



**VNiVERSiDAD
D SALAMANCA**

Departamento de Estadística

Máster en Análisis Avanzado de Datos Multivariantes y Big Data

Trabajo Fin de Máster

**Revisión sistemática del impacto del
sesgo de publicación en metaanálisis
y la influencia de la heterogeneidad
en los métodos de detección**

Autor : FRANCISCO JOSÉ RODRÍGUEZ ALCÁZAR

Tutor: DR. JAVIER MARTÍN VALLEJO

2017



Dpto. de Estadística

Universidad de Salamanca

DR. JAVIER MARTÍN VALLEJO

Profesor Titular del Departamento de Estadística de la Universidad de Salamanca

CERTIFICA que **D./D.^a Francisco José Rodríguez Alcázar** ha realizado en la Universidad de Salamanca, bajo su dirección, el trabajo que para optar título de Máster en Análisis Avanzado de Datos Multivariantes y Big Data presenta con el título ***Revisión sistemática del impacto del sesgo de publicación en metaanálisis y la influencia de la heterogeneidad en los métodos de detección***, autorizando expresamente su lectura y defensa.

Y para que conste, firma el presente certificado en Salamanca a 1 Septiembre de 2017.

Dr. Javier Martín Vallejo

**Revisión Sistemática Del Impacto Del Sesgo De Publicación En
Metaanálisis Y La Influencia De La Heterogeneidad En Los Métodos De
Detección**



Dpto. de Estadística

Universidad de Salamanca

Trabajo para optar al título de Máster en
Análisis Avanzado de Datos Multivariantes y
Big Data por la Universidad de Salamanca.

Presenta:

Francisco José Rodríguez Alcázar

Salamanca

2017

Agradecimientos

Este trabajo pone un punto y seguido a la realización del Master de Análisis de Datos Multivariantes y Big Data que imparte la Universidad de Salamanca. Mi deseo es haber podido plasmar la mayor parte de las competencias adquiridas durante su realización y que fueron aportadas por los distintos profesionales que han pasado por él.

En primer lugar, agradecer al Dr. Javier Martin Vallejo como tutor de este trabajo por sus conocimientos y tiempo dedicado, que con su apoyo ha sido posible la realización de esta revisión.

Agradecer el esfuerzo y la dedicación al Departamento de Estadística de la Universidad de Salamanca y otros integrantes colaboradores de este Máster con carácter multidisciplinar, al que he podido sobrevivir.

Silvia, Sara, Joaquín y los yayos por la paciencia que habéis tenido en todos los momentos de crisis. Esos días de lluvia, de frío y de calor donde estaba estudiando sin estar con vosotros, las reuniones del cole que me perdí o las tardes de parque en las que estaba ausente. Os quiero.

Mami y papi sin vosotros sería complicado tener la felicidad de ser lo que soy y el orgullo de haber llegado hasta aquí. Sois inolvidables.

María, Lola, Esther, Oscar, Laura, Gloria, Marta y Rosa por vuestro apoyo constante y sin el cual no habría podido llegar hasta este momento tan bonito. Soy enfermero.

Índice

1. Introducción	13
2. Objetivos.....	15
2.1. Objetivos secundarios	15
3. Material y Métodos.....	17
4. Definición de sesgo de publicación	25
5. Tipos de sesgo de publicación.....	27
6. Impacto del sesgo de publicación.....	29
7. Métodos de detección y corrección del sesgo de publicación	39
7.1. Herramientas para detectar el sesgo de publicación.....	39
7.2. Métodos estadísticos exclusivamente cuantitativos (método formal)....	40
7.3. Métodos representación datos de forma gráfica (método informal)	49
7.4. Métodos estadísticos cuantitativos con representación gráfica (método correcto)	54
8. Influencia de la heterogeneidad en el sesgo de publicación.....	61
9. Estudio de simulación para el análisis de la eficiencia de los métodos de detección del sesgo de publicación sin y con heterogeneidad.....	67
10. Conclusiones.....	97
11. Bibliografía	99

Resumen

Introducción: Los métodos de detección y corrección del sesgo de publicación realizados durante el proceso de metaanálisis pueden verse influidos en presencia de heterogeneidad. **Objetivo:** Revisión sistemática de la información existente sobre el problema del sesgo de publicación. **Material y método:** Búsqueda bibliográfica en base de datos on-line y libros, y una simulación en diferentes escenarios de heterogeneidad y sesgo de publicación. **Resultado:** Los métodos evaluados funcionan con anormalidad en condiciones de heterogeneidad con tamaño del efecto grande y en metaanálisis que integran un número de estudios pequeño. **Conclusión:** El grado de heterogeneidad y el número de estudios integrados en los metaanálisis influye en el resultado de los métodos de detección y corrección del sesgo de publicación.

PALABRAS CLAVE: Sesgo de Publicación, Metaanálisis como Asunto, Heterogeneidad, Impacto.

Introduction: Methods of detecting and correcting publication bias during the meta-analysis process can be influenced by the presence of heterogeneity. **Objective:** Systematic review of existing information on the problem of publication bias. **Material and method:** Bibliographic search in on-line database and books, and a simulation in different scenarios of heterogeneity and publication bias. **Result:** The evaluated methods work with the abnormality under conditions of heterogeneity with the size of the large effect and in the meta-analysis that integrate a small number of studies. **Conclusion:** The degree of heterogeneity and the number of studies integrated in the meta-analysis influence the results of methods of detection and correction of publication bias.

KEY WORDS: Publication Bias, Meta-Analysis as Topic , Heterogeneity, Impact.

1. Introducción

Antes de la realización de cualquier trabajo de investigación o ensayo, hay un arduo trabajo para conocer en que punto está el tema que a investigar. Este trabajo consiste en una revisión bibliográfica, desde un soporte físico como son los libros hasta contactar con los distintos profesionales expertos en ese tema, pasando por el soporte on-line cada vez más extendido en donde obtener casi toda la documentación necesaria y disponible.

Después de realizar la búsqueda, los investigadores deciden en que momento la información es “suficiente”, para iniciar la aplicación del proyecto de investigación previamente diseñado. La revisión sistemática y el metaanálisis de los diferentes conjuntos de datos obtenidos en esta investigación son dos caminos para evaluar la información recogida.

Tanto la aplicación de un proyecto de investigación en sus diferentes modalidades de estudio (observacionales y experimentales) como la revisión sistemática y el metaanálisis pueden contener un déficit en la información recogida. Puede que en un estudio experimental se necesite la información solo de cómo se han proyectado los estudios y los resultados obtenidos, sin entrar en detalle en otros aspectos del estudio (Sackett et al., 1996).

Pero si es importante el diseño del estudio, más importante es poder integrar los resultados de esos estudios para conocer si sus resultados en su conjunto tienen efectos en la población en general. Esto es una definición más de metaanálisis, de las muchas que pueden ser encontradas (Glass, 1976).

Pero la información que se pretende integrar puede tener características similares o no (heterogeneidad vs. homogeneidad), ser completa esa información o presentar algún tipo de sesgo (publicación, recuperación,...) o si los resultados finales pueden ser extrapolados a la población en general.

La información relacionada con el metaanálisis y la revisión sistemática puede influir entre ellas. Para ello, es necesario realizar una validación de los resultados obtenidos en la revisión. Dos factores importantes en esta validación son la heterogeneidad y el sesgo de publicación (Cobos-Carbó & Augustovski, 2011; Higgins & Green, 2011).

El sesgo de publicación tiene diferentes causas que lo originan (Thornton & Lee, 2000). Se ha intentado objetivar el sesgo de publicación con formulaciones estadísticas, bastante complejas en algunas ocasiones, y en otras algo más sencillas de forma gráfica.

Incluso se ha analizado el sesgo de publicación, no solo en el déficit de datos no significativos, sino en el exceso de datos significativos (Begg, 1985; Hedges, 1984; Lau et al., 1992; Sugita et al., 1992).

Se constata que la cantidad de estudios significativos publicados es superior a los no significativos, provocando este exceso de publicación (Orwin, 1983; Rosenthal, 1979; Sterling, 1959).

No solo hay métodos para conocer si hay sesgo de publicación sino que podemos encontrar una serie de métodos correctores, en relación a la influencia que puede ejercer la heterogeneidad en los resultados. Distintos autores hablan de heterogeneidad auténtica-real, y afinando aún más, reconocen la heterogeneidad residual. Explican como la heterogeneidad puede afectar a la fiabilidad de los diferentes métodos de detección de sesgo de publicación (Thompson & Sharp, 1999).

Después de todo lo escrito hasta ahora, la gran cantidad de artículos, manuales hablando de sesgo de publicación y de cómo la heterogeneidad puede afectar a los resultados obtenidos al igual de cómo evaluar el impacto de esta heterogeneidad en el sesgo de publicación, es necesaria una revisión de puesta a punto del tratamiento del sesgo de publicación, así como, de sus redefiniciones, etc.

2. Objetivos

El objetivo general de este trabajo es la revisión sistemática de la bibliografía existente sobre el problema del sesgo de publicación.

2.1. Objetivos secundarios

- Revisión de los métodos cuantitativos y gráficos para detectar y corregir el sesgo de publicación y analizar su eficiencia en ciertos contextos.
- Revisar y valorar el impacto de la heterogeneidad en el sesgo de publicación.

3. Material y Métodos

Búsqueda bibliográfica

Para la realización de esta revisión sistemática se ha elegido una serie de artículos con características determinadas, las cuales están incluidas en todos ellos.

Las características a cumplir son:

- Estudios completos o “full text”
- Estudios en habla inglesa y española
- Estudios que investiguen relación heterogeneidad y sesgo de publicación
- Estudios que analicen métodos detección y corrección del sesgo de publicación
- Estudios que analicen el sesgo de publicación en diferentes campos de la ciencia.

La búsqueda bibliográfica también se ha realizado en diferentes libros relacionados con el metaanálisis y el sesgo de publicación, y en tesis doctorales que abordan estos temas (Borenstein et al., 2009; Hunter & Schmidt, 2004; Martín-Vallejo, 1995; Perez, 2009)

La búsqueda bibliográfica online ha sido realizada en Web of Science , PubMed y Google Academic, con una búsqueda determinada por las características previamente definidas:

- **La búsqueda realizada en Web of Science contenía el criterio de búsqueda:** *TOPIC: (Publication Bias) AND TOPIC: (Meta-Analysis) AND TOPIC: (Heterogeneity) AND TOPIC: (Impact)* obteniendo el resultado de 144 estudios.
- **La búsqueda realizada en PubMed contenía el criterio de búsqueda:** *((Publication Bias) AND (Meta-Analysis) AND (Heterogeneity) AND ((Impact))* obteniendo el resultado de 356 estudios.
- **La búsqueda realizada en Google Academic contenía el criterio de búsqueda:** *((Publication Bias) AND (Meta-Analysis) AND (Heterogeneity) AND ((Impact))* obteniendo el resultado de 71300 estudios.

El buscador PubMed tiene la opción de realizar la búsqueda utilizando TESAUROS (MeSH, Medical Subject Headings), de los cuales se seleccionó los siguientes:

- **Publication Bias** (MeSH Unique ID: D017594) y aparece en un Tree numbers (L01.737.813). Con subtítulo “statistics and numerical data”.
- **Meta-Analysis as Topic** (MeSH Unique ID: D015201) y aparece en cuatro Tree numbers (E05.318.780.500, E05.581.500.501, N05.715.360.780.515, N06.850.520.445.500).

Se ha contado con la ayuda del Departamento De Estadística De La Universidad De Salamanca con material bibliográfico y otro tipo de documentación valiosa para esta revisión sistemática.

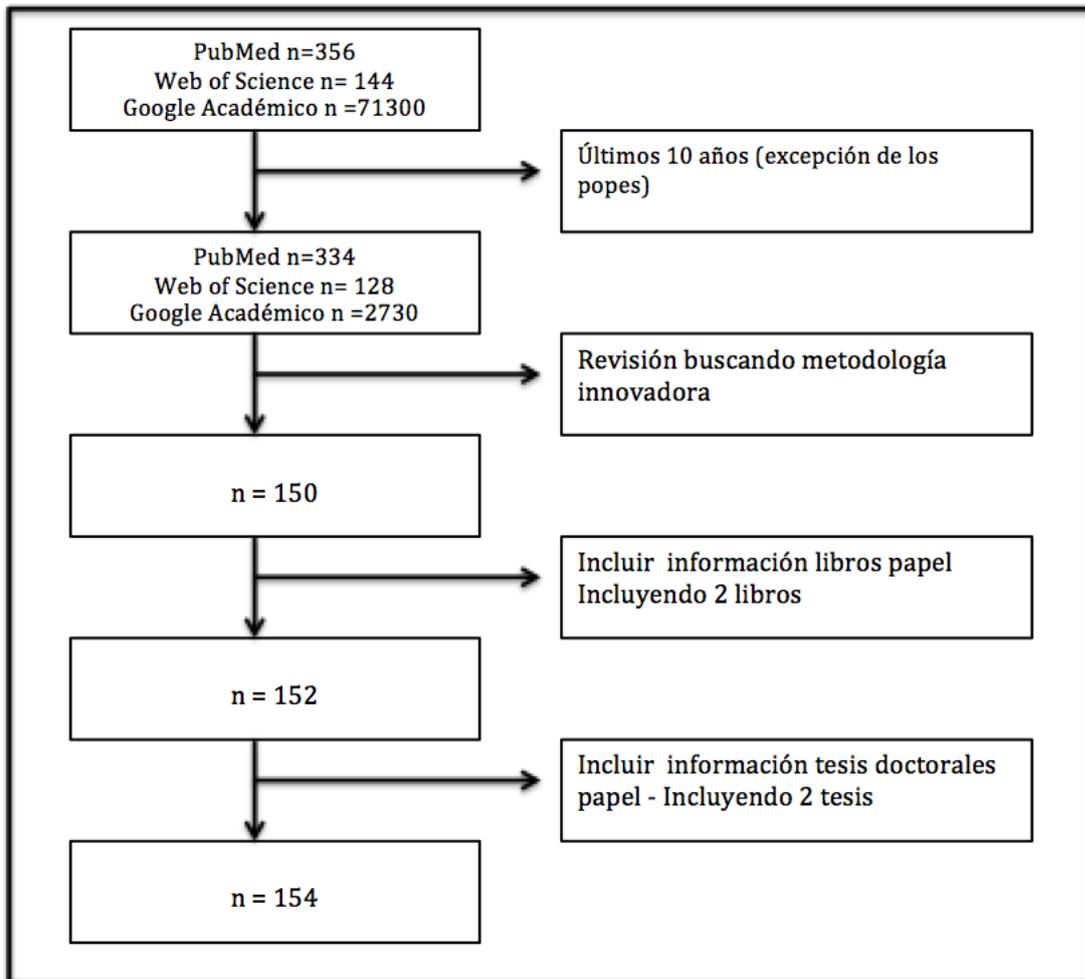


Figura 1. Diagrama de flujo de la revisión sistemática.

