ISSN (electrónico): 1695-2987 - ISSN (papel): 0003-3170

Angiología

Special Article

Methodological errors. Biases

Errores metodológicos. Sesgos

Antonio Martín Conejero[1], Víctor Quirós González[2]

[1]Department of Angiology and Vascular Surgery. Hospital Clínico San Carlos. Madrid, Spain. [2]Preventive medicine. Office of the 2020-2024 "Transforma 12" Strategy. Hospital Universitario 12 de Octubre. Madrid, Spain

Abstract

One of the most important aspects of research is being able to adequately communicate the results we obtain in our studies. However, reporting on studies that contains errors only serves to distort scientific reality, so it is essential to know which situations can negatively impact the quality of our studies. In this article, as a summary, we present some of them.

Keywords:
Bias. Confusion factor.

Resumen

Uno de los aspectos más importantes en la investigación es el hecho de ser capaces de comunicar adecuadamente los resultados que obtenemos en nuestros estudios. Sin embargo, comunicar sobre un trabajo que contiene errores solo sirve para distorsionar la realidad científica, por lo que es esencial conocer qué situaciones pueden afectar negativamente a la calidad de nuestros estudios. Presentamos en este artículo, a modo de resumen, algunos de ellos.

Palabras clave:
Sesgos. Factor de confusión.

Received: 04/23/2024 • Accepted: 04/23/2024

Conflicts of interest: the authors declared no conflicts of interest.

Artificial intelligence: the authors declare that they did not used any artificial intelligence (AI) or AI-assisted technologies to write this the article.

Martín Conejero A, Quirós González V. Methodological errors. Biases. Angiología 2024;76(4):233-240

DOI: http://dx.doi.org/10.20960/angiologia.00665

Correspondence:

Antonio Martín Conejero. Department of Angiology and Vascular Surgery. Hospital Clínico San Carlos. C/ del Prof. Martín Lagos, s/n. 28040 Madrid, Spain
e-mail: zafraangulo.juan@gmail.com

INTRODUCTION

In this chapter, we will explore some of the ways in which methodological errors can alter our research work.

Let's consider a couple of situations that a researcher might face:

- Imagine we want to conduct a study to compare whether smoking is a risk factor for AAA and for this purpose, we use patients from the urology ward, who smoke as much or more than our patients, as the control group. Will this have consequences on the tobacco-AAA association we are studying?
- And if, when we decide to study the patency of bypass and stent in the occlusive disease of the superficial femoral artery, we see that the percentage of diabetics is much higher in the stent group, will this influence the patency of the comparison groups?

Systematic errors, also called biases, affect the internal validity of the study. That is, they negatively impact our ability to measure what we really want to measure in our study. They prevent us, for example, from knowing whether a risk factor is associated with a disease or whether one treatment is different from another.

However, not all the errors we can make are problems of validity. Some are precision problems related to chance and, therefore, to sample size. Imagine this classic situation, which you are probably familiar with (Fig. 1): we have targets and are asked to throw a few darts with the following results.

In target A, we do not make any errors. We hit the target, which means we have conducted a perfect study. In target D, we are totally wrong. We do not hit one single target correctly; perhaps research is not our thing.

Targets B and C allow us to differentiate a random from a systematic error. In situation B, the researcher knows what the actual target really is; that is, the goal he is pursuing. The problem is that he is unable to reach it. It would be like asking someone who can't shoot to do so. They would shoot anywhere, and if we asked them to shoot many times, they might eventually hit the target. Target C is different. It would be as if we gave the world's best shooter a rigged gun.

