SPECIAL ARTICLE: EDUCATION

Errors and biases in clinical research☆

Errores y sesgos en investigación clínica

D.M. González de la Cuesta [a,b]

[a] *Instituto de Investigación Sanitaria Aragón-IISA, Hospital Universitario Miguel Servet, Zaragoza, Spain*
[b] *Departamento de Fisiatría y Enfermería, Facultad de Ciencias de la Salud, Universidad de Zaragoza, Zaragoza, Spain*

Introduction

A bias is the ''obliquity or twisting of a thing to one side, or in the cut, or in the situation, or in the movement'' according to the dictionary of the Royal Spanish Academy and refers to the deviations that occur in the usual practice in any aspect of life.

Thus, one can speak of different types of biases: cognitive, statistical, contextual, law enforcement bias, media bias, conflict of interest and even prejudice.[1] In many of these cases, a bias is something harmful and negative and we must be alert to try to neutralise them should we become aware of them; however, in some cases a bias can be positive, as in the case of a cognitive bias, where our brain, faced with a shortage of information or time to make a decision that involves our survival, makes irrational decisions, that which we call ''intuition'' and allows us to move away from the collision course of a vehicle, for example.[2]

Biased ideas or thoughts see only one side of reality, one side or part of it, and therefore lack impartiality.[3]

However, when this whole process is taken to clinical research, any deviations that may occur at any point in the process compromise the results of the research and, therefore, the conclusions reached.

Research

Any research process deals with answering the question posed in a valid and precise way, without errors. It is about measuring what you want to measure, and measuring it properly. That is to say, to guarantee the validity of the conclusions, since these research results are those that will be applied in clinical practice, and to vouch for their validity and reliability.[4–6]

Validity is the ability to actually measure what it seeks to measure, it expresses the degree to which the phenomenon of interest is actually measured. There are variables that are more valid than others for measuring a given phenomenon; for example, the glycaemic control of a diabetic patient is better observed with the measurement of glycosylated haemoglobin than with an isolated measurement of glycaemia.

The validity of a study consists of both internal and external validity.

Internal validity refers to the degree to which the results of a study are free of error for the sample studied; it indicates the intrinsic quality of a study, and its main threats are systematic errors and confounding factors.

In contrast, *external validity* refers to the degree to which the results can be generalised to populations other than those studied (the target population).
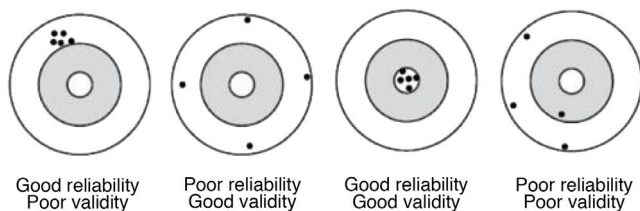
Figure 1 Validity is synonymous with accuracy; reliability is synonymous with precision.
Source: ''Método epidemiológico''. Escuela Nacional de Sanidad (ENS) Instituto de Salud Carlos III - Ministerio de Ciencia e Innovación.

Reliability or precision indicates the extent to which the same values are obtained when the measurement is made on more than one occasion and under similar conditions; i.e. it expresses the degree of reproducibility of a measurement procedure.

Measurement always involves some degree of error. Errors in measurement may be due to factors associated with individuals, the observer or the measuring instrument, and therefore there may be variations in measurements. For instance, in the measurement of body temperature there may be errors due to the patients' condition (agitation, blinding); the thermometer used may be faulty, or the observer may make a reading, transcription, or rounding error that differs from another observer.

The accuracy of a measurement does not guarantee its validity. For example, if 2 consecutive measurements of a patient's blood pressure are made with a poorly calibrated sphygmomanometer, the values obtained will be similar (the measurement will be reliable), but totally inaccurate (therefore invalid) (Fig. 1).

Systematic and random errors

Therefore, it can be said that there are 2 types of errors, systematic errors and randomised errors (Table 1).

Randomised errors consist of a divergence between an observation made in the sample and the true value in the population. It is due to chance, and occurs for two reasons: because we are working with samples and not with entire populations (and this gives rise to a degree of individual variability) and because of the variability inherent in the measurement process, both in the instrument used and in the observer.

In the first case, working with samples, this can be minimised by increasing the sample size and using randomised sampling.

In the second case, variability due to the measurement process may be attributable to measurements that change throughout the day, known as biological variability (average blood pressure varies over the course of the day because of circadian rhythms), which would be mitigated by taking several measurements and using averages This can also be a function of the instrument used or the observer; in these cases, in order to reduce it, measurements must be standardised and researchers must be well trained in how to measure each variable.[7]

Randomised error is closely related to the concept of precision.

Systematic error is what is actually known as ''bias''; it is an error in the design of the study that leads to an incorrect estimate of the effect or parameter being studied.[8]

There are 3 classes of biases or systematic errors (Table 2):

- Selection biases: in the selection of subjects.
- Information biases: in measuring the variables.
- Confounding bias: this occurs when there are variables that alter the relationship between the dependent and independent variables, and lead to confusion in the interpretation of the results obtained.

Biases affect the study's validity, and their effect is not modified by increasing the sample size, as the error is to be found in the design itself and cannot be controlled for in the analysis. In these cases, only the direction of the bias can be estimated: to know its possible effect on the observed results.

Selection biases

Looking at it more slowly, selection bias occurs when a sample is selected in a study that is not representative of the target population. It occurs when some subjects are more likely to be selected than others, e.g., when choosing sick individuals admitted to a hospital, the most seriously ill are selected. It affects external validity: the results may not be applicable to subjects with the disease (diabetes) in follow up with primary care.