Proceso de simulación

Después de la búsqueda bibliográfica se procede a realizar una simulación en diferentes escenarios para evaluar si se comportan de forma diferente los métodos de detección de sesgo de publicación y corrección en presencia o ausencia de heterogeneidad. Se decide plantear dos escenarios globales que son la ausencia de heterogeneidad (homogeneidad) y presencia de heterogeneidad. Cada uno de ellos tendrá una presencia del 0%, 50% y 100% de sesgo de publicación en estudios no significativos ausentes.

a) **Ausencia de heterogeneidad.** Se parte de metaanálisis con un tamaño k de 10 y 100 estudios donde se evalúa el tamaño del efecto (d) en las cantidades 0,25; 0,40 y 0,80. Cada uno de ellos tendrá una presencia del 0%, 50% y 100% de sesgo de publicación en estudios no significativos ausentes.

Se utiliza la herramienta informática Microsoft® Excel® 2011 para MAC (versión 14.4.4) para crear una hoja de cálculo donde en cada una de las columnas se insertaran unas fórmulas para conseguir los datos necesarios para realizar la simulación en sus diferentes escenarios.

En primer lugar se simularon los diferentes tamaños muestrales que presentarían los estudios a integrar. Para ello se parte de una distribución Weibull, que generalmente es representativa del análisis de supervivencia. Se eligió esta distribución porque es una distribución muy flexible a la hora de describir el comportamiento de aparición de un evento cambiando los parámetros de la distribución. La función de distribución de Weibull es:

$$F(x; \alpha, \beta) = 1 - e^{-\left(\frac{x}{\alpha}\right)^\beta}$$

para $x \geq 0$, siendo nula cuando $x < 0$.

Siendo su ecuación para la función de densidad de probabilidad de Weibull la siguiente :

$$f(x; \alpha, \beta) = \frac{\beta}{\alpha^\beta} x^{\beta-1} e^{-\left(\frac{x}{\alpha}\right)^\beta}$$

Se considero una distribución de tamaños muestrales asimétrica a la derecha con valores de n para cada grupo, con frecuencias altas por debajo de 50 individuos y con baja probabilidad de estudios con tamaños muestrales por grupos superiores a 100 individuos.

Por ello se marcaron los parámetros $\alpha = 50$ y $\beta = 2$, donde α es el parámetro de escala y β es el parámetro de forma. Esta distribución es asimétrica, con valor mediano poblacional de tamaños muestrales con estos parámetros de 24 y valor medio de 50.

Los resultados obtenidos para esta distribución serán tratados como datos continuos, procediendo a su redondeo con posterioridad.

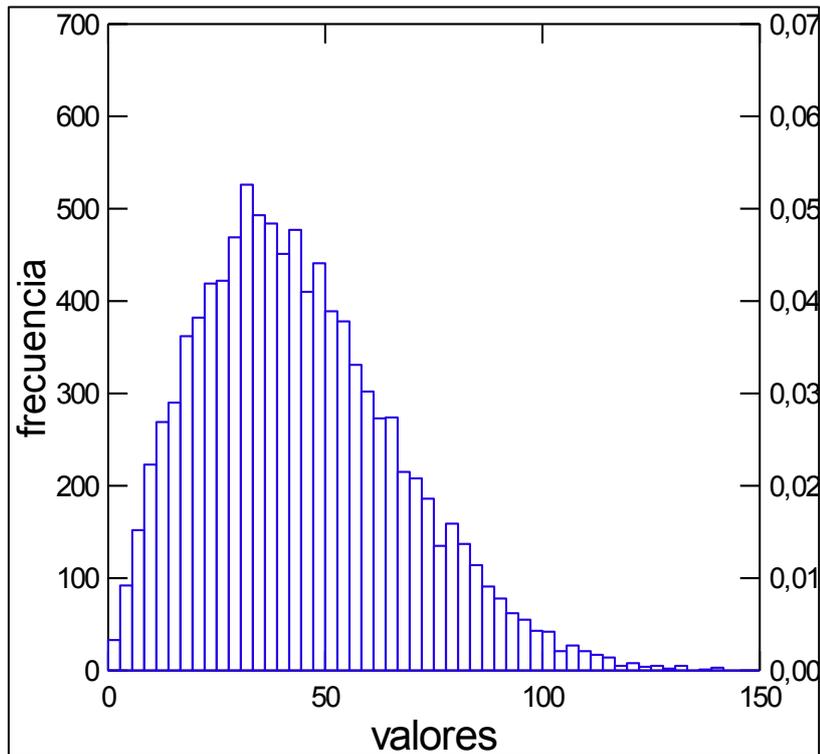


Figura 2. Distribución Weibull.

También se considera que los estudios son balanceados, es decir, los dos grupos presentan el mismo tamaño muestral.

El tamaño del efecto elegido para el estudio de la simulación es la diferencia de medias estandarizada propuesta por Hedges (Hedges, 1981) :

$$d = \frac{\hat{\mu}^T - \hat{\mu}^C}{S}$$

donde μ^T es la media del grupo tratamiento, μ^C es la media del grupo control y el valor S se calcula como la raíz cuadrada de la media ponderada de las cuasivarianzas:

$$S = \sqrt{\frac{(n_T - 1)s_T^2 + (n_C - 1)s_C^2}{n_T + n_C - 2}}$$

Se parte de la distribución Normal donde el grupo tratamiento tiene parámetros 3 y 1. A partir de esa distribución y de los tres escenarios del tamaño del efecto (0,25; 0,4 y 0,8), se calcula la media del grupo control considerando que la desviación típica es igual en ambos grupos.

Posteriormente se procede a simular las medias del grupo tratamientos y sus desviaciones típicas a partir de la distribuciones normales y Chi-cuadrado, respectivamente.

Para cada una de las tres situaciones anteriores según el valor del tamaño del efecto poblacional δ (0,25; 0,40 y 0,80) y los dos diferentes tamaños de los metaanálisis con k igual a 10 y 100, se procede a calcular el valor del estadígrafo (t) y el p-valor unilateral para conocer si el estudio es significativo o no .

Una vez calculado la significación de los estudios se plantean tres escenarios de sesgo de publicación; que no se publiquen el 100% de los estudios no significativos, que no se publiquen un 50% y que no exista sesgo de publicación.

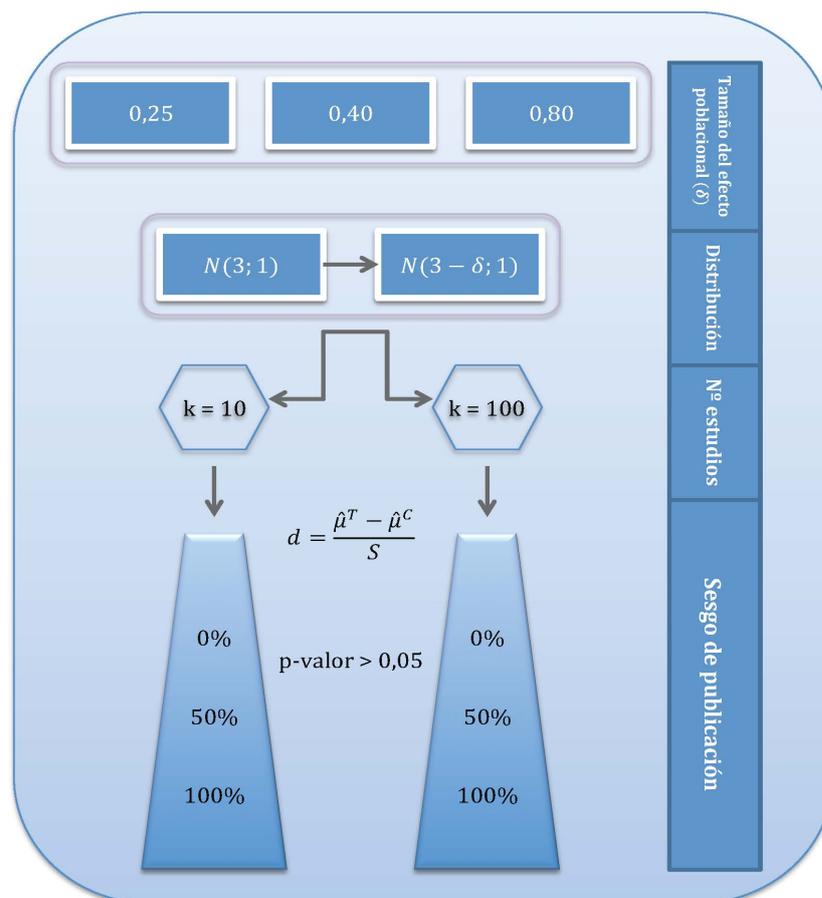


Figura 3. Simulación ausencia de heterogeneidad.

b) **Presencia de heterogeneidad.** Este supuesto se divide a su vez en otros dos supuestos en función de si la heterogeneidad este provocada por un factor experimental o bien exista una variabilidad intrínseca entre los estudios. En este caso el escenario del número de estudios a integrar se considera exclusivamente 100, ya que 10 sería difícil de analizar el efecto de la heterogeneidad en el sesgo de publicación. Los diferentes tamaños del efecto poblacional son δ (0,25; 0,40 y 0,80):

- En el caso que la heterogeneidad esté provocada por un factor, los 100 estudios se dividen en un tercio cada uno, con tres diferentes tamaños del efecto. Un tercio de estudios con el tamaño del efecto poblacional δ , otro tercio con $(\delta - 0,2)$ y otro tercio con $(\delta + 0,2)$.

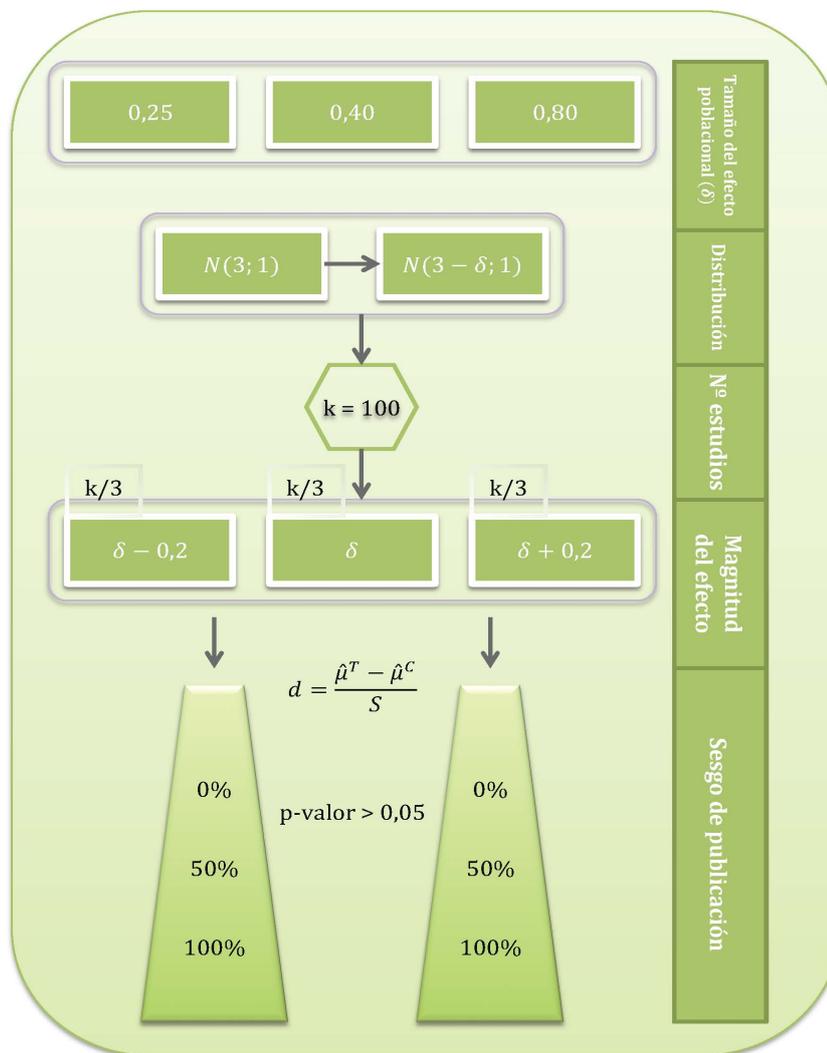


Figura 4. Simulación presencia heterogeneidad provocada por un factor.

