Confounding in Observational Studies Explained

Bikaramjit Mann[*,1] and Evan Wood[2,3]

[1]*Department of Medicine, University of Calgary, Canada*

[2]*Department of Medicine, University of British Columbia, Canada*

[3]*BC Centre for Excellence in HIV/AIDS, Canada*

Abstract: Practical and ethical constraints mean that many clinical and/or epidemiological questions cannot be answered through the implementation of a randomized controlled trial. Under these circumstances, observational studies are often required to assess relationships between certain exposures and disease outcomes. Unfortunately, observational studies are notoriously vulnerable to the effect of different types of "confounding," a concept that is often a source of confusion among trainees, clinicians and users of health information. This article discusses the concept of confounding by way of examples and offers a simple guide for assessing the impact of is effects for learners of evidence-based medicine.

Keywords: Confounding, observational studies, critical appraisal, evidence-based medicine.

INTRODUCTION

On the pyramid of evidence in the evidence-based medicine world, meta-analyses of randomized controlled trials (RCTs) sit atop other forms of evidence. In the absence of meta-analyses, RCTs still sit above other study designs since the process of randomization generally accounts for known and unknown confounders being evenly distributed between comparison groups [1]. Unfortunately, however, there are many questions in medicine and public health which will never be answered by a double-blind randomized placebo controlled trial. The affects of smoking and the development of lung cancer is the classic historical example, although there are many other exposure (e.g. pollution) disease associations that have proven causal through assessment in observational studies.

Among observational studies, cohort studies can be thought of as natural experiments in which outcomes are measured in real world rather than experimental settings [2]. As clinicians, we commonly must rely on observational studies to answer questions for which the implementation of a RCT would be inappropriate or impossible, while simultaneously recognizing that observational studies have a greater propensity for bias. Unfortunately, observational studies are notoriously vulnerable to the effect of "confounding," a concept that is often a source of confusion among medical students, residents, clinicians and users of public health information. This article outlines the concept of confounding by way of examples and offers a simple guide for assessing the impact of its effects for learners of evidence-based medicine.

CONFOUNDING

The term "confounding" is used commonly in the medical literature, but when asked to define the concept,

clinicians often encounter difficulty. As will be described below, there is variability in the impact of confounding variables in observational studies. In general, for a variable to be a confounder in a particular study, the formal definition requires that it must meet two criteria: the first is that it must be related to the outcome of interest in terms of prognosis or susceptibility. The second criteria, which explains why it is more commonly seen outside the setting of RCTs, is that the distribution of the confounding factor is different in the groups being compared [3]. More broadly, confounding can be defined as a mixing of effects between the exposure of interest, the disease, and a third factor (i.e. the confounder) that is associated with the exposure that independently affects the risk of developing the disease.

When assessing the impact of an exposure on an outcome, we often summarize the magnitude of association by using an "effect estimate." Depending on the study design and statistical method being used, an effect estimate may be calculated using relative risks, absolute risks, odds ratios, or a hazard ratio to represent the magnitude of association between an exposure and outcome of interest. When considering the potentially confounding effect of the third factor, it is also critical to remember that a confounding variable can create a spurious association, or this variable may be distributed between study groups so as to mask a causal association.

Fig. (**1**) demonstrates this concept by way of a well-known historical example. Many years ago, investigators reported an association between coffee drinking and pancreatic cancer in an observational study [3]. If we take coffee as our exposure of interest and correlate it with an increased development of pancreatic cancer there is the potential, as was the case with these investigators, to be misled if there is a third causal factor, such as cigarette smoking, that was more common among those who reported drinking coffee. With this example in mind, one can easily see how an unmeasured variable – when not evenly distributed between study groups – can easily "confound" a

*Address correspondence to this author at the Department of Medicine, University of Calgary, Calgary, Alberta, T2N 2T9, Canada; Tel: 587-215-3743; E-mail: mannb@ucalgary.ca

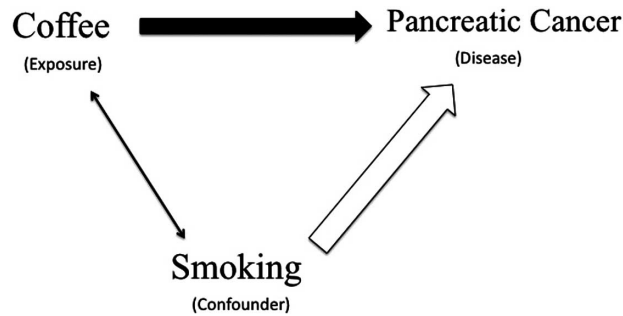study by leading to a spurious association that is solely due to the effect of this third variable.

Fig. (1). demonstrates that coffee (exposure) was associated with an increased risk of developing pancreatic cancer (disease) with the dark arrow. A third factor, smoking, which is a confounder is actually correlated with an increased risk of developing pancreatic cancer (light arrow). Coffee was spuriously correlated with increasing the risk of developing pancreatic cancer and once the confounding variable, smoking is taken into account the correlation between coffee and pancreatic cancer disappears. It should be noted that both smoking and coffee are also both correlated for confounding to occur (two-way arrow).

Post-marketing surveillance has become a powerful tool in determining the safety profile of many medications currently on the market and a prime example may be the association between bisphosphonates and osteonecrosis of the jaw [4-7]. The pathogenesis of osteonecrosis of the jaw is not well understood and has been associated when bisphosphonates are administered intravenously for skeletal complications of malignancy [5]. A similar link has not been demonstrated in patients with osteoporosis when these agents are administered orally and in low doses [5]. However, when considering post-marketing surveillance data one must not overlook the impact that confounding may play. For example, a recent population-based case-control study found that "ever-use" of oral alendronate was associated with an increased risk of incident atrial fibrillation [8]. However, since osteoporosis affects older people who are at higher risk of atrial fibrillation, subsequent studies - which adjusted for age - demonstrated no evidence of an overall long-term increased risk of atrial fibrillation associated with alendronate or risedronate [9].

