

## Heterogeneidad entre los estudios incluidos en un meta-análisis

Preparado por Luis M. Molinero (Alce Ingeniería)  
CorreoE: [bioestadistica@alceingenieria.net](mailto:bioestadistica@alceingenieria.net)

 [www.seh-lelha.org/stat1.htm](http://www.seh-lelha.org/stat1.htm)

 [Artículo en formato PDF](#)

Diciembre 2003

### Introducción

En un [artículo anterior](#), en el que se efectuaba una introducción a la metodología estadística de los meta-análisis, ya se comentó la importancia de cuantificar la heterogeneidad entre los estudios que componen el meta-análisis, ya que de su magnitud se deriva no sólo el método utilizado para combinar los resultados individuales, sino también la propia validez de las conclusiones globales.

Cuando hablamos de heterogeneidad podemos distinguir dos aspectos: por un lado el relativo a las diferencias existentes entre los estudios en cuanto a características de los pacientes incluidos, la metodología utilizada, el tiempo de seguimiento, las dosis empleadas, la localización geográfica, etc. y por otro lado el concepto de heterogeneidad estadística, que únicamente cuantifica la variabilidad entre los resultados de los estudios, y que puede ser debida a las diferencias reales de planteamiento y ejecución entre los estudios incluidos, o a otras causas.

La elaboración de guías de práctica clínica se fundamenta cada vez más en los resultados de revisiones sistemáticas, en las que los meta-análisis constituyen una herramienta crucial, y puesto que la validez de la conclusión global de un meta-análisis depende en gran medida de la homogeneidad de los estudios incluidos, resulta de extraordinaria importancia disponer de algún parámetro que cuantifique la heterogeneidad.

Uno de los aspectos de la heterogeneidad, el relativo a las diferencias clínicas o biológicas entre estudios y a las diferencias de procedimientos, es en primer lugar un problema metodológico, ya que habrá que decidir si las diferencias entre los estudios, que siempre existen, permiten o no combinarlos, independientemente de los resultados que en ellos se haya obtenido, y por lo tanto es previo a la ejecución del meta-análisis. Mientras que la heterogeneidad estadística trata de cuantificar la variabilidad del resultado medido en los diferentes estudios con respecto al resultado global promedio, y determinar si dicha variabilidad es superior a la que sería esperable por puro azar.

### Prueba estadística para verificar la existencia de heterogeneidad

La prueba estadística más ampliamente utilizada para verificar la posible existencia de heterogeneidad superior a la esperable por puro azar se denomina [Q de Cochran](#), y se basa en calcular la suma de las desviaciones cuadráticas entre el resultado individual de cada estudio y el resultado global, ponderadas por el mismo peso con el que cada resultado interviene en el cálculo global:

$$Q = \sum w_i (T_i - \bar{T})^2$$

En la hipótesis de homogeneidad  $Q$  se distribuye aproximadamente como una  $\chi^2$  con  $k-1$  grados de libertad, siendo  $k$  el número de estudios.

Sin embargo el empleo de esta prueba no está exento de problemas, ya que si el número de estudios es pequeño su capacidad para detectar heterogeneidad es muy baja (poca potencia de contraste), mientras que, por el contrario, cuando el meta-análisis combina gran número de estudios, el resultado puede ser

estadísticamente significativo incluso cuando la magnitud de la heterogeneidad no sea de relevancia clínica. Esto no es nada nuevo: son los problemas inherentes a la metodología de las pruebas de contraste estadístico.

Puesto que la situación más habitual en los meta-análisis es la de un reducido número de estudios, se recomienda utilizar un nivel de rechazo en esta prueba al menos de 0.1, en lugar del tradicional 0.05.

### Otras alternativas para cuantificar la heterogeneidad en un meta-análisis

Las limitaciones de la prueba  $Q$  anteriormente comentadas, así como el hecho de que su resultado no sirva para comparar diferentes meta-análisis, en los que intervienen diferente número de estudios, ha llevado a buscar otros índices que permitan cuantificar la heterogeneidad de manera que el parámetro calculado sirva para comparar diferentes meta-análisis y así, de alguna manera, ponderar la validez del efecto medio en ellos calculado. En esa línea se ha propuesto recientemente un nuevo índice denominado  $I^2$ , que comienza a aparecer ya en las revisiones sistemáticas y del que se dispone de una muy buena descripción en un reciente [artículo de BMJ](#).

Partiendo de la idea de que la pregunta de interés no es si existe o no heterogeneidad estadísticamente significativa, sino cómo afecta ésta a las conclusiones del meta-análisis, los autores trataron de buscar un índice que permita la comparación entre diferentes meta-análisis y por tanto que no dependa ni de la unidad de medida utilizada para cuantificar el efecto, ni del número de estudios incluidos, y que además sea fácilmente interpretable por personal no muy experto en estadística. Como veremos seguidamente, el índice propuesto cumple esos requisitos.