- Para el supuesto de una heterogeneidad intrínseca, se parte de una población Normal donde la media corresponde a la media del grupo tratamiento y variabilidad τ^2 . Son tres los escenarios generados según el τ^2 y diferentes tamaños del efecto poblacional de δ (0,25; 0,40 y 0,80). Se genera una muestra (metaanálisis) compuesta por 100 estudios con distribución Normal, donde la media procede de la distribución anterior y la varianza tiene valor 1.

Una vez calculado la significación de los estudios se plantean tres escenarios de sesgo de publicación; que no se publiquen el 100% de los estudios no significativos, que no se publiquen un 50% y que no exista sesgo de publicación.

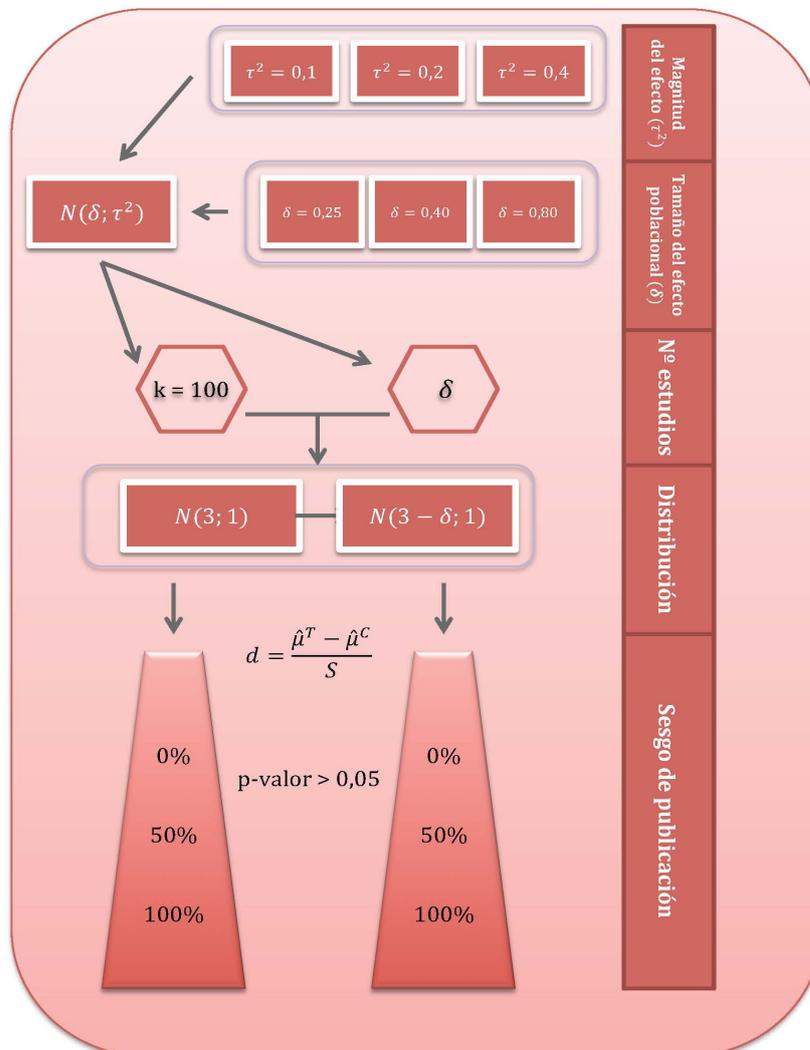


Figura 5. Simulación presencia heterogeneidad intrínseca.

Los métodos estadísticos para detección y corrección de sesgo de publicación evaluados con la aplicación informática MIX v.1.7 y 2.0. son los descritos en la **Tabla 1**.

Métodos estadísticos de salida numérica	Métodos estadísticos de salida gráfica
Begg's rank correlation	Contour-enhanced Funnel Plot
Egger's regression	Funnel Plot
Macaskill's regression	Trim and Fill
Fail-safe N	
Trim and Fill	

Tabla 1. *Métodos detección y corrección del sesgo de publicación. Extraído de : Bax, L., Yu, L., Ikeda, N., Tsuruta, H., & Moons, K. G. M. (2006). Development and validation of MIX: comprehensive free software for meta-analysis of causal research data, 11, 1–11. <https://doi.org/10.1186/1471-2288-6-50>.*

4. Definición de sesgo de publicación

En su origen, el sesgo de publicación aparece como “publication policy” donde el autor describe como un gran número de artículos publicados son los que rechazan la H_0 en sus resultados (Sterling, 1959).

Desde su inicio, los científicos exponen que en el campo de las ciencias del comportamiento no se publican todos los trabajos de investigación que se realizan realmente Bakan (1967), McNemar (1960), Smart (1964) y Sterling (1959). La visión extrema de este problema, el “*the file drawer problem*”, es que las publicaciones tienen el 5% de los estudios que muestran errores de Tipo I, mientras que el cajón está lleno del 95% restante de los estudios que no muestran importancia (Rosenthal, 1979).

Hay autores que mantienen que este asunto no es baladí, a pesar de que muchos compañeros puedan sostener la no importancia o inexistencia del sesgo de publicación (Rosa, 2016).

Podríamos decir que el sesgo de publicación se ha visto minimizado en los últimos años por el acceso a un mayor número de artículos e información a través de la red. Pero todavía existen muchas revistas y portales de la red en donde queda demasiado encorsetado el tipo de artículo con posibilidad de ser publicado, sin poder ser reducido el sesgo de publicación en la actualidad.

Otra definición de sesgo de publicación es la situación que se presenta cuando una serie de trabajos, estudios, ensayos,... no han sido publicados en aquellas fuentes de conocimiento por una serie de motivos entre los que destaca ser trabajos con resultados no significativos que en términos generales no interesan a las revistas donde intentan ser publicados, pudiendo provocar conclusiones erróneas en el área de trabajo.

Algunos autores definieron el sesgo de publicación como el término que ocurre cuando los trabajos que aparecen en la literatura publicada son sistemáticamente no representativos del total de estudios realizados. En pocas palabras, cuando la investigación que es fácilmente disponible difiere en sus resultados, de los resultados de toda la investigación que se ha hecho en un área, los lectores y los revisores de esa investigación están en peligro de sacar una conclusión equivocada (H. Rothstein, Sutton, & Borenstein, 2005).

La red Cochrane que se define como una red global independiente de investigadores, profesionales, pacientes, cuidadores y personas interesadas en la salud dice

que el sesgo de publicación es la publicación o no publicación de los hallazgos de la investigación, según la naturaleza y la dirección de los resultados (Higgins & Green, 2011).

Como puede comprobarse hay diferentes definiciones a lo largo del tiempo de que es el sesgo de publicación, encontrando que una de las más correctas y completas es la realizada por Rothstein, Sutton, & Borenstein (2005).

5. Tipos de sesgo de publicación

En la literatura, desde los inicios de las actividades investigadoras, se han descrito diferentes tipos de sesgo de publicación:

a) *Por la línea editorial de rechazar estudios no significativos*

Desde el inicio del estudio del sesgo de publicación, es uno de los analizados y que es debido a la selección de resultados significativos por parte de las editoriales y revistas (Rosenthal, 1979; Sterling, 1959).

Existe una comparativa de artículos enviados a diferentes editoriales, con el resultado de cuantos son rechazados y otros tantos aceptados, demostrando de una forma evidente lo que se apreciaba de forma subjetiva (Coursol & Wagner, 1986).

b) *Por el tipo de diseño del ensayo e informe de resultados*

Según cómo se ha diseñado el ensayo, sus medidas para evaluar los resultados y de cómo se comparan esos datos para plasmar en un informe final, puede aparecer este tipo de sesgo. Finalmente, el artículo aparecerá publicado o no (Higgins & Green, 2011; Pocock, Hughes, & Lee, 1987).

c) *El tiempo hasta la publicación.*

Puede existir retraso desde que se concluye un estudio hasta que finalmente es publicado. Ocurre en estudios prospectivos pero no en todos los casos, o al igual que puede sufrir un retraso los estudios con un resultado negativo (Misakian & Bero, 1998; H. Rothstein et al., 2005).

d) *Sponsor*

Otra realidad es que la realización de estudios con sponsor, patrocinio o apoyos de instituciones y empresas farmacéuticas orientados a la búsqueda de una positiva significancia en los estudios, hace que estos estudios sean publicados en mayor medida, descompensando y agravando la presencia de sesgo de publicación.

Puede ocurrir que una empresa que patrocina un estudio sobre uno nuevo producto comparando ventajas o desventajas frente a productos de la competencia, tenga los derechos de impedir la publicación de los resultados si son negativos para su nuevo producto (Higgins & Green, 2011; Rennie, 1997).

e) Porque el autor decida no compartir el resultado

En bastantes ocasiones, el autor enjuicia su artículo en virtud de los puntos anteriormente expuestos sobre el sesgo de publicación, porque en su evidencia diaria encuentra que los artículos con un resultado no significativo no se publican. En otros casos, se encuentra con que el tipo de ensayo no pudiera cumplir los estándares de calidad óptimos y finalmente no se envía a publicar.

Hay autores que debido a las exigencias de patrocinio, falta de sponsor o normas de publicación deciden no publicar su artículo.

f) Idioma de la publicación.

La preferencia por publicar los estudios en lengua inglesa y por lo tanto, tendencia a enviar los artículos en dicha lengua, provoca que aparezcan revisiones y metaanálisis con poca significación. El resultado es totalmente diferente cuando se incluyen los estudios relacionados con el tema a analizar en diferentes lenguas que no sea la inglesa.

Algunos autores referencian la presencia de estudios de significancia negativa en editoriales y publicaciones que son en lengua distinta a la inglesa. Por lo tanto, la presencia de este sesgo de publicación puede bascular en las técnicas de análisis del sesgo de publicación de diferentes formas (Egger et al., 1997).

6. Impacto del sesgo de publicación

Hay varios intentos en la bibliografía de analizar el impacto del sesgo de publicación en los diferentes campos de investigación como el clínico, el psicológico, el económico, etc. El primer intento de analizar el impacto del sesgo de publicación se debe a Coursol & Wagner (1986) donde se compara el número de estudios significativos y no significativos publicados. Todo parte de los resultados del metaanálisis de Smith y Glass (1977) donde diferentes autores dudan o plantean la existencia de sesgo de publicación por el conocido efecto de publicar los estudios significativamente positivos frente a los que no (Coursol & Wagner, 1986; Smith & Glass, 1977).

Partiendo de esta idea, decide hacer una encuesta entre profesionales (1000 y respondieron 609) para conocer qué estudios habían realizado, si había sido positivo su resultado en al menos uno y si finalmente habían sido aceptados y/o publicados.

La valoración del impacto sobre el sesgo de publicación se hace en proporciones, de una forma descriptiva, que deja en evidencia el hecho de que se publicaba más los estudios positivamente significativos.

	Publicados	No publicados	Total
Estudios significativos	85(65,8%)	44(34,1%)	129
Estudios neutros/ no significativos	14(21,5%)	51(78,4%)	65
Total estudios	99	95	194

Tabla 2. Estudios publicados vs. no publicados. Extraído de: Coursol, A., & Wagner, E. E. (1986). Effect of positive findings on submission and acceptance rates: A note on meta-analysis bias. *Professional Psychology: Research and Practice*, 17(2), 136–137. <https://doi.org/10.1037/0735-7028.17.2.136>.

El 21,5% de los estudios publicados tenían resultados no significativos. Del total de 194 estudios aceptados por parte de la editorial, solo el 7%(14) de los estudios no significativos son publicados frente a un 43,8%(85) en el otro extremo.

Se estudió el sesgo de publicación en las ciencias sociales mediante el análisis de una población conocida de los estudios realizados - 221 en total - en el que hay un control importante de lo que se publica (Franco et al., 2014).

Los resultados proporcionados sobre evidencia directa de sesgo de publicación en la etapa de producción de la investigación sobre si los autores deciden no escribir y presentar resultados nulos, fueron bastantes concluyentes.

	Resultado estadístico		
	Nulo	Mixto	Fuerte
No escrito	64,6%	12,2%	4,4%
Escrito no publicado	14,6	39,0	34,1
Publicado(de no calidad)	10,4	37,8	38,5
Publicado(de calidad)	10,4	11,0	23,1
Total	100,0	100,0	100,0

Tabla 3. Circuito estudio según grado significancia de su resultado. Extraído de : Franco, A., Malhotra, N., & Simonovits, G. (2014). *Publication bias in the social sciences: Unlocking the file drawer. Science, 345(6203), 1502–1505.* <https://doi.org/10.1126/science.1255484>.

Podemos comprobar como nuevamente los estudios poco significativos no aparecen en las publicaciones. Al tener acceso a todos los proyectos, estos investigadores dan una clave importante. No solo se publica lo significativo sino que a priori el autor del estudio decide no escribir el artículo al no obtener un resultado positivo.

Quince autores decidieron no redactar el artículo debido a que obtuvieron resultados nulos. Notifican que abandonaron el proyecto porque creían que los resultados nulos no tienen potencial de publicación, e incluso se les preguntó si ellos encontraron los resultados interesantes a nivel personal. Nueve de estos autores rechazaron los resultados nulos, quitando prioridad de redactar el estudio en cuestión y concentrarse en otros proyectos.

Tal vez lo más interesante es que dos de esto quince autores, cuyos estudios "no funcionaron", finalmente publicaron artículos que apoyaban sus hipótesis iniciales usando hallazgos obtenidos en muestras pequeñas y de su interés particular.

El sesgo de publicación también ha sido descrito en el área de la psicológica. La tendencia en la ciencia de la psicológica a evitar la publicación de resultados nulos produce una situación que limita la asunción de "replicabilidad de la ciencia", ya que no puede ser significativa sin el potencial reconocimiento de repeticiones fallidas (Ferguson & Heene, 2012).