When potential confounding variables have been measured (e.g. smoking), the possible role of confounding can be assessed when examining for a change in the strength of the effect estimate before and after statistical adjustment. If the confounding variable fully explains the association between an exposure and disease in its entirety, this can reduce the likelihood of residual confounding being present. In the bisphosphonate example above, this would be demonstrated by an estimate of the strength of an association (e.g. odds ratio) going from 5.0 to 1.0 after consideration of the confounding variable (e.g. age). Unfortunately, since there may be a myriad of measured and unmeasured factors in observational studies, the interpretation of confounding is not always this simple.

A decision about the impact of confounding generally assumes that a confounding variable has been perfectly measured because there is also a risk of "residual confounding." In the case of variables like gender or age, exact measurement is usually possible and we can rule out residual confounding. For other confounders, such as cigarette smoking or alcohol consumption, it may be very difficult to accurately estimate the actual exposure. For example, simply asking about previous and current smoking status does not capture the diverse range of tobacco exposures (e.g. cigarettes per day, light *vs* regular cigarettes, filtered *vs* unfiltered, second hand smoke exposure, depth of inhalation, etc). Under these circumstances residual confounding might be suspected when an effect estimate is diminished by statistical adjustment, but the remaining association is felt to be due to imprecise measurement of the confounder or unmeasured confounding variables that were not measured.

A famous example of residual confounding was seen in an evaluation of a needle exchange program in Montreal, Canada [10]. Although needle exchange programs have been shown to reduce the spread of HIV infection among intravenous drug users [11] when this study was adjusted for age, sex, and language it reported a risk of HIV infection among drug users of 2.6 (95% Confidence Interval: 1.7 – 4.0) in comparison to those that did not use the exchange [10]. However, when adjusted for addiction treatment and drug use, it was found that the association diminished to 1.8 (95% Confidence Interval: 1.1 – 2.9), and the effect estimate became 1.7 (95% Confidence Interval: 1.0 – 2.7) after further adjustment [10]. Although there remained a statistically significant 1.7 times higher rate of HIV infection among those using the needle exchange, the overall higher risk profile of exchange users and the difficulty in precisely measuring risk profiles, implies that the remaining difference was a result of residual confounding. In reality, it is often difficult to know if a diminished statistical association is due to partial or residual confounding and the key for practitioners of evidence-based medicine is to examine how an effect estimate changes after statistical adjustment or stratification across a known confounder.

Critically appraising medical literature requires clinicians and public health practitioners to develop a healthy level of skepticism along with an understanding of fundamental epidemiologic concepts. Confounding is an often confusing notion that is often misunderstood by learners and clinicians alike, however it is a real world phenomenon that one must be aware of and account for to avoid spurious results. Without the knowledge of the concept of confounding learners of evidence-based medicine are at a disadvantage in interpreting an association reported in a research study. Observational studies may be vulnerable to the impact of confounding, particularly if investigators are not aware of their implications and how to identify them. Clinicians may be mindful of confounding and also aware of the impact of residual confounding so they may critically appraise the medical literature.

CONFLICT OF INTEREST

None declared.

REFERENCES

[1]   Altman DG, Bland JM. Treatment allocation in controlled trials: why randomise? Br Med J 1999; 318: 1209.

[2]   Rochon PA, Gurwitz JH, Sykora K, *et al*. Reader's guide to critical appraisal of cohort studies: 1. Role and design. Br Med J 2005; 330: 1143.

[3]   MacMahon B, Yen S, Trichopoulos D, Warren K, Nardi G. Coffee and cancer of the pancreas. N Eng J Med 1981; 304: 630-3.

[4]   Edwards BJ, Hellstein JW, Jacobsen PL, Kaltman S, Mariotti A, Migliorati CA. Updated recommendations for managing the care of patients receiving oral bisphosphonate therapy: an advisory statement from the American Dental Association Council on Scientific Affairs. J Am Dent Assoc 2008; 139: 1674-7.

[5]   Khan AA, Sandor GK, Dore E, *et al*. Canadian consensus practice guidelines for bisphosphonate associated osteonecrosis of the jaw. J Rheumatol 2008; 35: 1391-7.

[6]   Edwards BJ, Gounder M, McKoy JM, *et al*. Pharmacovigilance and reporting oversight in US FDA fast-track process: bisphosphonates and osteonecrosis of the jaw. Lancet Oncol 2008; 9: 1166-72.

[7]   Ruggiero SL, Mehrotra B, Rosenberg TJ, Engroff SL. Osteonecrosis of the jaw associated with the use of bisphosphonates: a review of 63 cases. J Oral Maxillofac Surg 2004; 62: 527-34.

[8]   Heckbert S, Li G, Cummings S, Smith N, Psaty B. Use of alendronate and risk of incident atrial fibrillation in women. Arch Intern Med 2008; 168: 826-31.

[9]   Grosso A, Douglas I, Hingorani A, MacAllister R, Smeeth L. Oral bisphosphonates and risk of atrial fibrillation and atrial flutter in women: a self-controlled case-series safety analysis. PLoS One 2009; 4: e4720.

[10] Bruneau J, Lamothe F, Franco E, *et al*. High rates of HIV infection among injection drug users participating in needle exchange programs in Montreal: results of a cohort study. Am J Epidemiol 1997; 146: 994-1002.

[11] Wodak A, Cooney A. Do needle syringe programs reduce HIV infection among injecting drug users: a comprehensive review of the international evidence. Subst Use Misuse 2006; 41: 777-813.