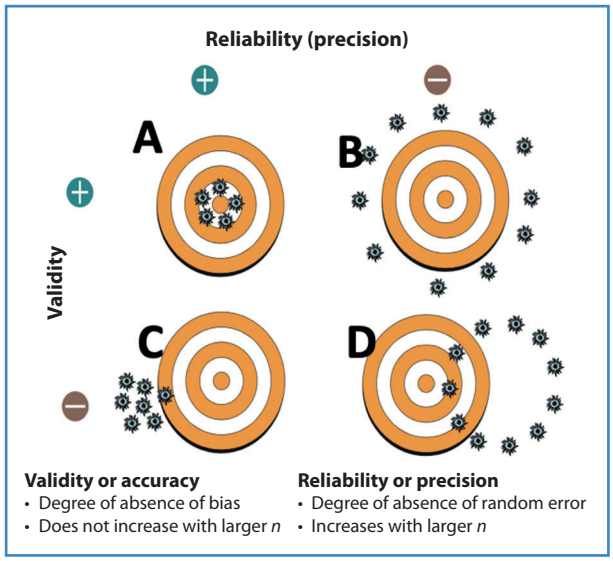

Figure 1. *Difference between random error and systematic error.* n: *sample size.*

He would shoot beautifully but always in the wrong direction. No matter how many times he would shoot, he would always miss in the same direction.

Situation B represents a random error; that is, an error due to chance. If we analyze the relationship between smoking and acute myocardial infarction (MI), a random error would imply saying that we did not find an association between the two. We do not "hit" the target. This is what happens when we have a non-significant p-value or obtain a very wide confidence interval: in both situations, this error can be mitigated by increasing the sample size.

Situation C affects validity. If we analyze the relationship between smoking and AMI, bias could lead to finding a protective association between the two parameters, which we will not correct even if we increase the sample size.

There are many biases we can fall into when conducting an epidemiological study. In this article, we will be looking at some of the most common ones occurring in our routine research, especially when conducting case-control studies.

SELECTION BIAS

Selection bias occurs when subjects are allocated to the control group who significantly differ in some key characteristic from the problem group. This type

of bias can be controlled through a randomization process in the formation of the different study groups. We choose the subjects who will be part of our study poorly, for various reasons.

Some of the selection biases that are most frequently made depending on the type of study conducted are listed below (Table I):

Table I. Selection bias and study type

Prospective designs	Retrospective designs
Healthy worker bias	Neyman's fallacy
Incomplete follow-up of participants	Berkson's bias
Self-selection bias	Diagnostic suspicion bias

Self-selection or volunteer bias

The participation or self-referral of the individual to the study compromises its validity. Imagine a terminal cancer patient asking to participate in a trial with a promising new treatment. What would happen if we included patients with such a poor prognosis in our study? The therapeutic response would not be good, not because the drug is ineffective, but most likely because of the advanced stage of the disease in our patients. That is, volunteer bias implies that the "self-inclusion" of these patients conditions a different therapeutic response than that of the "usual" patient.

Diagnostic bias or Berkson's bias

This occurs when a hospital sample is chosen as the control and the risk factor being studied is associated with a higher probability of hospitalization (Fig. 2). Let's see an example. Suppose we want to analyze the association between alcohol consumption and the development of liver cirrhosis. For this purpose, we select the cases, our patients, from the gastroenterology ward, asking them about their history of alcohol exposure. Logically, most will tell us that, indeed, they abused alcohol. The problem is if we select the controls (healthy without cirrhosis) from the hospital environment such as the psychiatry ward. We ask the patients there about their history of alcohol consumption, so it is possible that many also have a history of this exposure. Many exposed cases (the usual) and many exposed controls (the problem) will condition the minimization of the association between alcohol and cirrhosis, given the equality between cases and controls with alcohol exposure.