El parámetro  $I^2$  indica la proporción de la variación entre estudios respecto de la variación total, es decir la proporción de la variación total que es atribuible a la heterogeneidad:

$$I^2 = \frac{\tau^2}{\tau^2 + \sigma^2}$$

donde  $\tau^2$  corresponde a la variabilidad entre estudios y  $\sigma^2$  a la variabilidad interna en los estudios, ésta última debida fundamentalmente a la variabilidad entre pacientes. Como vemos esta expresión recuerda al coeficiente de [correlación intraclase](#).

Quien esté interesado en saber cómo se deriva su cálculo puede consultar el [artículo](#) que se indica en las referencias. Su cálculo es muy sencillo, a partir del valor de la prueba  $Q$ , se obtiene mediante la siguiente expresión:

$$I^2 = \frac{Q - k + 1}{Q}$$

siendo  $k$  el número de estudios. Cuando el valor resulta negativo se tomará como igual a cero. Se suele expresar en porcentaje %, para lo cual el resultado de la expresión anterior se multiplicará por 100.

## ● Cálculo del intervalo de confianza para $I^2$

Para el cálculo del intervalo de confianza de  $I^2$  utilizaremos otro parámetro relacionado y que vamos a denominar  $H^2$ . Viene dado por la expresión:

$$H^2 = \frac{Q}{k-1}$$

$H^2$  constituye también un parámetro que nos permite cuantificar la heterogeneidad, ya que indica el exceso en el valor de  $Q$  respecto del valor esperado si no hubiera heterogeneidad, dicho valor esperado es precisamente  $k-1$ . Luego  $H^2=1$  si no existe heterogeneidad y tanto mayor cuanto mayor sea ésta. Con unas operaciones elementales se puede comprobar que:

$$I^2 = \frac{H^2 - 1}{H^2}$$

Calculamos en primer lugar el intervalo de confianza para  $H^2$  y utilizando la fórmula anterior se deducirán los límites para  $I^2$ .

Dicho intervalo se calcula con la siguiente expresión

$$\exp(\ln H \pm Z_{\alpha} \cdot ES[\ln H])$$

siendo  $Z_{\alpha}$  el valor correspondiente del percentil de la distribución normal = 1.96 para  $\alpha=0.05$ .

$ES[\ln(H)]$  error estándar de  $\ln H$ , viene dado por las siguientes fórmulas:

$$\frac{1}{2} \frac{\ln Q - \ln(k-1)}{\sqrt{2Q - \sqrt{2k-3}}} \quad \text{Si } Q > k$$

$$\sqrt{\frac{1}{2(k-2)} \left( 1 - \frac{1}{3(k-2)^2} \right)} \quad \text{Si } Q \leq k$$

## ● ¿Qué valor $I^2$ indica una heterogeneidad elevada?

En el [artículo de BMJ](#) indicado en las referencias, los autores proponen los valores de 25%, 50%, y 75% como posibles marcas para indicar una heterogeneidad *baja*, *moderada* o *alta* respectivamente, y efectúan un cálculo de  $I^2$  en 509 meta-análisis en Cochrane Database of Systematic Reviews, encontrando que sólo una cuarta parte de los mismos presenta valores superiores al 50 % y casi la mitad de ellos (250) tienen una excelente homogeneidad ( $I^2=0\%$ ).

## ● Epílogo

El índice  $I^2$  recientemente propuesto es un parámetro muy fácil de calcular y también sencillo de interpretar, permitiendo comparar meta-análisis en cuanto a la heterogeneidad estadística que presentan los estudios que incluyen, por lo que seguramente se irá ampliando su difusión y utilización, y por lo tanto conviene conocerlo.

Una vez más hay que resaltar que la ausencia de heterogeneidad estadística no garantiza que los estudios sean homogéneos en cuanto a sus características clínicas, biológicas o metodológicas, y por lo tanto no garantiza la validez de proceder a combinarlos. Es pues obligado investigar y comentar esos aspectos de los estudios incluidos. Las técnicas estadísticas son herramientas al servicio del razonamiento, no significan nada por sí mismas. Así todos sabemos que no es lo mismo correlación que causalidad (relación causa-efecto) y que a veces la correlación no es mas que mera casualidad, o la relación existente debida a una tercera variable, que

al desaparecer hace que también se desvanezca la correlación.

El análisis de la heterogeneidad en un meta-análisis debe ser pues exhaustivo y además del sencillo índice aquí descrito, disponemos de otros mecanismos para su estudio, ya comentados en un [artículo anterior](#).

---

## Referencias

- [Measuring inconsistency in meta-analyses](#)  
Julian P T Higgins, Simon G Thompson, Jonathan J Deeks, and Douglas G Altman  
BMJ, Sep 2003; 327: 557 – 560
- [Quantifying heterogeneity in a meta-analysis.](#)  
Higgins JPT, Thompson SG.  
*Stat Med* 2002; 21: 1539–58.



[Indice de artículos](#)

[Principio de la página](#)