Estos autores afirman que a menudo se construyen argumentos para bloquear la publicación y la interpretación de resultados nulos. Así, estos resultados nulos desaparecen por prácticas éticamente dudosas por parte de los investigadores.

Otros autores investigan el sesgo de publicación en dos grandes áreas del comportamiento político: el voto económico y los efectos de la publicidad negativa. En ciencias políticas encontraron presencia de sesgo de publicación en las dos literaturas de comportamiento político especificada, siendo más prevalente en las revistas dominantes.

El nivel de sesgo de publicación puede estar subestimando ya que en las últimas etapas de un estudio puede existir un incentivo para descubrir efectos estadísticamente no significativos a través de, por ejemplo, subgrupo y modelo de selección de especificaciones concretas (Gerber et al., 2010).

En estudios patrocinados, el sesgo de publicación puede tener un impacto importante sobre todo en el campo de la medicina y la farmacología que mueven ingentes cantidades de dinero desde investigación a marketing de los productos.

El campo de las estatinas, medicamentos para controlar los niveles de colesterol, es un campo de estudio que no solo se centra en alimentos buenos o malos, sino también en el control farmacológico.

Hay un estudio donde concluyen que los hallazgos obtenidos de los estudios de estatinas patrocinados por la industria parecen similares en magnitud a los de fuentes no relacionadas con la industria. Existen diferencias reales en la eficacia de las estatinas individuales en varias dosis que explican las discrepancias observadas previamente entre estudios patrocinados por la industria y no relacionados con la industria (Nacie et al., 2014).

	Origen sponsor	
	Industria	No industria
Control placebo	43,8%	43,2%
Multi-armed	22,6%	16,2%
Comparación única dosis-estatina	21,2%	83,7%
TOTAL (n° de estudios)	146	37

Tabla 4. Origen financiación relacionado tipo ensayo en estatinas. Extraído de : Naci, H., Dias, S., & Ades, A. E. (2014). Industry sponsorship bias in research findings: a network meta-analysis of LDL cholesterol reduction in randomised trials of statins. *Bmj*, 349(oct03 1), g5741–g5741. <https://doi.org/10.1136/bmj.g5741>.

Los porcentajes sobre quien patrocina el tipo de ensayo son similares en ambos casos. Dentro del estudio, se comparan un total de 31 estudios con características similares entre estudios patrocinados por la industria y otros que no.

Los autores concluyen que a pesar de conocer la influencia ejercida por el patrocinio en el resultado de los estudios, las herramientas para evaluar el sesgo de publicación que están a su alcance no son eficaces para realmente encontrarlo en sus estudios. (Bekelman & Gross, 2015; Lexchin et al., 2003).

El sesgo de publicación esta presente, eso es un hecho claro. Ni todos lo artículos han sido publicados ni todos han sido escritos por sus autores. Otro sesgo de publicación

relacionado directamente con la línea editorial de las revistas las cuales deciden que artículos publican y cuales no, fue analizado por un grupo de investigadores (Van Lent, Overbeke, & Out, 2014).

Parece que este sesgo de publicación no aparece en las revistas seleccionadas para su análisis, sin existir una tendencia clara a publicar preferiblemente manuscritos de Estudios de Control Aleatorizado de fármacos con resultados positivos. Interpretan que el sesgo está antes de la presentación del manuscrito, planteándose el autor el dilema de si su artículo será publicado en virtud del resultado.

Por el contrario, este estudio encuentra relación altamente significativa entre estudios patrocinados que fueron publicados frente a los que no. Sugieren la relación de resultados positivamente significativos en los estudios con patrocinio.

Existe una revisión en The Cochrane Collaboration® donde se estudia si la forma en la cual los médicos practican la medicina y los tratamientos ofrecidos a los pacientes afecta a los resultados obtenidos en estudios clínicos sobre fármacos. La investigación clínica es cada vez más patrocinada por empresas, ya sea porque las empresas realizan los estudios directamente, o los financian total o parcialmente. En investigaciones anteriores han encontrado que los estudios patrocinados por la industria farmacéutica tienden a favorecer los medicamentos de los patrocinadores, más que los estudios con otras fuentes de patrocinio (Bero, 2013).

Los estudios de fármacos y dispositivos patrocinados por la industria con más frecuencia tuvieron resultados de eficacia favorables en comparación con los estudios sobre fármacos y dispositivos no patrocinados. También son a menudo favorables a los productos del patrocinador frente a los estudios de fármacos y dispositivos no patrocinados sin poder ser explicados con las herramientas de evaluación estándar de riesgo de sesgo.

No encontraron una diferencia entre estudios patrocinados por la industria y no relacionados con la industria con respecto a factores estándar que pueden aumentar el riesgo de sesgo, excepto para el cegado. Tampoco encuentran una diferencia entre los estudios de drogas y dispositivos sobre la asociación entre el patrocinio y las conclusiones.

En la industria alimentaria y en el campo médico, los probióticos se utilizan de manera muy diferente. En ambos casos los probióticos se utilizan, como su nombre indica, para efectos beneficiosos. Irónicamente, la mayoría de los estudios han usado sin rigor el mismo probiótico de la industria alimentaria de animales de granja para estudiar también su papel en los seres humanos. De hecho, el mismo capítulo de un libro, que evalúa el uso de

probióticos en animales y hombres, mencionó que la misma especie, se utiliza en los lechones para aumentar el peso, y en los niños para reducir los episodios de infecciones intestinales

Que hay sesgos de publicación en documentos relacionados con probióticos, es otro hallazgo relacionado con el patrocinio. Muy comúnmente los conflictos de interés no se informan, incluso cuando los autores son patrocinados directamente por la industria alimentaria. Como se sugiere en la serie más grande disponible hasta el momento analizando más de 11,000 fuentes (<http://www.ahrq.gov/clinic/tp/probiotiptp.htm>), el efecto a largo plazo de los probióticos en los seres humanos puede ser analizado mediante una investigación independiente. Así se evitaría lo sucedido con el uso de probióticos y pacientes con pancreatitis (Million & Raoult, 2012).

El sesgo inducido por los intereses comerciales puede ser contrarrestado de dos maneras. El primero es erigir un cortafuegos entre el dinero y las personas que hacen la investigación y el análisis de datos. El otro enfoque es desarrollar una fuente de financiación completamente independiente que también lo sea de la industria farmacéutica (Lexchin, 2012).

Una parte de los estudios sobre como influye el sesgo de publicación en la interpretación de los resultados de un metaanálisis insisten en la influencia de la industria farmacéutica en los estudios realizados y sus resultados.

Evaluar si las restricciones en el idioma de publicación afectan las estimaciones de la efectividad de una intervención es otro tipo de sesgo de publicación investigado y que tiene repercusiones en la interpretación de los metaanálisis.

Los resultados de este estudio sugieren que limitar el lenguaje de los informes de los estudios en inglés en revisiones sistemáticas de las intervenciones de la Medicina Convencional no cambia las estimaciones de la efectividad de una intervención (Pham et al., 2005).

Por el contrario, en el estudio de Medicina Alternativa y Complementaria(CAM) encuentra la presencia de sesgo de publicación relacionado con la lengua de publicación. Las revisiones sistemáticas que evaluaban la efectividad de las intervenciones de CAM, produjeron diferentes estimaciones del efecto de la intervención cuando la síntesis de los datos se limita a estudios únicamente en inglés, en comparación con las publicaciones en otras lenguas no inglesas(LOE). La estimación del efecto protector fue 63% menor (en promedio).

Esto indica que los resultados de los estudios en otra lengua distinta al inglés son importantes y deben ser considerados. La exclusión de los estudios clínicos publicados en otros idiomas distintos del inglés puede introducir sesgos y reducir la precisión de las estimaciones combinadas de los efectos del tratamiento (Juni et al., 2002).

Extrajeron las proporciones de las estimaciones de los efectos del tratamiento de los ensayos de idiomas no ingleses comparados con los estudios en inglés para 50 metaanálisis. Las estimaciones del efecto del tratamiento fueron, en promedio, un 16% más beneficiosas en estudios que no son en inglés. Sin embargo, hubo heterogeneidad considerable entre los metaanálisis ($p = 0,003$).

En resultados de análisis estratificados, el efecto del lenguaje parecía ser más pronunciado en la medicina complementaria y menos pronunciado en los estudios con intervenciones de control activo, pero ninguna estadísticamente significativa.

El impacto de los estudios de idiomas no ingleses disminuyó de 8,34 a 7,68 después de la exclusión de los estudios de idioma no inglés (precisión promedio de las estimaciones del efecto del tratamiento). El número de estudios publicados en idiomas distintos al inglés varió de 1 hasta 14 estudios, y del 4,3% al 72,7% de todos los estudios incluidos. Los estudios en lengua no inglesa contribuyeron en un 17,5% al peso en metaanálisis individuales.

Sin embargo, los estudios publicados en otros idiomas pueden ser más difíciles de localizar y pueden requerir la traducción, lo que aumentará los costos y retrasará la conclusión de una revisión.

Los metaanálisis en lengua no inglesa pueden excluir estudios relevantes de especialidades médicas que suelen publicarse en lengua inglesa, como ocurre con la especialidad de Cardiología u Obstetricia y Ginecología.

La importancia de los estudios publicados en revistas no inglesas también es conocida en la medicina especializada. Dentro de una determinada especialidad médica puede tener un gran impacto el sesgo de publicación. En la especialidad de neurología, el 80% de los estudios incluidos en los metaanálisis publicados eran en lengua inglesa. Pues bien, una enfermedad neurológica como es el accidente vascular, es el tema principal de publicación en lengua china. De un total de 166 artículos relacionados con la especialidad neurológica, 116 eran sobre accidente vascular. El impacto que tiene no hallar esta información para un metaanálisis en una enfermedad a estudio en concreto es elevado (He L, 1997).

Otro estudio nos lleva al otro extremo en relación al uso de artículos en determinados lenguajes. No proporciona evidencia en sus resultados de que los metaanálisis restringidos en el lenguaje conduzcan a estimaciones sesgadas de la eficacia de la intervención (Moher et al., 2000).

En su resultado principal obtuvo una diferencia de 2% (ROR = 0,98; CIs = 0,81-1,17), en promedio, entre las estimaciones de tratamiento con versus sin restricciones explícitas en el idioma de publicación de los estudios incluidos.

Estos mismos autores en estudios previos ya habían confirmado que no hubo diferencias estadísticamente significativas en los informes de estudios publicados en idiomas distintos del inglés y los publicados en inglés en cuanto a asignación al azar, doble ciego, abandonos y retiros, o puntuación total general (Moher et al., 1996).

La exclusión de la literatura gris¹ de los metaanálisis también puede conducir a estimaciones exageradas de un 12% de la eficacia de la intervención. En general, los metaanalistas deben intentar identificar, recuperar e incluir todos los informes, tanto gris como publicados, que cumplan con los criterios de inclusión predefinidos. En el análisis de sensibilidad encontraron que si eliminan los resúmenes(abstract) de las muestras, la sobreestimación aumenta a un 38% (McAuley et al., 2000).

Esto nos da una idea de lo importante que es tener acceso a todo lo publicado, indistintamente en que lengua sea presentado para concluir de una forma acertada y, en campos de mayor sensibilidad, de forma segura el resultado final de nuestro metaanálisis.

También se hace una reflexión de por qué hay que incluir la literatura gris, buscarla e incluirla en revisiones sistemáticas. Sin embargo, el 30% de los editores encuestados no publicarían un metaanálisis que incluyera material inédito, aunque recibiera una revisión favorable. Son varios los motivos que podrían llevar a esta respuesta desde evaluación de estudios no realizado por pares, a cuestiones de si la literatura gris es de baja calidad,... (McAuley et al., 2000)

La búsqueda bibliográfica está relacionada con el sesgo de publicación, así podemos ver en las conclusiones extraídas, donde sugieren que el uso de una única base de datos

¹ La literatura gris (también denominada como literatura no convencional, literatura semi-publicada o literatura invisible) es cualquier tipo de documento que no se difunde por los canales ordinarios de publicación comercial, y que por tanto plantea problemas de acceso (Auger, 1989; García Santiago, 1998; Guimaraes Almeida, 1998) .

electrónica para identificar estudios clínicos controlados en inglés es inadecuado porque los estudios de calidad metodológica similares a los obtenidos serían pasados por alto. Incluso si EMBASE y MEDLINE se utilizan en combinación, la recuperación de estudios es incompleta, y la búsqueda manual en revistas seleccionadas sería necesario para realizar una búsqueda exhaustiva (Suarez-Almazor et al., 2000).

Como nota de color diferente sobre sesgo de publicación, ciertos países tienden a publicar más resultados significativamente positivos que otros. En un estudio realizado encontraron esta tendencia en el área de la acupuntura, donde se publican más los resultados positivos (Vickers et al., 1998).

Los 51 estudios realizados en China y Asia oriental y listados en Medline en el período estudiado favorecieron la acupuntura, lo que sugiere según estos investigadores que los estudios clínicos de acupuntura que se originan en ciertos países tienen una proporción inusualmente alta de resultados positivos.

Por otro lado, se puede pensar que los criterios de los revisores son diferentes en esos países.

Una combinación de sesgo de publicación por lengua de publicación y resultados significativos fue el que concluyó que los estudios clínicos controlados aleatorios tienen más probabilidades de ser publicados en una revista en inglés si contienen resultados significativos, mientras que otros estudios se publicaron en revistas nacionales en lengua alemana (Egger et al., 1997).

En otros estudios se revela que en al menos en uno de los 36 metaanálisis consecutivos publicados en revistas revisadas por pares, la exclusión de documentos por razones lingüísticas produjeron resultados diferentes de los que habrían sido obtenidos si no se hubieran cumplido los criterios usados. Aunque los cambios de 0,70 a 0,67 en el estimación puntual y de 1,09 a 0,95 en el límite superior del IC del 95% del Odds ratio pueden considerarse triviales, es importante darse cuenta de que los resultados de las publicaciones son a menudo evaluados y aplicados a la práctica clínica (Grégoire, Derderian, & Le Lorier, 1995).