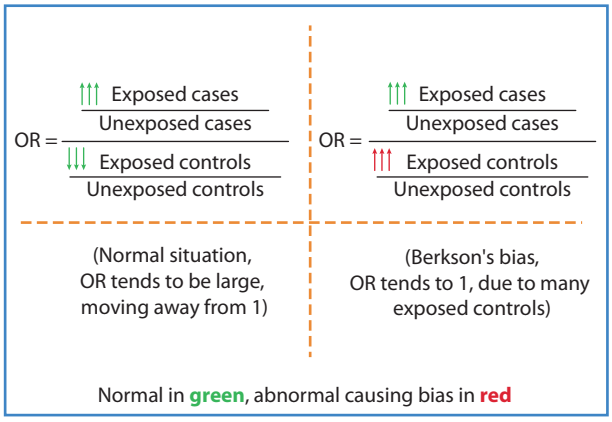

Figure 2. *Influence of Berkson's bias on the OR.*

Healthy worker bias

The exit of the sick worker from the labor market compromises the validity of the study. Imagine we want to study whether there is a higher risk of mesothelioma in workers exposed to insulators such as asbestos. The paradox could be looking for sick people at a factory where asbestos is used and there are no sick people. We would mistakenly conclude that asbestos is not associated with mesothelioma, when what really happens is that there are no patients with mesothelioma working at the time of choosing the sick because they are on sick leave.

Neyman's fallacy

This occurs in case-control studies (Fig. 3) by selecting prevalent (already existing) instead of incident (new) cases. This means that cases are less likely to be exposed to those risk factors that decrease survival; that is, cases would represent individuals with traits of greater resistance to the disease or milder forms of it. For example, if we want to find out the association between mushroom consumption and the occurrence of fulminant hepatitis. If we select patients

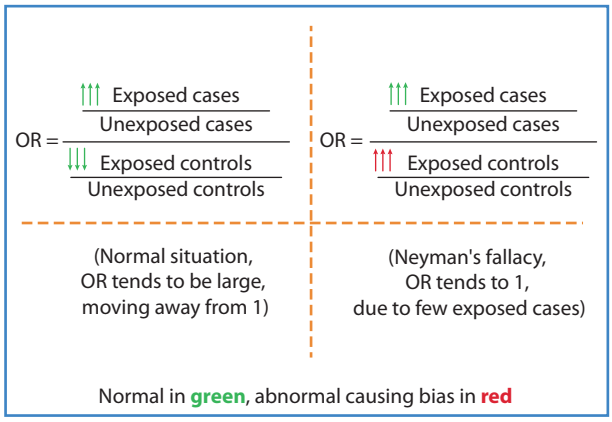

Figure 3. *Influence of Neyman's fallacy on the OR.*

with fulminant hepatitis with a history of mushroom consumption from the hospital, we will paradoxically see that there are few patients. This is due to the rapid lethality of this disease. Again, we will find that the association between the risk factor and the disease is reduced, in this case because we find few exposed cases (due to the lethality of the clinical condition) and few exposed controls (the usual).

Loss to follow-up

It is logical to think that this bias is common in prospective designs, in which subjects are lost in one of the groups. The problem arises when losses or drop-outs can influence the final variable.

Suppose we design a study to evaluate the association between smoking and the development of esophageal cancer. For this purpose, we propose a cohort study in which the exposed are smokers and the non-exposed, non-smokers. We follow them over time expecting the appearance of incident cases. What conditions this bias is that the follow-up is much closer in the smoking group, so that incident cases in that group will not be missed. However, in non-smokers, minimal dysphagia may go unnoticed, so cases (cases not selected) are lost in the non-exposed group.

INFORMATION, CLASSIFICATION, OR RECALL BIAS

In the previous point, we explained what happens when we incorrectly or incompletely select study subjects (selection biases). Now let's put ourselves in a different situation: we have selected our sample perfectly, but we commit a systematic error in the information collection process. Therefore, we define information or classification biases (also found in the literature as observation biases) as the incorrect classification of study participants concerning the variables collected by the researcher. There are multiple reasons for this incorrect classification of subjects. The 3 main ones are:

- *Measurement procedural failure:* This can be a calibration problem with the instrument, the person incorrectly recalling their exposure/disease, or using health records for data collection (especially when they are not computerized).
- *Use of surrogate variables (called proxy variables in epidemiology):* Despite providing greater convenience to the researcher, these variables reduce data validity. For example, to determine a person's purchasing power, would we trust more their self-reported socioeconomic status (proxy variable) or a review of their bank assets and other possessions? Similarly, to know a person's weight, should we ask them (proxy variable) or should we use a scale?
- *Poorly defined variables:* When conducting studies associating obesity with various conditions, we define "person with risk factor" as someone with a BMI > 30 kg/m^2, as established by the World Health Organization (WHO). But, what is the threshold between social classes? And between races? As we see, the threshold is not easy to see in some of the variables.

