. The analysis of case-control studies*. Lyon (France): International Agency for Research on Cancer; 1980.
6. Oxford Centre for Evidence-Based Medicine. Levels of evidence and grades of recommendation. Available: www.cebm.net/levels_of_evidence.asp (accessed 2005 Mar. 4).
7. Urbach DR, Bell CM. The effect of patient selection on comorbidity-adjusted operative mortality risk: implications for outcomes studies of surgical procedures. *J Clin Epidemiol* 2002;55:381-5.
8. Upshur REG. Are all evidence-based practices alike? Problems in the ranking of evidence. *CMAJ* 2003;169(7):672-3.
9. Chari R, Chari V, Eisenstat M, Chung R. The author replies [letter]. *Surg Endosc* 2001;15:224. Carbajo Caballero MA, Martin del Olmo JC, Blanco Alvarez JJ. Laparoscopic incisional hernia repair [letter]. *Surg Endosc* 2001;15:223-4. Chari R, Chari V, Eisenstat M, Chung R. A case controlled study of laparoscopic incisional hernia repair. *Surg Endosc* 2000;14:117-9.