Aquí proponen el nombre de "sesgo de Torre de Babel" cuya razón sería que:

- Los autores que han completado un ensayo clínico que da resultados negativos podrían estar menos seguros de tenerlo publicado en una gran

revista internacional de difusión escrita en Inglés y luego enviarlo a una revista local.

- Si estos investigadores trabajan en un país que no habla inglés, el artículo se publicará en su propio idioma en una revista local.
- Por lo tanto, es más probable que los resultados positivos de los autores de países que no hablan inglés se publiquen en inglés, y los resultados negativos.

Es evidente que los criterios de exclusión lingüística, además de ser incompatibles con el principio de "visión sistemática de la totalidad de las pruebas de todos los estudios aleatorios no confiados pertinentes", pueden ser una fuente de error.

Si bien la inclusión de datos no publicados en reseñas científicas sigue siendo controvertido, la mayoría de los investigadores directamente implicados en el metaanálisis creen que los datos no publicados no deben ser sistemáticamente excluidos (Cook et al., 1993).

	N° (%)
Total de metaanálisis evaluados	150(100)
Metaanálisis que no buscaron datos no publicados	50(33.3)
Metaanálisis que buscaban datos no publicados	80 53.3)
➤ Metaanálisis que recuperaron datos no publicados	➤ 50(33.3)
○ Metaanálisis que incluyeron datos no publicados en el análisis primario	○ 46(30.7)
No se pudo determinar si se realizaron búsquedas de datos no publicados	20(13.3)

Tabla 5. Utilización de Datos no Publicados en Metaanálisis evaluados. Extraído de : Cook, D. J., Guyatt, G. H., Ryan, G., Clifton, J., Buckingham, L., Willan, A., ... Oxman, A. D. (1993). *Should in Unpublished. Jama*, 269(21), 2749.

La síntesis más válida de la información disponible se obtendrá cuando los metaanalistas sometan el material publicado y no publicado a la misma evaluación metodológica rigurosa presentando los resultados con y sin fuentes inéditas de datos (Dickersin, 1990).

Los revisores pueden utilizar rasgos de las publicaciones que pueden indicar la presencia o no de sesgo de publicación (Borenstein et al., 2009):

- A. Es probable que los estudios de gran tamaño se publiquen sin importar la significación estadística, ya que implican grandes compromisos de tiempo y recursos.

- B. Los estudios de tamaño moderado corren el riesgo de perderse, pero con un tamaño de muestra moderado, incluso los efectos modestos serán significativos, por lo que solo se pierden aquí algunos estudios.
- C. Los estudios pequeños están en mayor riesgo de ser perdidos. Debido al pequeño tamaño de la muestra, es probable que los efectos más grandes sean significativos, con los efectos pequeños y moderados probablemente no publicados.

Las respuestas a esas preguntas son contestadas con los diferentes métodos que veremos a continuación, pero la última es quizás la más difícil de estudiar por venir derivadas de multitud de posibles escenarios. El tercer enfoque intenta estimar el impacto que tuvo el sesgo y estimar cuál sería el tamaño del efecto en ausencia de sesgo. La esperanza es clasificar cada metaanálisis en uno de tres grandes grupos, según la siguiente clasificación:

- I. El impacto del sesgo es probablemente trivial. Si se incluyeran todos los estudios pertinentes, el tamaño del efecto probablemente permanecería prácticamente inalterado.
- II. El impacto del sesgo es probablemente modesto. Si se incluyeran todos los estudios pertinentes, el tamaño del efecto podría cambiar, pero la conclusión clave (que el efecto es o no es de importancia sustantiva) probablemente permanecería sin cambios.
- III. El impacto del sesgo puede ser sustancial. Si se incluyeran todos los estudios pertinentes, el hallazgo clave (que el tamaño del efecto es, o no, de importancia sustantiva) podría cambiar.

Jackson investigó el impacto del sesgo de publicación en las estimaciones de la varianza entre estudios (τ^2 -estadística) en el metaanálisis (Jackson, 2006). El sesgo de publicación no solo puede afectar al tamaño del efecto sino también a la estimación de la varianza entre estudios.

7. Métodos de detección y corrección del sesgo de publicación

7.1. Herramientas para detectar el sesgo de publicación

Las herramientas, o métodos estadísticos, que podemos utilizar para la detección de sesgo de publicación son muchos y no exentos de complicación según que casos, debido al enfoque que cada autor/creador haya dado al método.

Como se ha indicado en la exposición inicial, son multitud las variables que pueden intervenir en la aparición del sesgo de publicación que van desde que un artículo no se publique por falta de significancia estadística de su resultado hasta que una publicación no esté en el idioma habitual de publicación como es el inglés.

En la bibliografía se han encontrado diferentes formas de enunciar y/o clasificar los métodos de análisis del sesgo de publicación. Hay autores que hacen un recorrido de los principales métodos por su relevancia diferenciando entre métodos informales(gráficos) y formales(cuantitativos) (Martín-Vallejo, 1995).

Otros autores hacen un recorrido de los métodos que aplican mediante estudios comparativos (Niemeyer, Musch, & Pietrowsky, 2014).

Finalmente, hay otros autores (Thornton & Lee, 2000) que hacen un recorrido por un gran número de métodos, aunque algunos sean menos utilizados. Estos últimos realizan una clasificación similar a Martín-Vallejo (1995), en el sentido que clasifica en métodos “*Funnel Graph*” y “*Statistical Methods*”, añadiendo por delante de ellos una categoría/método única llamada “*Proportion of significant studies*” donde se evalúa la proporción de estudios significativos dentro del metaanálisis.

La red Cochrane hace otra clasificación más laboriosa con múltiples subniveles donde se presentan los dos tipos de “*Funnel Graph*” como son, el Gráfico de embudo (Funnel Plot) y el Gráfico de embudo de contorno mejorado (Contour-enhanced Funnel Plot). Posteriormente exponen dos sub-clasificaciones principalmente relacionadas sobre que métodos explican y miden la asimetría que presenten los gráficos anteriores y el grado de sensibilidad de estos métodos (Higgins & Green, 2011).

Finalmente, hablan del método “Recortar y rellenar” (Trim And Fill), Fail-safe N y otros métodos relacionados con el número de estudios significativos que no pueden crear entre ellos una subcategoría común, sino subcategorías independientes.

Las clasificaciones utilizadas están elaboradas y con criterio, pero se ha decidido fusionar las diferentes metodologías de clasificación para intentar obtener una que recoja las virtudes de todas intentando clarificar aún más el origen y el por qué de la utilización de cada uno de los métodos de detección del sesgo de publicación.

A continuación, y debido al número tan grande de métodos en la detección de sesgo de publicación, se realiza una pequeña descripción exponiendo en que consiste cada uno de estos métodos y las diferentes dificultades o problemas que pueden presentar según diferentes investigaciones y autores.

7.2. Métodos estadísticos exclusivamente cuantitativos (método formal)

En la siguiente tabla (**Tabla 6**) aparecen los métodos que se han considerado en este primer grupo, exponiendo sus ventajas y limitaciones.

Método	Ventaja	Limitaciones
Proporción de estudios significativos (Sterling, 1959)	Fácil	No cálculo teórico No tipo de sesgo
Fail-safe N (Rosenthal, 1979)(Orwin, 1983)	Fácil Indica número de estudios ausentes	Ausencia modelo estadístico Gran dependencia estimación del efecto de la intervención
Método de Hackshaw (Hackshaw, Law, & Wald, 1997)	Fácil Indica número de estudios ausentes	Similares a Fail-safe N Asociarse a otros métodos ausentes
Métodos truncados (Hedges, 1984)		Asume solo publicación estudios no significativos
Modelo de función ponderada (Cooper & Hedges, 1994a; Vevea & Woods, 2005)	Mejorar en reducir número estudios	Complejidad Algunos necesitan muchos estudios
Muestreo invariante en banco de datos (Simes, 1987)	Teóricamente alcance a todas las publicaciones	Banco datos incompleto actualmente
Método de Sugita (Sugita et al., 1992)	Posible corrección sesgo publicación	Estudio y Odds ratio homogéneas
Método bayesiano (Givens, Smith, & Tweedie, 1997)	Posible corrección sesgo publicación	Complejidad Demostración eficacia método
Modelo fijo p-uniforme (van Assen, van Aert, & Wicherts, 2015)	Gran potencial	Solo considera estudios significativos Desarrollo escaso

Tabla 6. Métodos estadísticos exclusivamente cuantitativos (método formal).

7.2.1. Proporción de estudios significativos (Sterling, 1959)

En este método analiza el número de estudios significativos publicados en relación a los no significativos publicados. Lo que se encontró de una forma descriptiva es que los estudios significativos publicados eran mayores que los no significativos. Lo que suponen que es debido al sesgo de publicación, aunque no pueden cuantificar su impacto teóricamente y tampoco detectar el tipo de sesgo de publicación.

7.2.2. Fail-safe N (Orwin, 1983; Rosenthal, 1979)

Analiza cual sería el número exacto de estudios nulos/negativos no publicados necesarios para producir un cambio de significación. Sin embargo, la estimación del Fail-safe N depende en gran medida del efecto medio de la intervención que se asume para los estudios no publicados (Iyengar 1988), y los métodos disponibles conducen a estimaciones muy diversas del número de estudios adicionales siendo propenso a la mala interpretación y al mal uso (Becker, 2006; Higgins & Green, 2011).

Una manera de combinar los p-valores es la transformación en valores Z . A partir de esa Z combinada se puede obtener el p-valor combinado y se puede estimar el número de artículos que sería necesario para pasar de unos resultados significativos a no significativos. El resultado es en si mismo una Z que puede meterse en una tabla para encontrar la asociación de p-valores (Rosenthal, 1979):

$$Z_c = \sum_{i=1}^k \frac{Z_i}{\sqrt{k}}$$

donde Z_c es la nueva Z combinada, k es el numero de estudios combinados, y \bar{Z}_i es el valor Z obtenido a partir del p-valor del i ésimo estudio a combinar.

Rosenthal, a partir de esta expresión, define k_0 como el número de p-valores necesarios para que el Z_c se convierta en el valor crítico Z_α , (al 5% en el contraste unilateral es 1,65) y por lo tanto, pase de un valor significativo a un valor no significativo. La expresión para obtener ese conjunto de estudios no publicados que cambia la significación viene dado por:

$$k_0 = \frac{k(Z_c^2 - Z_\alpha^2)}{Z_\alpha^2}$$

Donde k_0 es el número de estudios necesarios para cambiar la significación; k es el número de estudios recuperados, Z_c es el valor Z a partir de la combinación de los p-valores; Z_α es el valor Z que marca la significación a un nivel de significación α .

El autor se plantea cuál sería el resultado del test que indicara un número poco probable de estudios nulos/negativos no publicados. Para ello plantea lo que él llama el nivel de tolerancia:

$$\text{Nivel de tolerancia} = 5 \times k + 10,$$

donde k es el total de estudios integrados en el metaanálisis y 10 es el número mínimo de estudios nulos/negativos no publicados cuando $k = 1$.

Por lo tanto, el resultado del test de Fail-safe N en una cantidad superior al nivel de tolerancia indicaría probable ausencia de sesgo de publicación.

La crítica de este método es que considera que todos los estudios no recuperados presentan un resultado nulo. Iyengar & Greenhouse (1988) sugieren una modificación de este método donde no es tan restrictivo en relación a los estudios no recuperados. Se basa en una distribución Normal truncada.

Sin embargo, la combinación de p-valores es una metodología muy poco utilizada para combinar los resultados de los artículos en metaanálisis. El tamaño del efecto es el valor más utilizado en las integraciones metaanalíticas. Por ello, Orwin (1983) propone un método similar al de Rosenthal basado en la diferencias de medias estandarizada como tamaño del efecto. Su propuesta pretende calcular el número de estudios para reducir el tamaño del efecto obtenido en la integración a un valor pequeño que sea acorde con la hipótesis nula de no diferencias. Aunque se propuso para la diferencia de medias estandarizada, se puede utilizar en otros tamaños del efecto como el coeficiente de correlación o el Odds ratio:

$$d_c = \frac{k_0 (\bar{d}_0) + k_{fs} (\bar{d}_{fs})}{k_0 + k_{fs}}$$

d_c es el tamaño del efecto que se obtendría si hubiera k_0 estudios con un tamaño del efecto medio nulo (d_{fs}); \bar{d} es el tamaño del efecto medio obtenido en la integración y k el número de estudios integrados. Si se despeja k_0 la expresión queda de la siguiente manera:

$$k_0 = \frac{k(d - d_c)}{d_c}$$

Su principal desventaja es que depende de la estimación, en gran medida, del efecto medio de la intervención que se asume para los estudios no publicados y los métodos disponibles conducen a estimaciones muy diversas del número de estudios adicionales.

7.2.3. Método de Hackshaw (Hackshaw et al., 1997)

Es un método que intenta calcular, si no existe auténtica asociación entre estudios, cuantos estudios no significativos serían necesarios para conseguir un nivel de significación de $\rho = 0.05$. Es un método similar al Fail-safe N, donde, por ejemplo, si se detectan 7

estudios significativos y considerando independientes, la probabilidad de que tal resultado surja por casualidad es ($p < 0,05$). Asumen no asociación, exponiendo que la probabilidad de que tal resultado surja por casualidad es 1 en 40, por lo que se requeriría un total de 280 estudios (7×40) para generar los siete significativos. Concluyen que es inverosímil que haya hasta 241 estudios inéditos de 39 publicados.