Incorrect classification may or may not affect all study groups equally. Below, through examples, we will be reviewing the main classification and information biases.

Differential misclassification

The probability of misclassification of exposure (in a case-control study) affects healthy and sick individuals unequally, or the classification of sick or non-sick (in a cohort study) is done differently based on exposure to the study factor. This type of bias underestimates or overestimates the effect of the exposure factor on the disease.

Diagnostic bias. Exposure suspicion bias

Imagine a case-control study at our center, where the principal investigators (PIs) are, aiming to highlight the relationship between the use of certain drugs and the occurrence of cardiovascular events. For this purpose, we'll select patients admitted to our cardiology department with specific diagnoses in the past year as cases. Controls will be healthy individuals, and information will be obtained via telephone interviews. For diagnosed patients (included in the hospital electronic health records), we perfectly review all drug prescriptions made over the past year. Since many controls have not required hospitalization, we will limit ourselves to asking what drugs they have used in the past, so the history of drug exposure (risk factor) will primarily affect the controls.

Recall bias

Having a disease makes one more motivated to remember possible exposure history. Let's consider a situation far from typical research scenarios: a traffic accident. Imagine you live 20 kilometers from the hospital where you work and need your vehicle to get there every morning. One day, at 12:00 PM, I ask you how did your drive to work go. Since you do it every day, you've lost your sense of surprise, and you wouldn't be able to give me any specific details about that morning. Now, let's put ourselves in a more unpleasant situation: that morning, on your way to work, you had a collision at a roundabout that broke your left side mirror. At 12:00 PM, I ask you again: how did your drive to work go? Besides telling me you had an accident, you'll describe the exact position where you entered the roundabout, the color of the vehicle in front of yours, whether the car that hit you used indicators, if it was raining or sunny, etc.

In conclusion, cases always remember exposure better than controls, so the effect is overestimated. The classic example in epidemiology is case-control studies, particularly those relating drug use during pregnancy to congenital malformations in newborns. Mothers of children with these conditions will better remember what drugs they took during that period than mothers of healthy children. Thus, we may classify women of healthy children (controls) as non-exposed.

Attention bias or Hawthorne effect

Between 1927 and 1937, various studies on worker satisfaction were conducted at Western Electric's Hawthorne plant, a power plant based in Chicago, IL, United States within a general American context in pursuit of constant productivity. After introducing various labor improvements, productivity skyrocketed. However, one clever researcher thought the improvements were far greater compared to the few changes made by the company. What was happening? Were they improving because of the company new strategy? The answer was no. They were getting better because, for the very first time, they felt watched and observed. The same thing happens in many of our studies, where study subjects feel observed, which eventually makes them change their exposure, for example, out of fear of repercussions. Imagine that inclusion in a clinical trial depends on whether patients smoke or not. Participants will likely continue smoking, even if they say they wouldn't do so. This means that in our study, we'll be categorizing them as non-exposed when they are, actually, exposed.

Regression to the mean bias

A variable may be extreme in its first measurement for physiological reasons but can return to normal later. For example, the white coat effect, by which a subject is classified as exposed when they are really not. If the measurement is repeated several times, the subject's blood pressure will tend to normalize. This means that a single measurement will incorrectly classify subjects as hypertensive when they are really are not.

