



CIRUGÍA ESPAÑOLA

www.elsevier.es/cirugia



Methodological letter

Validity of a study: Internal and external validity[☆]



Validez de un estudio: validez interna y externa

Maria Sala Serra,^{*} Laia Domingo Torrell

Servicio de Epidemiología y Evaluación, Instituto Hospital del Mar de Investigación Médica, Barcelona, Spain

If we conduct a study to find out whether one new surgical technique is better than another, we need to know whether the observed differences between the two techniques are due to chance or due to study design and conduct in order to rule out alternative explanations for the observed effects. Or the other way around, if we do not observe differences, we can be sure that they do not exist because there really is no causal relationship. Because when faced with the results of a study, we have to ask ourselves whether they are correct, without error or bias, whether they can be attributed to chance, and whether they are applicable to other contexts¹.

Two types of study validity have been defined: internal and external validity.

Internal validity assesses the degree to which the results obtained in a particular study are correct for the individuals studied in this study. It determines whether the observed outcome can be attributed to the new surgical technique being evaluated. There are two main errors that could threaten the internal validity of a study. They are biases (also called systematic errors) and random errors.

External validity (or generalisability) assesses the degree to which the conclusions drawn from a study can be extrapolated or generalised beyond the sample population studied. It depends on the size and characteristics of the sample and the context in which it is to be applied.

Lack of internal validity negatively influences the quality of evidence that can be derived from a study. A study with internal validity may or may not have external validity, but a study without internal validity cannot have external validity.

Errors limiting internal validity

Systematic errors

Systematic error is an error that consistently occurs in every measurement. It is not due to chance, nor does it depend on the sample size. It occurs when an error is introduced in the design of the work, either in the selection of individuals, in the information collected, or in its analysis and interpretation. Systematic biases can lead to an overestimation or underestimation of the true effect of an intervention or treatment under study².

One of the most widespread classifications proposes to group biases into three broad categories: selection, information and confounding³.

Selection bias: the sample does not represent the target population to be studied. They occur mainly in case-control studies and retrospective cohort studies. However, loss to follow-up of participants in these studies can also lead to selection bias in prospective cohorts. On the other hand, randomised controlled clinical trials (RCTs) are also susceptible to selection bias, depending on when the patient is informed and invited, whether before or after randomisation and the possibilities to give informed consent^{4,5}. Failure to take into account non-consenting patients who are part of the non-response rate is another cause of selection bias. So is not taking into account characteristics that may be associated with prognosis. Randomisation may minimize this problem.

[☆] Please cite this article as: Sala Serra M, Domingo Torrell L. Validez de un estudio: validez interna y externa. Cir Esp. 2022;100:649–651.

^{*} Corresponding author.

E-mail address: msalaserra@psmar.cat (M. Sala Serra).

Information bias: non-random errors in the mediation of information. It occurs when the necessary information is collected in a systematically different way between study groups. This category includes interviewer bias, respondent bias, recall bias, etc. All study designs are susceptible to information bias.

We can introduce bias, for example, by paying more attention to patients operated on with a new technique and that this results in fewer or less serious complications because they are detected earlier. Or, conversely, that by being more closely observed, more complications are detected and recorded. If the way of obtaining or interpreting the information differs depending on the group of patients, especially if the person obtaining or giving the information is aware of the hypotheses and objectives of the study, there is a possibility of information bias. A patient who is aware of the treatment he or she is receiving may be more aware of side effects and report them to a greater extent.

These biases can be partly avoided with good protocols that define comprehensively what information is to be recorded and how, including masking techniques, and applying them exactly the same for all participants.

However, studies in surgery have the added difficulties inherent to the discipline⁶, including the selection of a control group (often involving historical groups that are difficult to compare); the difficulty of giving placebo; the difficulty of carrying out double-blind trials; the effect of the caregiver or perioperative care; the experience of the surgeon, etc.⁷. Although there are strategies to minimise these biases, such as performing sham operations, blinded outcome assessment or the use of a second surgical team, they are not always possible, so a balance must be sought between the risk of bias and the feasibility of conducting the study under the best possible conditions².

