CIRUGÍA ESPAÑOLA

www.elsevier.es/cirugia

Methodological letter

Assessment of heterogeneity in meta-analysis☆

Evaluación de la heterogeneidad en el metaanálisis

Guillermo Lillo-Albert,[a] *Andrea Boscà i Robledo,*[b] *Salvador Pous-Serrano*[a,*]

[a] *Unidad de Cirugía de Corta Estancia y Pared, Hospital Politècnic i Universitari La Fe, Valencia, Spain*
[b] *Unidad de Cirugía Hepatobiliopancreática y Trasplantes, Hospital Universitario y Politécnico La Fe, Valencia, Spain*

One of the greatest challenges and limitations of meta-analysis is deciding which studies are or are not eligible. Clearly combinable are those in which the research hypothesis completely coincides. Unfortunately, studies differ for various reasons: in the definition and severity of the pathology; the study country; the study design characteristics, and the participating populations, etc., with the last two being of greatest significance.[1–3]

The concept of heterogeneity arises from these differences and may be defined as the inter-study variability existing between the effect size estimates.[2–4]

Study design characteristics may have an effect here[1,3]:

1) The intervention assignment process: a) the selection of the cohort through pragmatic studies, which have high external validity, or explanatory studies with low internal validity; b) randomisation; c) masking; d) the type of analysis, which, if by protocol, overestimates the differences or if by intention to treat, underestimates them.
2) Monitoring of patients during the study.

The characteristics of the study's participating populations may also have an effect[1,3]:

1) Variations in the same surgical technique not reflected in the studies.
2) The representation of special risk subgroups.
3) The type of effect related to the exposure or intervention. The same intervention can have different results or effects due to patient variability.

As a result, all meta-analysis must analyse the amount of heterogeneity existing among the included studies. There are several statistical and graphical methods to evaluate the degree of heterogeneity.

Regarding statistical methods, the most used are the following[1–5,6]: (Fig. 1)

- **Cochrane's Q test**: is a statistic based on a Chi-square test that considers the deviations between the results of each study and the global result, according to the weight of the study with respect to the global result. Its limitations include loss of power when the number of studies is small. It thus tends to be compensated by considering it statistically significant and heterogeneous from p values under .10.
- I^2: expresses the percentage of heterogeneity in the meta-analysis, with values between 0% (homogeneous) and 100% (totally heterogeneous). Values below 25% are considered to

☆ Please cite this article as: Lillo-Albert G, Boscà i Robledo A, Pous-Serrano S. Evaluación de la heterogeneidad en el metaanálisis. Cir Esp. 2024. https://doi.org/10.1016/j.ciresp.2024.03.009

* *Corresponding author.*
E-mail addresses: pous_sal@gva.es (S. Pous-Serrano).

http://dx.doi.org/10.1016/j.cireng.2024.03.010

2173-5077/© 2024 AEC. Published by Elsevier España, S.L.U. All rights reserved.

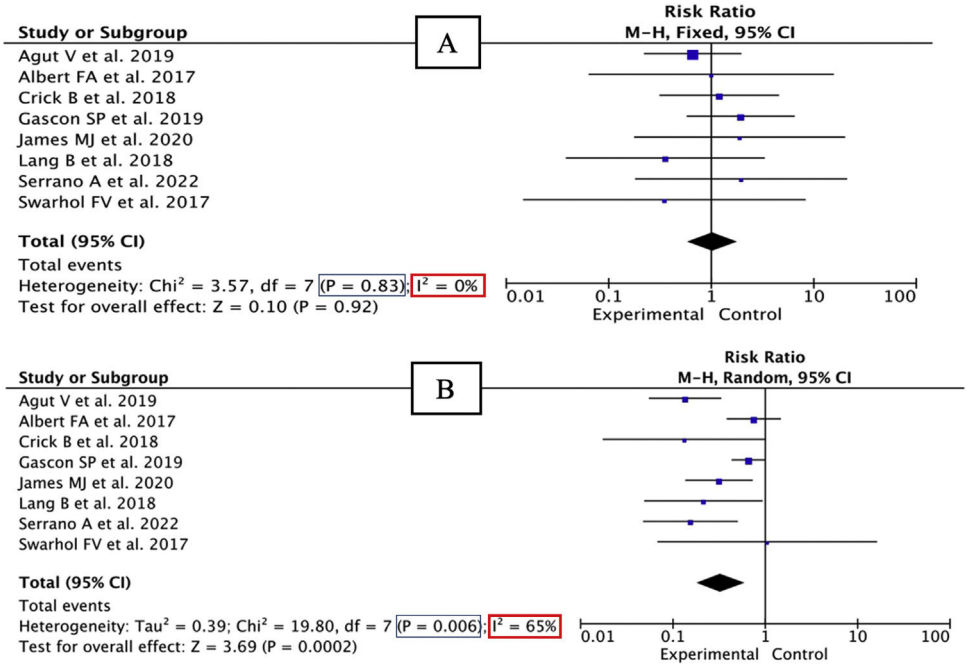

Fig. 1 – Statistical interpretation and Forest Plot of heterogeneity using simulated data and authors. A) Homogeneous study (Cochrane's Q *P* > .1 and I^2 = 0). B) Heterogeneous study (Cochrane's Q *P* < .1 and I^2 with intermediate values).

indicate low heterogeneity and values above 75% to indicate high heterogeneity. If values are between 25%–75% Cochrane's Q test may be used as support.