Non-differential misclassification

The probability of misclassification occurs similarly across all study groups. Non-differential misclassification causes an underestimation (brings the association measure closer to 1) of the effect of the exposure factor studied in the disease. Imagine we want to study the association between hypertension and stroke. To do this, we use a blood pressure cuff in our clinic and divide the sample into

hypertensive and non-hypertensive based on a diastolic pressure > or < 90 mmHg. What happens if the cuff sometimes erroneously measures high values and sometimes low values? The result will be that some normotensive patients will end up in the hypertensive group and some hypertensives will fall into the normotensive group. This equalizes both groups, underestimating the association between the risk factor and the disease.

CONFOUNDING FACTOR

One of the most important factors due to its presence in research is undoubtedly the concept of confounding, defined as the distortion of the estimated effect of the exposure on the disease due to the introduction of a new effect caused by a strange factor (a so-called confounding factor) that is not intended to be studied and causes a mix of effects. In other words, we have a strange variable to the study that modifies the results obtained. Every confounding factor must meet 3 requirements (some examples are the *tobacco risk factor,* the *alcohol confounding factor,* and *disease esophageal cancer*):

- Be a risk factor for the disease. Alcohol, the confounding factor, is associated with the disease.
- Be associated with the exposure. Alcohol is more common in smokers vs non-smokers.
- Not be an intermediate step between exposure and the disease. Alcohol is not a link between tobacco and cancer.

Despite being an error with significant implications for our results, confounding bias can be quantified (unlike selection and information biases). Therefore, we will directly compare the crude or raw effect estimate (ORb, in the total sample) with the estimate without confounding (ORe, in the stratified group). We can calculate the size of the confounding using the formula:

$$\frac{ORe - 1}{ORb - 1}$$

Let's imagine that for the association between a risk factor and a disease, an OR of 3 is obtained. When stratifying into 2 categories based on a variable that could be distorting the effect, 2 ORs of 2 are obtained. Applying the formula:

$$\frac{ORe - 1}{ORb - 1} = \frac{2 - 1}{3 - 1} = \frac{1}{2} = 0.5$$

The size of the confounding is $(1 - 0.5) \times 100 = 50\ \%$ $(1 - 0.5) \times 100 = 50\ \%$ of the effect. That is, the true or real effect, once confounding is controlled, is 50 % less (just half) than the estimate obtained from the crude analysis, which was indeed distorted. We consider that confounding is sufficiently important to need to control it when the crude OR differs from the adjusted one (or in our case, the *ORe*) by more than 10 %, as is the case here.

To prevent confounding biases, various techniques exist, applicable both in the design phase and in statistical analysis. This is one of the great advantages of this systematic error: it is controllable afterward. In the design phase, we find randomization (experimental studies), matching, and restriction. In the statistical analysis phase: stratified analysis (categorizing global data into 2 groups based on the presence or absence of the confounding factor, as we will see later on) and multivariate analysis or regression techniques, which are beyond the scope we want to address in this article. It is important to assume that perfection does not exist in research methodology. There will always be residual confounding, understood as that which persists after failed attempts to control it.

Closely related to confounding, other changes of the estimated effect should be known: *interaction*. Unlike the confounding factor, interaction does not change the estimated effect incorrectly. Among the main common points, we find the need for assessing external variables to achieve a correct interpretation of our study. Let's see everything discussed above in the example illustrated by figure. 4.

We wish to study the association between esophageal cancer and smoking. To do this, we design a 2 × 2 table with the variables "cancer (yes / no)" and "tobacco (smokes / does not smoke)". After calculating the cross-product ratio, we obtain an OR, which we will call the crude OR, of 2.5.