Confusion

Confounding occurs when the observed association is due, at least in part, to differences between the groups studied, other than the exposure or intervention under study, that independently affect the risk of the outcome of interest. It can be limited during study design by using techniques that avoid imbalances between study groups on potentially confounding variables. Matching, restriction and especially randomisation are tools to prevent confounding in study design. And stratified analysis and multivariate analysis allow to control for it in the analysis.

Random errors and reliability

Random error is due to the fact that we work with samples of individuals and not with the whole population. It arises from the inherent variability of sampling and depends on the sample size: as the sample size increases, the error decreases. Statistics allows us to quantify random error and we speak of statistical power.

External validity

External validity refers to the extent to which the results of observational trials or studies provide a sound basis for generalisation to other populations or other clinical circumstances. Internal validity is a prerequisite for external validity. The results of an internally valid study may not be applicable to other populations, either because the sample is not representative due to biases or restrictions in inclusion criteria, or because the methodological conditions of the study cannot be easily replicated in routine practice, or because of other factors such as the patient-physician relationship, the information patients have about the treatment they are receiving, or their values and preferences. Furthermore, people (patients and researchers) react differently when observing or being observed in the context of a study than when they are in normal conditions^{8,9}. Other problems limiting the generalisability of results are due to factors affecting the interpretation and incorporation of evidence into decision-making, as described by García-Alamino and López Cano in their methodological letter (publication pending in *Cirugía Española*).

Conclusions

The assessment of internal validity involves ruling out sources of bias and random error and is concerned with methodological quality. External validity is concerned with assessing the applicability of the results to the real-world population, and includes methodological aspects, but also the setting or context. Reporting guidelines have been developed to improve the quality of studies, including CONSORT (for clinical trials), PRISMA (for systematic reviews), QUADAS and START (for diagnostic tests) and STROBE (for observational research). There are new proposals, some in the validation phase, adapted to specific research settings, including surgery.

REFERENCES

1. Kwaan MR, Melton GB. Evidence-based medicine in surgical education. *Clin Colon Rectal Surg.* 2012;25:151-5. <http://dx.doi.org/10.1055/s-0032-1322552>.
2. Gurusamy KS, Gluud C, Nikolova D, Davidson BR. Assessment of risk of bias in randomized clinical trials in surgery. *Br J Surg.* 2009;96:342-9. <http://dx.doi.org/10.1002/bjs.6558>.
3. Rothman KJ, Greenland SLT. *Modern Epidemiology*, 3th ed. Philadelphia: Lippincott Williams & Wilkins; 2008.
4. Schulz KF, Grimes DA. Allocation concealment in randomised trials: defending against deciphering. *Lancet.* 2002;359:614-8. [http://dx.doi.org/10.1016/S0140-6736\(02\)07750-4](http://dx.doi.org/10.1016/S0140-6736(02)07750-4).
5. Hotter B, Ulm L, Hoffmann S, Katan M, Montaner J, Bustamante A, et al. Selection bias in clinical stroke trials depending on ability to consent. *BMC Neurol.* 2017;17:206. <http://dx.doi.org/10.1186/s12883-017-0989-9>.
6. Demange MK, Fregni F. Limits to clinical trials in surgical areas. *Clinics (Sao Paulo).* 2011;66:159-61. <http://dx.doi.org/10.1590/s1807-59322011000100027>.

-
7. Mcculloch P, Taylor I, Sasako M, Lovett B, Griffin D. Randomised trials in surgery: problems and possible solutions obstacles to randomised trials in surgery. *BMJ*. 2002;324:1448-51.
 8. Nguyen VN, Miller C, Sunderland J, McGuiness W. Understanding the Hawthorne effect in wound research—a scoping review. *Int Wound J*. 2018;15:1010-24. <http://dx.doi.org/10.1111/iwj.12968>.
 9. Sedgwick P, Greenwood N. Understanding the Hawthorne effect. *BMJ*. 2015;351. <http://dx.doi.org/10.1136/bmj.h4672>.