En un método que tiene similares inconvenientes a los vistos en el Fail-safe N, con determinadas presunciones a priori en sus estudios que pueden estar condicionadas.

7.2.4. Métodos truncados (Hedges, 1984)

Calcula un estimador máximo verosímil del tamaño del efecto a partir de una función de probabilidad basada en el muestreo truncado y bajo el supuesto de publicación de todos los estudios con resultados significativos.

Se parte de la función de densidad $f(x|\delta, n, \alpha)$ para el estimador del tamaño del efecto, d , bajo el supuesto de partida y considerando tamaños muestrales iguales:

$$f(x|\delta, n, \alpha) = \begin{cases} \frac{h(x|\delta, n, \alpha)}{A(\delta, n, \alpha)} & \text{si } |x| > \sqrt{\frac{2}{n}} t_{\alpha; 2n-2}, \\ 0 & \text{en otro caso} \end{cases}$$

donde $h(x|\delta, n, \alpha)$ es la función de densidad de una distribución $\sqrt{\frac{2}{n}}$ veces una t no central con $(2n - 2)$ grados de libertad y como parámetro central $\delta\sqrt{\frac{2}{n}}$ y $A(\delta, n, \alpha)$ es la probabilidad que una t no central, con $(2n - 2)$ grados de libertad, con parámetro central $\sqrt{\frac{2}{n}} \delta$, sea mayor que $t_{\alpha; 2n-2}$ o menor que $-t_{\alpha; 2n-2}$.

A partir de la distribución truncada se construye una función de verosimilitud para estimar el tamaño del efecto poblacional, δ :

$$L(\delta) \prod_{i=1}^k \frac{h(d_i|\delta, n_i, \alpha)}{A(\delta, n_i, \alpha)}$$

El planteamiento de este modelo es el problema del sesgo inducido por la no publicación de las estimaciones del tamaño del efecto correspondientes a diferencias significativas estadísticamente no significativas, lo cual significa asumir que solo no se

publican los estudios no significativos. Esto puede ser un error pues pueden aparecer estudios significativos no publicados a priori (Martín-Vallejo, 1995).

7.2.5. Modelo de función ponderada (Cooper & Hedges, 1994a; Vevea & Woods, 2005)

Teoría de función ponderada, es una generalización del muestreo truncado, que se basa en la premisa de que un estudio se incluye en el análisis con una probabilidad que se determina por su resultado, siendo las probabilidades de selección relacionados con diferentes resultados posibles a través de una función ponderada (Cooper & Hedges, 1994a).

Trabajos recientes han examinado la posibilidad de evaluar la solidez en un rango de funciones de ponderación, evitando la necesidad de un gran número de estudios (Veeva & Woods, 2005).

En este estudio, asumen un modelo estadístico que consta de dos partes: un modelo para el tamaño del efecto y un modelo para el proceso de selección. Aquí la parte relacionada con el tamaño del efecto puede ser para modelo efectos fijos y aleatorios, y la selección del proceso tiene parámetros con valores particulares (Cooper & Hedges, 1994) .

Cualquiera de los modelos de metaanálisis puede ser modificado pensando en el parámetro θ como representación de la media de una distribución de efectos. Es habitual (y a menudo razonable) asumir una distribución normal de los efectos, por lo que:

$$T_i \sim N(\theta_i, \sigma^2 + \tau^2),$$

donde τ^2 es la varianza entre estudios. La variabilidad total T_i ahora incluye un componente σ_i^2 que cuantifica la variación asociada con las muestras de personas u otras unidades principales dentro del estudio, así como la varianza τ^2 , la cual cuantifica la variación que aumenta desde las muestras de las población en estudio a la distribución de la población posible.

Cuando θ_i tiene una valor común θ . para todos i , esto es el modelo efectos aleatorios simple. Si θ representa la media de una distribución del modelo efectos aleatorios que es normal con una τ^2 . Si θ_i es el resultado de un modelo lineal, el resultado es un análogo de efectos aleatorios de un modelo de efectos fijos con co-variables, llamado modelo efectos mixtos.

Para la selección, dada una densidad $f(x)$ y una función ponderada negativa $w(x)$, la función:

$$\frac{w(x)f(x)}{\int w(x)f(x)dx}$$

será la densidad que representa la densidad original ponderada de $w(x)$, que depende de la magnitud y la varianza de los tamaños del efecto individuales de sus valores de una cola probabilística. Y esto queda:

$$w(p_i) = \begin{cases} \omega_1 \text{ if } 0 < p_i \leq a_1; \\ \omega_j \text{ if } a_{j-1} < p_i \leq a_j; \text{ y} \\ \omega_k \text{ if } a_{k-1} < p_i \leq 1, \end{cases}$$

donde a_1, \dots, a_k representan puntos de corte, p_i el valor probabilístico de una cola del i ésimo estudio, y los $\omega_1, \dots, \omega_k$ son parámetros a estimar (el primer peso fue fijado 1.0 para tratar una indeterminación).

Tiene una gran dificultad para el cálculo, y no implementado en el software que generalmente se utiliza para el estudio de datos metaanalíticos (Higgins & Green, 2011).

7.2.6. Muestreo invariante en banco de datos (Simes, 1987)

En este trabajo se desarrolla un modelo para reunir los resultados de los estudios clínicos, que está libre de publicación parcial. Se propone que se establezca un registro internacional de todos los estudios clínicos con los objetivos y puntos finales de cada ensayo claramente definidos en el registro. De esta manera, para cada tema terapéutico, los investigadores pueden seleccionar una cohorte de estudios clínicos independientemente de los resultados del ensayo. El enfoque se ilustra utilizando el registro del International Cancer Research Data Bank (ICRDB).

Actualmente la red Cochrane está haciendo un esfuerzo para conseguir una base de datos importante de estudios y revisiones sistemáticas, además de otras fuentes de información on-line cada vez más extendidas pero que no impiden que ciertos estudios puedan no aparecer publicados en ellas o aparezcan en revistas de menor calado o nichos limitados.

7.2.7. Método de Sugita (Sugita et al., 1992)

Este método se basa en la razón de probabilidad resumida calculada a partir de las Odds ratios de los estudios publicados en un metaanálisis pudiendo estar sobreestimada debido al sesgo de publicación.

Este método permite calcular la relación de probabilidades resumida de estudios hipotéticos no publicados de Odds ratios de estudios publicados indirectamente, empleando un método que asume el logaritmo natural de las Odds ratio que se distribuyen normalmente.

La Odd ratio Or_H y sus intervalos de confianza (CI) (Or_{UH}, Or_{LH}) de los estudios no publicados hipotéticamente puede ser obtenido indirectamente de los Odds ratios Or_i y sus CIs (Or_{Ui}, Or_{Li}) de los estudios publicados, empleando un método basado en momentos. A partir de la resolución de un conjunto de ecuaciones simultáneas que contiene el segundo, tercero y cuarto momento, entonces estimar la Odds ratio resumen \overline{Or}_A con sus CIs (Or_{UA}, Or_{LA}) en todos los estudios, incluyendo no solo estudios publicados sino estudios no publicados en el metaanálisis.

Los aspectos técnicos se pueden ver en Sugita et al. (1992).

La limitación de este método es que solo sirve para cuando estos estudios son homogéneos en calidad y sus Odds ratios son homogéneos en cantidad. No es un método implementado en el software generalmente utilizado, ni tampoco está validado.

7.2.8. Método bayesiano (Givens et al., 1997)

Este enfoque estima y ajusta para el sesgo de la publicación. La estimación se basa en un principio de aumento de datos dentro de un modelo jerárquico, y el número y los resultados de los estudios no observados se simulan usando métodos de muestreo de Gibbs. Esta técnica se utilizó para producir un ajuste cuantitativo para el metaanálisis del tabaquismo pasivo. Asume que la probabilidad de publicación depende solamente del p-valor.

El modelo de efectos aleatorios tiene por termino:

$$Y_j = \Delta + \beta_j + \varepsilon_j,$$

donde $\beta_j \sim N(0, \tau^2)$ introduce la heterogeneidad entre estudios. Partiendo de aquí, la probabilidad de que los datos observados $Y = (Y_1, \dots, Y_n)$ es:

$$p(Y|\Delta, \tau^2, \sigma^2)$$

$$\propto \prod_{j=1}^n \exp\left(-\frac{1}{2} \frac{(Y_j - \Delta)^2}{(\tau^2 + \sigma_j^2)}\right) / \sqrt{\tau^2 + \sigma_j^2}.$$

Estiman el logaritmo del riesgo relativo del j ésimo estudio perdido anotado como Z_i para $j = (n+1), \dots, (n + m)$, y anota $Z = \{Z_i\}$. Lo mismo para X y para Y . Asumiendo modelo efectos aleatorios con estudios perdidos:

$$Z_j = \Delta + \beta_j + \epsilon_j,$$

donde $\beta_j \sim N(0, \tau^2)$, y $\epsilon_j \sim N(0, \sigma_j^2)$ son mutuamente independientes. Anotamos que ahora σ^2 incluye la varianza de los estudios latentes al igual que al de los estudios observados.

Sus autores creen no estar muy lejos de la verdad, hay evidencia de sesgo de publicación en su estudio, pero su impacto probablemente no es especialmente fuerte en este metaanálisis debido a la aparente tendencia general de los datos hacia un pequeño efecto positivo del tabaquismo pasivo sobre el riesgo de cáncer de pulmón.

Continúan diciendo que las limitaciones de la calidad de los datos y la aparente debilidad del efecto del tabaquismo pasivo significa que el análisis está lejos de ser concluyente y es poco probable que otros estudios observacionales puedan afectar esta conclusión general.

Además, todos ellos se basan en supuestos que en la práctica pueden no ser ciertos. Otro problema añadido es la imposibilidad de probar rigurosamente los métodos, debido a la falta de un "estándar de oro" en el conjunto de estudios.

7.2.9. Modelo fijo p-uniforme (van Assen et al., 2015)

P-uniforme es un método para realizar metaanálisis que permite probar el sesgo de publicación y estimar un tamaño de efecto fijo bajo sesgo de publicación. La idea básica de P-uniforme es que la distribución de los p valores es uniforme. Además puede usarse como un análisis de sensibilidad para detectar el sesgo de publicación en metaanálisis.

Hace dos suposiciones:

- En primer lugar, al igual que en otros métodos, se considera un modelo de efectos fijos en lugar de aleatorios. La heterogeneidad puede ser más común en los estudios observacionales.
- En segundo lugar, supone que todos los estudios con hallazgos estadísticamente significativos tienen la misma probabilidad de ser publicados e incluidos en el metaanálisis. P-uniforme no hace suposiciones acerca de la magnitud de la probabilidad de publicación o la probabilidad de que los estudios estadísticamente insignificantes sean publicados.

P-uniforme del sesgo de publicación se basa en la desviación de la distribución del valor $\rho^{\hat{\mu}}$ de la distribución uniforme, donde $\hat{\mu}$ es igual a la estimación del tamaño del efecto del metaanálisis tradicional de efecto fijo.

Este test está basado en contrastes unilaterales de hipótesis nulas donde $\mu = \hat{\mu}$, si el tamaño del efecto de la población es igual al tamaño del efecto estimado de un metaanálisis de efectos fijos. La idea básica es que la hipótesis es rechazada si la distribución del valor $\rho^{\hat{\mu}}$ es una distribución condicional. Asumen un test de $\mu = \mu^*$ para definir esta distribución condicional. La definición usa la distribución muestral del tamaño del efecto $M_i^{\mu^*}$ de todos los estudios i , asumiendo de $\mu_i = \mu^*$. La distribución condicional del p-valor $p_i^{\mu^*}$ se define entonces :

$$p_i^{\mu^*} = \frac{p(M_i^{\mu^*} \geq \hat{\mu}_i)}{p(M_i^{\mu^*} \geq M_i^{CV})}.$$

$p_i^{\mu^*}$ representa la probabilidad de observar el efecto $\hat{\mu}_i$, o mayor, condicional tanto en un efecto de población μ^* y un p-valor significativo (cuando se probó contra la hipótesis nula de ningún efecto).

El método solo considera estudios con un efecto estadísticamente significativo, y por lo tanto descarta aquellos con un efecto no significativo.

El método todavía necesita un mayor desarrollo, pero puede convertirse en la técnica para examinar el sesgo de publicación y estimar los efectos de la población en la revisión metaanalítica.

7.3. Métodos representación datos de forma gráfica (método informal)

En la siguiente tabla (**Tabla 7**) aparecen los métodos que se han considerado en este segundo grupo, exponiendo sus ventajas y limitaciones.

Método	Ventaja	Limitaciones
Funnel Plots (Light & Pillemer, 1984)	Fácil	Subjetividad interpretación
Contour-enhanced funnel plot (Peters et al., 2008)	Fácil Zonas significación claramente delimitadas	Subjetividad interpretación
Metaanálisis acumulado (Borenstein et al., 2009; Lau et al., 1992, 1995)	Fácil Muestra la magnitud del cambio	No es un método de análisis

Tabla 7. Métodos representación datos de forma grafica (método informal).

7.3.1. Funnel Plot (Light & Pillemer, 1984)

Un Funnel Plot es un diagrama de dispersión del efecto del tratamiento contra una medida de precisión del estudio. Este diagrama lo que trata de representar es la existencia gráfica de sesgo de publicación visualizando presencia o no de asimetría en la representación.

El diagrama va a representar la precisión de la estimación del efecto de la intervención que aumenta cuando se incrementa el tamaño del estudio. Por lo tanto, las estimaciones del efecto de los estudios pequeños se dispersarán más ampliamente en la parte inferior del gráfico, y la dispersión se estrechará entre los estudios más grandes.