To prevent selection bias, probability sampling should be used for the selection of subjects for the study.

In general, they occur in the following situations:

- *Biases in the selection of the control group:* In cohort studies, the exposed and control cohorts must be similar in all but the factor of exposure under study; in clinical trials, randomisation makes the groups very likely to be similar. Biases in the selection of the control group occur especially in case-control studies and in retrospective studies if the control cohort is not similar to the case cohort.
- *Loss-to-follow-up bias:* In longitudinal studies, this bias occurs when subjects are lost to follow-up who are more likely to develop the outcome of interest than those who are not (e.g. in a study of cardiovascular disease, where more subjects are lost to follow-up among smokers than among non-smokers).
- *Loss to follow up on-response bias:* This occurs in surveys and cross-sectional studies when there is a suspicion that individuals who respond to these surveys have different characteristics from non-respondents.
- *Selective survival bias:* This bias is present when newly diagnosed, more benign, or milder cases, which have longer survival rates, are included. In this case, the sample is not representative of the full spectrum of disease severity and the results cannot be transposed to all those affected by the disease.
- *Non-representative sample bias or Berkson bias:* This is more common in cross-sectional studies, if the sample

Table 1 Difference between randomised error and systemic error.

	Random error	Systematic error (bias)
Cause	• Sampling • Parameter variability	Design, execution, and analysis: • Selection of study subjects • Information collection • Presence of distorted external variables
Decreases by increasing sample size	Yes	No
Affecting	Accuracy	Validity
Assessment	Statistical concept linked to informed judgement elements	The main elements involved are informed judgement

Source: «Método epidemiológico». Escuela Nacional de Sanidad (ENS) Instituto de Salud Carlos III - Ministerio de Ciencia e Innovación.

Table 2 Classification of biases.

Biases	
Selection biases	Control group selection bias
	Loss to follow-up bias
	Non-response bias
	Relative survival bias
	Berkson's bias
	Detection bias
	Self-selection bias
Information biases	Non-differential misclassification
	Differential misclassification
	Memory, amnesia or recall bias
	Unacceptability bias
	Interviewer bias
Confounding bias	An effect is attributed to a variable without being due to it.

Source: Created by the authors.

does not represent the target population; for example, if we take only subjects admitted to hospitals or from health centres, etc.

- *Detection error biases:* These biases affect clinical trials, particularly when the response is evaluated differently according to the treatment group. To avoid them, it is very important that the evaluator be blinded to the treatment group.
- *Volunteer participation bias:* This happens when the volunteers may have a different profile to those who do not participate. It is a self-selection mechanism.

Information biases

These occur when information about the study factor or response variable is collected erroneously or has been collected differently among different study groups, if any. It affects both internal and external validity.

They may be due to the use of inappropriate measurement instruments, imprecise definitions or errors of the enumerators or respondents.

Basically, we will discuss 2 types of errors:

- Non-differential classification error: This occurs when the proportion of misclassified subjects is similar in each of the study groups; for example, if an insensitive instrument is used to measure the main variable in the different groups under study, it leads to an underestimation of the true association and can give rise to discrepancies between the results of different studies. It is a bias of lesser importance than differential classification biases.
- Differential misclassification error: In this case, the proportion of errors in the classification of disease and exposure is not the same in the different study groups; examples include the following types of biases.
  - *Memory bias:* individuals with a health problem have better recall of their exposure history than those without a health problem. This bias is not uncommon in retrospective studies and in case controls.
  - *Interviewer bias:* This arises if there is a systematic difference in how data are collected or interpreted from study participants depending on the group to which they belong.
  - *Unacceptability bias:* This is the result of when study subjects have misgivings about certain exposures that are socially frowned upon, such as excessive alcohol consumption or the use of certain substances.

Conclusion

Biases can occur at all stages of a research project, from the literature review, to selecting only articles published in a particular language, to analysing data with incorrect statistical tests, to not publishing the results because you do not like the data obtained.

Although not all biases can always be avoided, at least every effort should be made to control and minimise them and, above all, to be aware of them.

One must be very careful in planning studies, since mistakes can always be made. Some errors can be overcome in the statistical analysis, for instance, but others cannot be fixed and can distort the results to the point of being inadmissible as evidence.

References

1. Kahneman D, Tversky A. Subjective probability: a judgment of representativeness. Cogn Psychol. 1972;3:430–54, http://dx.doi.org/10.1016/0010-0285(72)90016-3.

2. Baron J. Thinking and deciding. New York, NY: Cambridge University Press; 2007.

3. Ariely D. Predictably irrational: The hidden forces that shape our decisions. New York, NY: HarperCollins; 2008.

4. Abraira V. Sesgos en los estudios sobre pruebas diagnósticas. Semergen. 2006;32:24–6.

5. Rodriguez JL, Casado A. Doble ciego. El control de los sesgos en la realización de ensayos clínicos. Contradicciones, insuficiencias e implicaciones. Med Clin (Barc). 2002;118:583–4.

6. Lobiondo-Wood G, Harber J. Nursing research: methods, critical appraisal and utilization. St Louis: Mosby; 2006.

7. Sánchez P, Aguilar I, Fuentelsaz C, Moreno MT, Hidaglo R. Fiabilidad de los instrumentos de medición en ciencias de la salud. Enferm Clin. 2005;15:227–36.

8. Manterola C, Otzen T. Los sesgos en investigación clínica. Int. J. Morphol. 2015;33:1156–64.