However, we become quite concerned when reviewing the literature, as several studies mention

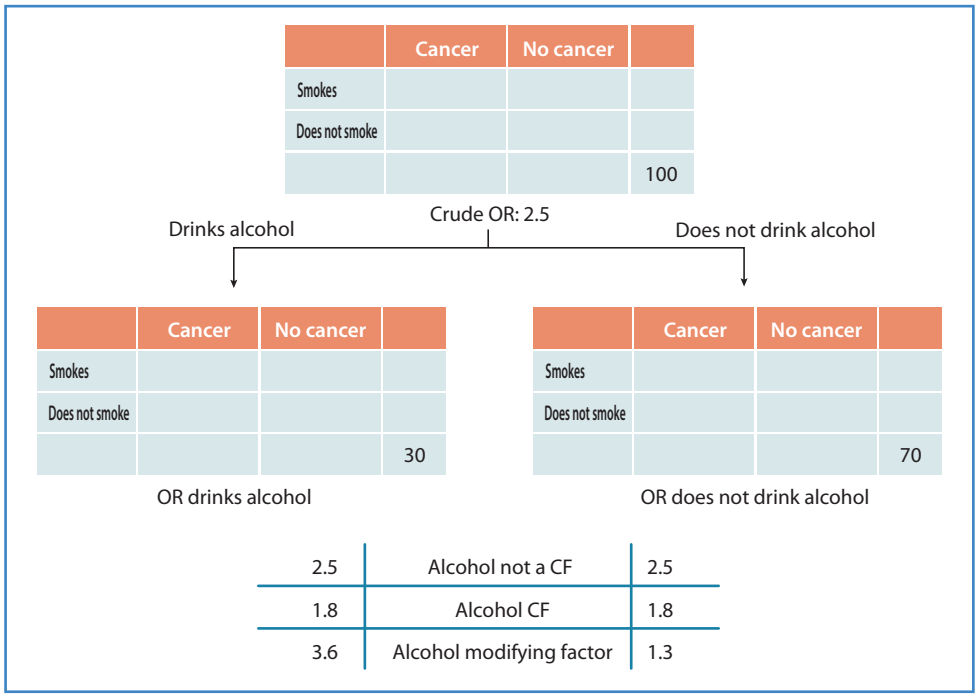

Figure 4. *Stratified analysis.*

the impact that alcohol consumption may have on this relationship. It could be a confounding factor in our study.

To clear up any doubts, we decide to stratify based on whether participants have the supposed confounding factor (in this case, alcohol consumption). We obtain 2 new tables that can lead us to 3 different situations:

- If the crude OR we had estimated is maintained when stratified (OR = 2.5 in the alcohol group and OR = 2.5 in the non-alcohol group), it means nothing is happening. More methodologically: alcohol is neither a confounding factor nor an effect modifier in our study.
- If the new ORs, one for each stratum, are 1.8 (that is, lower than the crude OR we initially calculated), we are facing a clear confounding factor. We will have to remedy this, or our study will be seriously compromised. Applying everything we've learned, let's quantify the problem using the formula:

$$\frac{ORe - 1}{ORb - 1} = \frac{1.8 - 1}{2.5 - 1} = \frac{0.8}{1.5} = 0.5$$

The size of the confounding is (1 – 0.53) × 100 = 47 % (1 – 0.53) × 100 = 47 % of the effect. That is, the true or real effect, once confounding is controlled, is 47 % less than the estimate obtained from the crude analysis, which was indeed distorted. Being higher than the agreed limit of 10%, we will consider that the confounding factor is invalidating the results. 3.

If the new ORs are lower than the crude OR, but with different values in both groups, we are dealing with an effect modifier. To recap, we spoke of interaction when the effect of the exposure on the disease is different at various levels of a third variable. In our example, we would say that tobacco always acts as a risk factor since the OR is > 1 in both strata. However, we will affirm that alcoholics have 3.6 times more risk from tobacco, while non-alcoholics have only 1.3 times more risk of cancer from tobacco. As we see, alcohol modifies the risk of cancer associated with tobacco.

If our readers were able to go through all of the above, that means that they are truly interested in this topic and should read books that address methodological errors in greater detail.

In this article, we wanted to provide just a few brushstrokes of some of the situations most frequently encountered in the routine of an aspiring researcher.