Clin Orthop Relat Res (2011) 469:1794–1796
DOI 10.1007/s11999-010-1538-7

IN BRIEF

Statistics in Brief

How to Assess Bias in Clinical Studies?

Jerome Lambert MD

Published online: 31 August 2010

Background

Owing in part to the development of evidence-based medicine (EBM), the ability to assess study validity and the relative potential for bias is of growing importance. Readers of clinical articles must assess internal and external validity of the study. These two concepts can be understood easily if translated into two questions: (1) can I rely on the conclusions of this study? (internal validity), and (2) can I apply these conclusions to my patients? (external validity).

Since the article by Sacket in 1979 which listed 35 biases [9], numerous articles have proposed listing all biases in clinical studies [1, 2, 7, 10]. In addition to being tedious for the reader, such an approach is difficult because of the number of biases and the absence of consensus regarding their definitions: the same type of biases often have differing names and differing levels of importance.

Question

The purpose of this article is to provide a theoretical framework and practical checklist that will ensure the reader can assess bias in clinical studies.

The author certifies that he has no commercial associations that might pose a conflict of interest in connection with the submitted article.

J. Lambert (✉)
Département de Biostatistiques et Informatique Médicale, INSERM-UMR-S 717, AP-HP, Université Paris 7, Hôpital Saint Louis, 1, Avenue Claude-Vellefaux, Paris Cedex 10 75475, France
e-mail: jerome.lambert@univ-paris-diderot.fr

Discussion

What is a bias and where can one find it? Bias is defined as "any process at any stage of inference which tends to produce results or conclusions that differ systematically from the truth" [9]. Biases sometimes are confounded with random error and cannot be identified. Random error is related to sampling variability, an unavoidable phenomenon that occurs when one deals with a sample of patients instead of the whole population. Random error can be reduced by increasing the sample size. However, some forms of bias can be systematic and independent of sample size. Bias can arise at three steps of the study: during initial enrollment of the participants, during implementation of the study, and during analysis of the findings. If well described, the Materials and Methods section of any article should provide the experienced reader with useful information regarding potential biases of the study. (The well known Levels of Evidence reflect relative risks of bias.)

The first source of bias arises from the absence of a control group in descriptive studies. Descriptive studies, such as cross-sectional studies and case series, select a group of patients based on a particular characteristic (eg, a type of disease or treatment) and describe their evolution, for example, the disease course with a new treatment. Contrary to analytic studies, such as case-control studies, cohorts and randomized controlled trials, there is no control group for comparison. Thus, if a certain recovery rate is observed, it not only can be related to the treatment effect but also to several other parameters. For example, initial characteristics of the patients, natural evolution of the disease, placebo effect and, in the case of a comparison between pretreatment and posttreatment values, regression

toward the mean could partially or totally explain the recovery.

Among analytic studies, there is some agreement on three major categories of bias: selection bias, classification bias, and confounding bias [5].

Selection Bias

Selection bias occurs if the study population does not reflect a representative sample of the target population. Thus, the external validity is questionable and the conclusion drawn by the study should not be extended to other patients. Even more serious bias arises if the selection bias is differential, ie, if the way patients were selected differs between two or more groups. Measures of association (rate ratio, odds ratio, risk ratio) are distorted and the internal validity is questionable. In randomized trials, proper randomization minimizes differential selection bias, although it is frequent in observational studies. For example, certain risk factors may be overrepresented in hospitalized control subjects compared with the general population, and these risk factors may confound the findings independently of the disease in question [8]. For example, a case-control study comparing patients hospitalized for knee arthroplasty attributable to osteoarthritis with control subjects hospitalized with Type II diabetes mellitus probably would not show obesity as a risk factor. Selection bias also can arise during implementation of the study. In observational and experimental studies, when losses or withdrawals are uneven in the exposure and outcome categories, the internal validity of the result may be affected. Such selection bias attributable to losses of followup is called attrition bias.

Classification Bias

Classification bias, also called measurement or information bias, results from improper, inadequate, or ambiguous recording of individual factors—either exposure or outcome variables. Owing to the fact that perfect tools to gather data are uncommon, most studies are subject to a certain degree of misclassification. If the misclassifications occur randomly, the bias is said to be nondifferential. On the contrary, if misclassifications are related to exposure, outcome status, or treatment allocation, the classification bias is differential. In clinical trials, blinding prevents differential classification bias. However, it is a major risk in observational studies, and in particular, case-control studies. For example, certain data, even regarding irrelevant exposures, often are remembered better by patients or/and underreported by control subjects, thus generating a memory, or recall, bias. Classification bias also can occur if different methods of diagnosis are used for the patients. For example, if aseptic loosening is detected by plain radiography for some patients and by arthrography for others, the latter are at greater risk of being classified as having loosening.

Confounding Bias

Confounding bias is a spurious association made between the outcome and a factor that is not itself causally related to the outcome, and occurs if the factor is associated with a range of other characteristics that do increase the outcome risk. Thus, for a characteristic to be a confounder, it must be related to the outcome in terms of prognosis or susceptibility and be unequally distributed among the compared groups. For example, if one tries to establish the effect of age on the risk of knee prosthesis loosening, physical activity could be a confounder. By balancing the different prognosis factors across the groups, randomization partially prevents confounding bias in randomized controlled trials. Randomization is not completely effective, however, because a certain amount of imbalance attributable to chance may occur. Confounding bias is a major risk in observational studies, especially owing to confounders that either are known but not considered or are unknown. Among all biases, confounding bias is the only one which can be corrected after completion of the study by statistical adjustment.