La heterogeneidad verdadera en los efectos de la intervención también puede dar lugar a la asimetría del gráfico de embudo. Como ejemplo, se puede observar un beneficio significativo solamente en los pacientes con alto riesgo para el resultado que esté afectado por la intervención, y habitualmente es más probable que estos pacientes con alto riesgo se incluyan en los estudios iniciales más pequeños (Davey Smith & Egger, 1994; Higgins & Green, 2011).

Hay una serie de condiciones a tener en cuenta con este tipo de gráfico. La primera sería la unidad de medida que escogeremos en cada uno de los ejes a representar según **Figura 5**. Aquí recorre un número de diferentes opciones justificando qué medida sería la más adecuada para representar en cada eje los valores de los estudios en cuestión en el caso de datos binarios.

Table 5.1 Possible choices of axis in funnel plots for binary outcomes: advantages, disadvantages and recommendations.		
Axis / measure	Advantages and disadvantages	Recommendations
Vertical axis		
Standard error	Funnel shape with straight 95 % confidence lines. Emphasis of the plot is on smaller studies where bias is more likely.	The best choice in most cases.
Precision (inverse standard error)	Plot is not funnel-shaped; 95 % confidence lines are curved. Emphasis of the plot is on larger studies; smaller studies are compressed at the bottom.	An option in studies which focus on a comparison of meta-analyses of small studies with subsequent large studies. Not recommended.
Variance (squared standard error)	Plot is not funnel-shaped; 95 % confidence lines are curved. Emphasis of the plot is on smaller studies where bias is more likely.	
Inverse variance	Plot is not funnel-shaped; 95 % confidence lines are curved. Emphasis of plot is on larger studies, smaller studies are compressed at the bottom.	An option in studies which focus on a comparison of meta-analyses of small studies with subsequent large studies, but precision would usually be better. Invalid choice.
Sample size or log sample size	Expected shape of plot in absence of bias is unpredictable.	
Horizontal axis		
Log odds ratio	Scale is not constrained and plots have the same shape whether the outcome is defined as occurrence or non-occurrence of the disease. Odds ratios may be misinterpreted as risk ratios.	The best choice in most cases.
Log risk ratio	Readily understood measure. Scale is naturally constrained so that heterogeneity may be introduced if the event rate is high.	Valid choice in many cases, but not recommended if the event rate is high. Can give different conclusions depending on outcome definition. Not recommended in most cases.
Risk difference	Readily understood measure. Often associated with increased heterogeneity which may result in funnel plot asymmetry which is not apparent when ratio measures are used.	

Figura 6. Posibles elección medidas ejes de un Funnel Plot. Extraído de : Rothstein, H., Sutton, A., & Borenstein, M. (2005). *Publication Bias in Meta-Analysis*. (H. R. Rothstein, A. J. Sutton, & M. Borenstein, Eds.), *BMC Medical Research Methodology* (Vol. 1). Chichester, UK: John Wiley & Sons, Ltd. <https://doi.org/10.1002/0470870168>

En caso de datos continuos, hay discusión sobre que medida debe ser elegida según que eje del gráfico. Lo que queda claro es que debe ser una medida que establezca la varianza de cada estudio. Se puede utilizar en el eje de coordenadas la inversa del error estándar (H. Rothstein et al., 2005).

Como ejemplo, se puede ver en la **Figura 6**, que en el eje abscisas se representa la medida del tamaño del efecto más adecuada en este caso la Odd ratio y en el lado del eje de ordenadas es el error estándar. En este gráfico se puede observar la asimetría en el lado inferior derecho del gráfico detectando la presencia del sesgo de publicación.

Los estudios mayores se sitúan en la zona alta del gráfico de forma poco dispersa, mientras que los estudios más pequeños van a la zona inferior estando alejados unos de otros.

Como principal ventaja destacar la facilidad en la interpretación, trasladada a otros métodos que arrancan desde este método. Continúa siendo conflictiva la interpretación subjetiva que de estos gráficos se realiza dando lugar a diferentes estudios, mejorándola con modelos gráficos como los que se exponen a continuación.

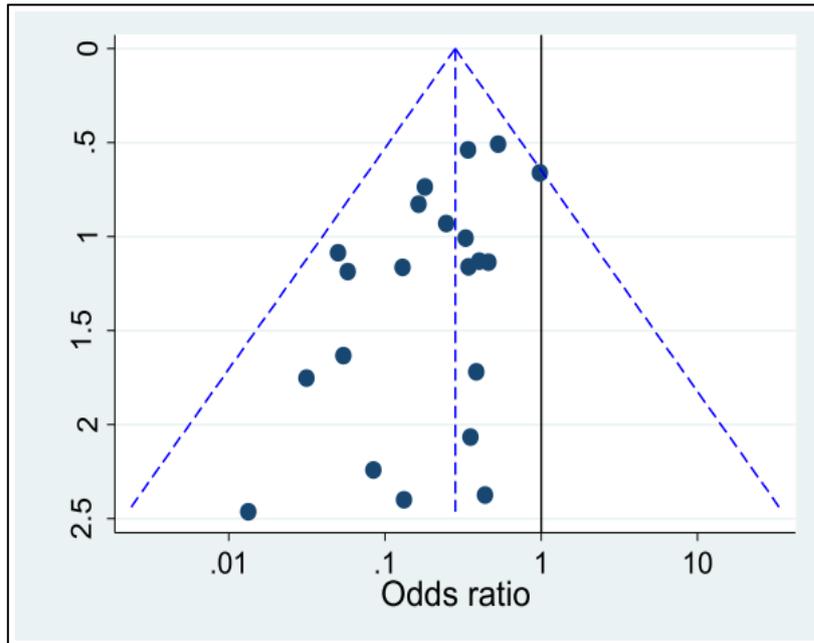


Figura 7. Ejemplo de Funnel Plot. Extraído de : Higgins, J. P. T., & Green, S. (2011). *Manual Cochrane de revisiones sistemáticas de intervenciones*. Cochrane, (March), 1–639. <https://doi.org/Available from www.cochrane-handbook.org>.

7.3.2. Contour-enhanced Funnel Plot (Peters et al., 2008)

Este gráfico propone establecer unas áreas relacionadas con la significancia de los estudios. Incluye áreas con la extensión de (p -valor = 0,01; 0,05; 0,1), como generales, pero se podrían establecer significación estadística menos estricta o más acorde con el tipo de estudios a analizar.

El objetivo de este tipo de gráfico es ver en que áreas de significancia se encajan los distintos estudios a analizar. Si los estudios parecen faltar en áreas no significativamente estadísticas, **Figura 8(a)**, puede sugerir que la asimetría se debe a un sesgo de publicación. Por el contrario, si los supuestos estudios faltantes se encuentran en áreas de mayor significación estadística, **Figura8(b)**, esto sugeriría que la causa de la asimetría puede ser debida a otros factores como la calidad de los estudios (Peters et al., 2008).

A pesar de ser gráficamente fácil de interpretar, continúa arrastrando el problema de la subjetividad de la interpretación. Por lo tanto, los autores intentan evaluar la asimetría del grafico con métodos estadísticos complejos ya enumerados pero que son de bajo poder estadístico o incluso solo sirven para determinado tipo de datos (continuos, dicotómicos,...)

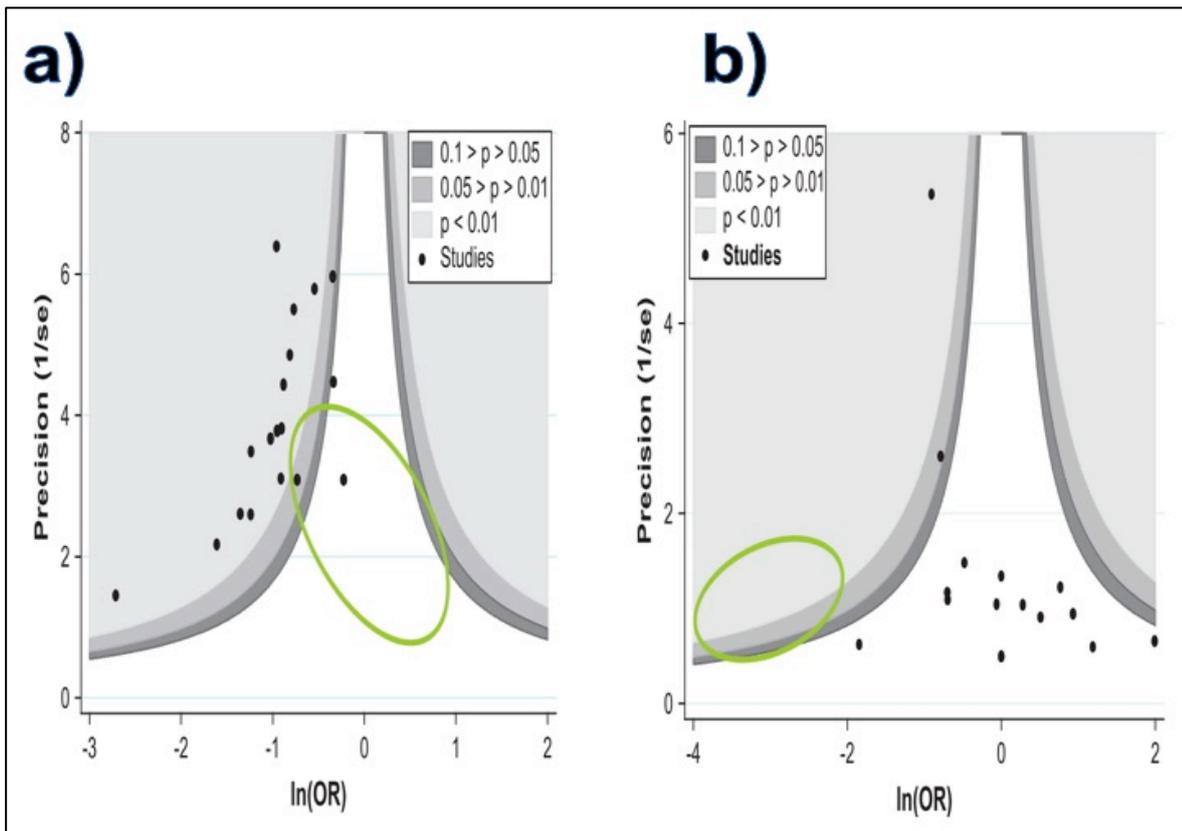


Figura 8. Ejemplo de Contour-enhanced Funnel Plot Extraído de : Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., & Rushton, L. (2008). Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *Journal of Clinical Epidemiology*, 61(10), 991–996. <https://doi.org/10.1016/j.jclinepi.2007.11.010>.

7.3.3. Metaanálisis acumulado (Borenstein et al., 2009; Lau et al., 1992, 1995)

El metaanálisis acumulado es un mecanismo para mostrar los resultados de una serie de análisis separados en una tabla. Los estudios suelen clasificarse cronológicamente, lo que muestra cómo se ha acumulado la evidencia (y posiblemente cómo han cambiado los resultados) con el tiempo. Sin embargo, los estudios también pueden ser ordenados por otras variables, para mostrar cómo los resultados cambian en función de otro factor (como la calidad del estudio).

Se puede utilizar metaanálisis acumulados para mostrar el sesgo de publicación y su posible impacto. El problema que se plantea es que los grandes estudios son imparciales, pero los estudios más pequeños pueden tender a sobreestimar el tamaño del efecto. Podríamos realizar un metaanálisis acumulado, comenzando con los estudios más grandes en la parte superior y agregar los estudios más pequeños en la parte inferior. Si el efecto fue inicialmente pequeño cuando teóricamente no hay sesgo de publicación (estudios grandes), y luego aumentó con los estudios más pequeños, nos muestra que el tamaño del efecto estuviera relacionado con el tamaño de la muestra.