Regarding graphical methods, the most used are the following[1–5]:

- **Forest plot or diagram**: if the confidence intervals of the effect measure used in the different studies do not overlap each other, it is heterogeneous. In other words, if the graphical representation of the studies adopts a ''stack of coins'' shape, it indicates homogeneity (Fig. 1).
- **Galbraith's radial plot**: this can be used in both observational and experimental studies. It represents the precision of each study (X-axis) against its standardised effect (Y-axis), along with the fitted regression equation line and confidence bands. The points outside these bands contribute the greatest heterogeneity to the analysis, whereas if all the points were located within the said bands, we would speak of homogeneity. The position of each study on the precision axis indicates the weight of its contribution to the overall study, with those of greatest weight being on the right side of the graph[7] (Fig. 2).
- **L'Abbé plot**: this is a useful tool when working with a dichotomous response in clinical trials. In it, the proportion of events in the control group (X axis) is represented against the proportion of events in the treatment group (Y axis) and their position with respect to the bisector. Each of the points in the plot represents the measure of association of each study, in such a way that, above the diagonal, there are studies favourable to the treatment, while, below it, there are those favourable to the control group. Heterogeneity is interpreted according to the dispersion of the points with respect to the bisector.[8] (Fig. 2).

The presence or absence of heterogeneity will determine the analysis method that should be used use[1–3,5]:

- If there is no heterogeneity or it is very low, the fixed-effects model analysis will be performed. This model assumes that study size and intra-study variability (variance) are the only determinants of the weight of each individual study within the meta-analysis.
- If heterogeneity exists, the random-effects model should be used. This model considers both intra-study and inter-study variability. It has two fundamental limitations: firstly, it tends to increase the variance of the meta-analysis result, reducing the possibility of finding a significant result (decreased power) and secondly it generates wider confidence intervals, thus granting excessive weight to studies with small sample size. Whenever the presence of heterogeneity is detected, its cause should be investigated. To do this, a subgroup analysis or a (more efficient) meta-regression is used, allowing us to explore whether the characteristics of the patients, the methodologies or other aspects of the studies may have generated heterogeneity.[1,3]

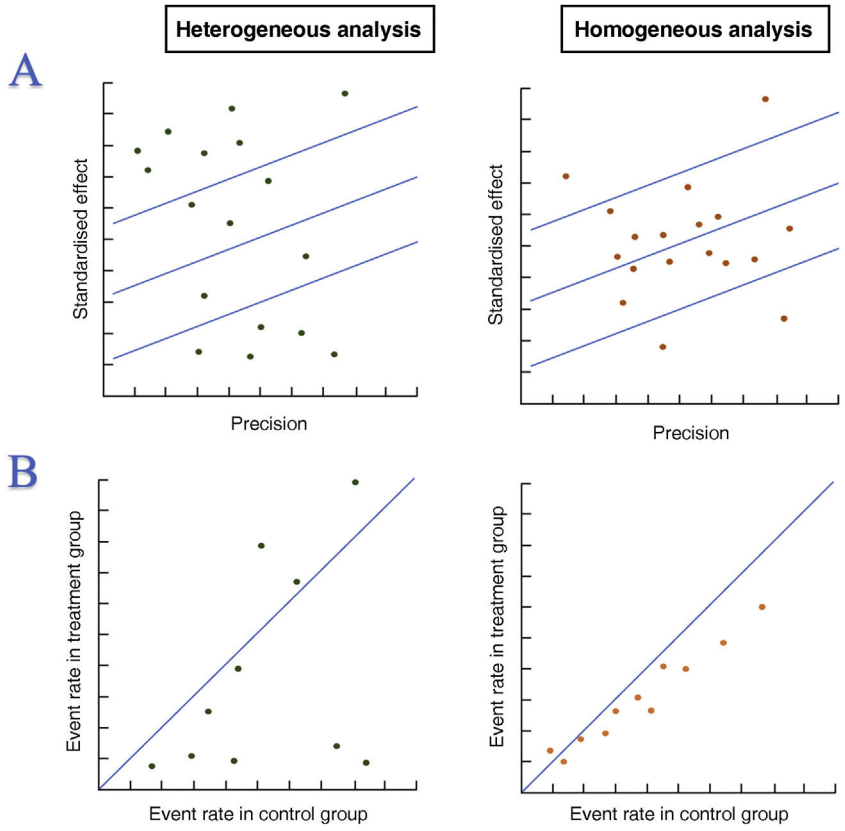

Fig. 2 – Graphical interpretation of heterogeneity using simulated data. A) Galbraith's plot B) L'Abbé plot.

REFERENCES

1. Delgado M. Revisión sistemática de estudios. Metaanálisis, 4.ª ed Barcelona: Signo; 2014.
2. Molina Arias M. Aspectos metodológicos del metaanálisis (1). Rev Pediatr Aten Primaria. 2018;20:297–302.
3. Martín Conejero A, Alonso García M, Quirós González V. Manual CTO. Metodología básica de la investigación clínica. Grupo CTO Editorial; 2019.
4. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med. 2002;21:1539–58. http://dx.doi.org/10.1002/sim.1186.
5. Pértega Díaz S, Pita Fernández S. Revisiones sistemáticas y metanálisis (II). Cad Aten Primaria. 2005;12:166–71.
6. Fletcher J. What is heterogeneity and is it important? BMJ. 2007;334:94. http://dx.doi.org/10.1136/bmj.39057.406644.68.
7. Galbraith RF. Graphical display of estimates having differing standard errors. Technometrics. 1988;30:271–81. http://dx.doi.org/10.1080/00401706.1988.10488400.
8. L'Abbé KA, Detsky AS, O'Rourke K. Meta-analysis in clinical research. Ann Intern Med. 1987;107:224–33. http://dx.doi.org/10.7326/0003-4819-107-2-224.