Other Types of Bias

In addition to the three types of bias described above, more specific biases exist that are related only to certain types of studies.

Diagnostic studies can have spectrum bias, a subtype of selection bias. The sensitivity and specificity of a diagnostic test can depend on who exactly is being tested. If only a section of the disease range is included in the study, for example, only the severe type, one may get a biased impression of how well a diagnostic test performs.

Lead-time bias is related to a study assessing the impact of a screening test on survival. Even if people die at the same time with or without early diagnosis, the length of time between early detection of the disease attributable to the screening test can appear as a gain in survival.

A final type of bias is not related to the study but to publication of the results. Numerous articles document the existence of ‘publication bias’: studies with significant results are more easily published than those with negative (ie, nonsignificant) findings [6]. This bias, perhaps more

appropriately called 'negative-outcome bias', can occur at several levels: authors do not submit negative-outcome studies as often, reviewers do not recommend acceptance of negative-outcome studies as often, and editors may not accept negative-outcome studies as often. Even when published, negative studies are quoted less frequently than positive studies [3], as are studies not published in English [4]. Publication and language biases can affect the results of literature reviews and meta-analyses.

Myths and Misconceptions

Two misconceptions are really common about biases: (1) Retrospective studies are more biased than every other type of study whereas randomized controlled trials do not experience bias owing to randomization**.** The presence of bias does not rely solely on the retrospective design, but rather, among other characteristics including the quality of data collection and extraction. In retrospective studies, data on exposition and history often have been collected before the study was performed (eg, in medical records), and therefore might be poorly standardized and more prone to classification bias. Nevertheless, a case-control study using well-standardized data should not experience more bias than a cohort study. Further, randomization does not totally prevent bias. Some biases do not depend on randomization, for example, attrition and classification biases which can be addressed by intent to treat analysis and blinding, respectively. Randomization only partially prevents selection and confounding biases; even if it usually produces comparable groups, a certain amount of covariate imbalance can still occur owing to chance, especially as the sample size is small.

(2) Biases tend to overestimate the treatment effect, or the link between a risk factor and the studied disease. In analytic studies, the null hypothesis that one seeks to reject is that of no difference between groups. In trials, this means no difference in outcome between the different treatments; in a cohort study, it means no difference in disease incidence between exposed and nonexposed patients; and in a case-control study, it means no difference between the histories of patients and control subjects. The bias effect can occur in two directions: either away or toward the null of no difference. When bias is not differential, ie, when opportunities of bias are equivalent in all study groups, the outcome measure is biased toward the null. When the bias is differential, ie, when opportunities of bias are different in the different study groups, the outcome measure can be biased in both ways, either toward or away from the null. For example, the recall bias of a case-control study, described above, tends to overestimate the link between the history of cases and the disease: this is an away from the null bias. Nevertheless, the direction of a differential bias cannot always be predicted.

Conclusions

Every clinical study has some bias. Nevertheless, their presence does not necessarily imply that a study should be disregarded. Rather, it is important to identify and assess the potential impact of biases on the conclusions of the study. Several aspects of the study design, conduct, and analysis should be assessed by the critical reader regarding the three main types of bias: (1) selection bias: are the study subjects corresponding to the target population, and are the subjects in the study arms comparable; (2) classification bias: how was the information regarding exposure and disease collected, and was the information collected in a comparable manner; and (3) confounding bias: has there been a systematic effort to identify and measure potential confounders, and is there information regarding how the potential confounders are distributed between the comparison groups?

References

1. Choi BC, Noseworthy AL. Classification, direction, and prevention of bias in epidemiologic research. *J Occup Med*. 1992;34:265–271.
2. Delgado-Rodriguez M, Llorca J. Bias. *J Epidemiol Community Health.* 2004;58:635–641.
3. Gotzsche PC. Reference bias in reports of drug trials. *Br Med J (Clin Res Ed).* 1987;295:654–656.
4. Gregoire G, Derderian F, Le Lorier J. Selecting the language of the publications included in a meta-analysis: is there a tower of Babel bias? *J Clin Epidemiol*. 1995;48:159–163.
5. Kleinbaum DG, Kupper LL, Morgenstern H. *Epidemiologic Research: Principles and Quantitative Methods*. New York, NY: John Wiley and Sons; 1982.
6. Newcombe RG. Towards a reduction in publication bias. *Br Med J (Clin Res Ed).* 1987;295:656–659.
7. Paradis C. Bias in surgical research. *Ann Surg*. 2008;248: 180–188.
8. Roberts RS, Spitzer WO, Delmore T, Sackett DL. An empirical demonstration of Berkson's bias. *J Chronic Dis*. 1978;31: 119–128.
9. Sackett DL. Bias in analytic research. *J Chronic Dis*. 1979;32: 51–63.
10. Sica GT. Bias in research studies. *Radiology*. 2006;238:780–789.