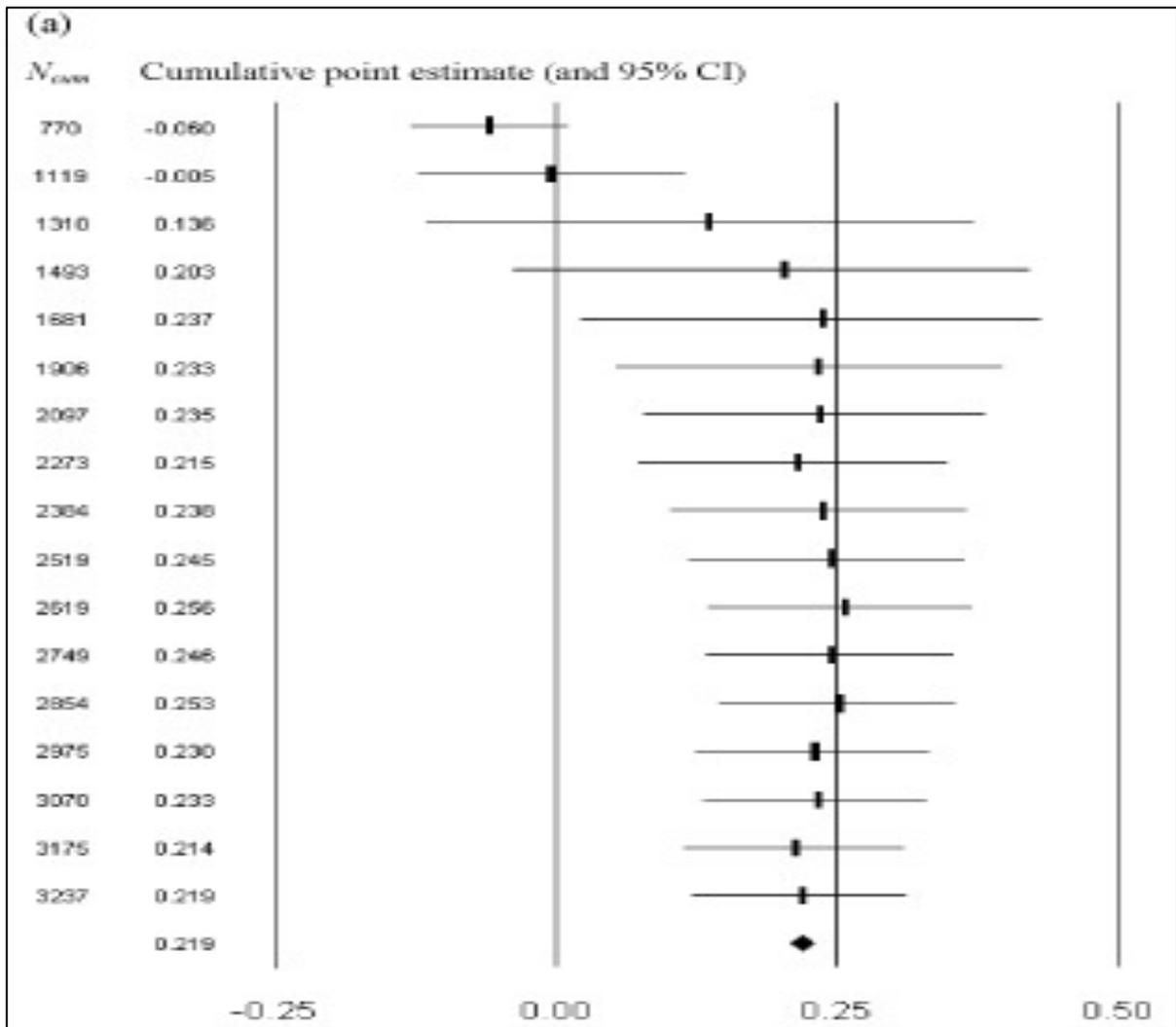


Figura 9. Ejemplo de Funnel Plot Acumulado. Extraído de Banks, G. C., Kepes, S., & Mcdaniel, M. A. (2012). *Publication Bias: A call for improved meta-analytic practice in the organizational sciences. International Journal of Selection and Assessment*, 20(2), 182–196. <https://doi.org/10.1111/j.1468-2389.2012.00591.x>.

Como se muestra en la **Figura 9**, para el metaanálisis acumulado, añaden primero los estudios más grandes y después se van sumando los estudios más pequeños. Hay una tendencia a la parte positiva en cuanto se van añadiendo estudios más pequeños, para finalmente estabilizarse. Hay una gran diferencia entre el resultado del metaanálisis acumulado con los dos primeros estudios y el metaanálisis acumulado final. Los autores contrastaron con el método Trim And Fill para evidenciar que existía sesgo de publicación tal y como demuestra la tendencia del metaanálisis acumulado de la **Figura 9**.

Una ventaja del metaanálisis acumulado es que muestra no solo si hay un cambio en el tamaño del efecto, sino que también la magnitud del cambio.

El metaanálisis acumulado es un mecanismo de visualización, no de análisis. Si nuestro objetivo fuera probar que el tamaño del efecto ha cambiado con el tiempo, el enfoque sería utilizar el análisis de subgrupos o meta-regresión.

7.4. Métodos estadísticos cuantitativos con representación gráfica (método correcto)

En la siguiente tabla (**Tabla 8**) aparecen los métodos que se han considerado en este tercer grupo, exponiendo sus ventajas y limitaciones.

Método	Ventaja	Limitaciones
Prueba de correlación de rangos ajustado Begg y Mazumdar (Begg & Mazumdar, 1994)	Análogo estadístico directo del Funnel Plot	Precaución metaanálisis pequeños
Regresión lineal de Egger (Egger et al., 1997)	Mayor poder estadístico que la prueba de Begg y Mazumdar's prueba de correlación de rango ajustado	No prueba directamente el sesgo de publicación
Regresión de Macaskill sobre gráfico de embudo (Macaskill, Walter, & Irwig, 2001)	Frecuencia de falsos positivos (error alfa) es mucho menor (OR) Orienta posibles causas subyacentes del sesgo de publicación	No es capaz de determinar si una heterogeneidad subyacente o residual
P-curve (Simonsohn & Nelson, 2013)	Visualmente interpretable	Estudios y los p-valores deben ser seleccionados apropiadamente
Trim and Fill (Duvall & Tweedie, 2000)	Fácil interpretación	No garantiza la intervención del método, al no conocer el mecanismo de sesgo de publicación

Tabla 8. Métodos estadísticos cuantitativos con representación gráfica (método correcto).

7.4.1. Prueba de correlación de rangos ajustado Begg y Mazumdar (Begg & Mazumdar, 1994)

La prueba de correlación de rangos de Begg y de Mazumdar utiliza la correlación de rangos, Tau de Kendall, entre los tamaños del efecto estandarizados y sus varianzas. En ausencia de sesgo de publicación, no hay una relación entre los tamaños del efecto y su varianza, pero si existe sesgo de publicación una gran varianza debería asociarse a un gran tamaño del efecto, ya que los estudios pequeños, que tienen varianzas más altas, tienen más probabilidad de ser publicados si muestran tamaños del efecto grande.

Por ejemplo, grandes estudios, a pesar de su menor varianza muestral, puede por razones de costo recolectar menos datos que en pequeños estudios, permitiendo así un menor control de confusión y, por lo tanto, una estimación de efecto diferente (Thornton & Lee, 2000).

En resumen, utilizan como medida de sesgo la diferencia entre la expectativa de la media ponderada de los tamaños de efecto bajo publicación selectiva y el tamaño del efecto real.

Anotemos $\{t_i\}$ y $\{v_i\}$ son los tamaño del efecto estimado y la varianza muestrales de k estudios en el metaanálisis, $i = 1, \dots, k$. Para construir un test de correlación de rango válido, es necesario estabilizar la varianza estandarizando por el tamaño del efecto:

$$t_i^* = (t_i - \bar{t}) / (v_i^*)^{1/2},$$

donde al final

$$v_i^* = v_i - \left(\sum v_j^{-1} \right)^{-1}$$

es la varianza de $t_i - \bar{t}$. La varianza será aproximadamente inversamente proporcional al tamaño muestral en los estudios, siendo el test similar a la correlación del tamaño del efecto con el tamaño muestral.

Se ha utilizado la prueba de correlación de rango basada en la tau de Kendall, lo cual implica enumerar el número de pares de estudios que se clasifican en el mismo orden con respecto a los dos factores (es decir, t^* y v). Si este número es x , y y es el número de pares de estudios clasificados en el orden opuesto, entonces la estadística de prueba normalizada es:

$$z = (x - y) / [k(k - 1)(2k + 5) / 18]^{1/2}$$

La selección del modelo la realiza con la premisa de una distribución normal

$$t \sim N(\delta, v_i).$$

Después de generar t_i , y el estudio es seleccionado para su inclusión en el metaanálisis con probabilidad determinada por el modelo de selección apropiado caracterizado por la función ponderada :

$$g_i(t_i) = \frac{f_i(t_i)w_i(t_i)}{\int_{-\infty}^{\infty} f_i(x)w_i(x)dx},$$

donde $f_i(t_i)$ es la densidad de probabilidad subyacente de t_i antes de la selección, y $w_i(t_i)$ es la función ponderada, la probabilidad que el estudio sea publicado como ejemplo del autor.

No obstante, la prueba debe interpretarse con precaución en pequeños metaanálisis. En especial, el sesgo no puede descartarse si la prueba es no significativa. Sus autores lo recomienda para estudios grandes.

7.4.2. Regresión lineal de Egger (Egger et al., 1997)

En este método calcula el tamaño del efecto estandarizado definido como la Odds ratio dividida por su error estándar y se hace una regresión sobre su precisión (definida como la inversa de su error estándar). Si la intersección no pasa por el origen, esto indica asimetría en la dispersión y se puede utilizar como medida del grado de asimetría.

Ecuación de regresión:

$$SND = a + b \times \text{precisión}$$

, donde SND es el tamaño del efecto estandarizado.

Esta regresión sería análoga a hacer la regresión ponderada de los tamaños del efecto frente a la varianza del estimador (es decir el error estándar al cuadrado) y ponderado por el inverso de la varianza del estimador. Pero en lugar de ver si la ordenada en el origen es significativa se mira si la pendiente es significativa. A partir de esta regresión se han propuesto otros métodos cambiando la variable independiente y su ponderación.

Esta regresión no prueba directamente el sesgo de publicación, sino la relación entre los tamaños del efecto observado y el error estándar. Si existe una relación, esto suele implicar asimetría en la gráfica, lo que a su vez puede indicar la presencia de sesgo de publicación (Viechtbauer, 2010).

Este método, si tiene mayor poder estadístico que la prueba de correlación de rangos ajustado de Begg y Mazumdar.

7.4.3. Regresión de Macaskill sobre gráfico embudo (Macaskill et al., 2001)

Macaskill et al. (2001) critican la regresión de Egger debido a que hay un aumento en el incremento de error tipo I, debido a que se incumplen el supuesto de homocedasticidad de la regresión. Por ello, proponen una nueva regresión donde en principio, mejora los resultados. Esta regresión lineal simple, donde el Funnel Plot representa la magnitud del efecto (t_i) como variable dependiente y el tamaño del estudio (n_i) como variable independiente. Las observaciones son ponderadas por el inverso de la varianza del estimador para permitir la posible heterocedasticidad (FIV) (Palma Pérez & Delgado Rodríguez, 2006).

Cuando no hay sesgo de publicación, la pendiente de regresión tiene un valor esperado de cero. Una pendiente distinta de cero sugeriría una asociación entre el efecto del tratamiento y el tamaño de la muestra, posiblemente debido al sesgo de publicación.

Nos orienta en las posibles causas subyacentes del sesgo de publicación, pero al mismo tiempo no es capaz de determinar si una heterogeneidad subyacente o residual es la que pueda influir en este sesgo como única causa.

Como desventaja el autor describe que existe una disminución de la potencia con el aumento del efecto del tratamiento, dado que la probabilidad de que un estudio no sea publicado disminuye a medida que el efecto del tratamiento se hace más grande. Los tamaños de muestra más grandes también reducen la potencia cuando el verdadero efecto es distinto de cero, porque se publican más estudios.

Existe una mejora de este método que sugiere una regresión lineal ponderada simple con $\ln OR$ como variable dependiente, y usar como variable independiente el inverso del tamaño de muestra total, presentado la ventaja de potencia estadística similar al método de Egger pero la frecuencia de falsos positivos (error alfa) es mucho menor. Esta es una modificación menor de la prueba de Macaskill, que utiliza la inversa del tamaño total de la muestra como variable independiente en lugar del tamaño total de la muestra (Peters, 2006).

Aunque Macaskill, en principio, podría utilizarse para todos los tamaños del efecto, el autor lo probó con Odds ratio y no está probado con la diferencia de medias estandarizada.

7.4.4. P-Curve (Simonsohn & Nelson, 2013)

La P-curve es la distribución de p-valores estadísticamente significativos para un conjunto de estudios ($p < 0,05$). Debido a que solo se espera que los efectos verdaderos generen P-curvas inclinadas a la derecha (que contienen valores de p más bajos (.01s) que altos (.04s) solo las p-curvas inclinadas a la derecha son el resultado del efecto esperado.

Al detectar que la información con un efecto verdadero ha sido seleccionada debido a este efecto, la P-curve ofrece una solución a problemas inferenciales causados por estudios y análisis fallidos que finalmente eran archivados. Por lo tanto, la P-curve es una forma de distinguir entre la información selectiva y la verdadera.

La distribución de p-valores estadísticamente significativos para un conjunto de hallazgos independientes es la P-curve. Su forma es diagnóstica del valor probatorio de ese conjunto de hallazgos e indica si un conjunto de conclusiones significativas contiene valor

probatorio. En ese momento se puede descartar si la presentación de informes selectivos sobre estudios realizados es la única explicación de estas conclusiones significativas.

Una representación de la P-curve puede inferir si un conjunto de hallazgos contiene valor probatorio examinando las distribuciones de p-valores. Esto es debido a que los p-valores mayores que 0,05(no significativos) son a menudo archivados y, por lo tanto, no publicados. Sin embargo, todos los p-valores por debajo de 0,05(significativos) son potencialmente publicables y, por lo tanto, observables. Esto nos permite hacer inferencias imparciales de la distribución de p-valores que caen por debajo de 0,05, es decir, a partir de la P-curve.

La principal desventaja es que los estudios y los p-valores deben ser seleccionados apropiadamente. Los p-valores seleccionados deben estar asociados con la hipótesis de interés, estadísticamente independientes de otros valores p seleccionados y distribuidos uniformemente bajo hipótesis nula.

Otra desventaja es que la P-curve no puede distinguir de manera fiable efectos verdaderos y efectos nulos con p-hacking² en la investigación observacional (para denotar cualquier estudio en el que no haya aleatorización en la comparación de los grupos de interés) (Bruns & Ioannidis, 2016).

7.4.5. Trim and Fill (Duvall & Tweedie, 2000)

Este método se basa en el gráfico de embudo interpretando la presencia de asimetría y posteriormente aplicando una formulación matemática donde se reponen/añaden estudios para compensar la presencia de asimetría o no. Indica el número de estudios necesarios para corregir esta asimetría.

Asumen que los n estudios observados, hay k_0 estudios relevantes no observados provocando sesgo de publicación. El valor de k_0 y el tamaño del efecto que esos k_0 podrían haber generado es desconocido.

La situación más difícil que plantea el metaanálisis es porque Δ es desconocido, así como el número y la posición de cualquier estudio perdido este correlacionado con el valor real de Δ .

² P-hacking es encontrar hallazgo estadísticamente significativo que puede reflejar información selectiva en lugar de un verdadero efecto o, lo todo lo contrario, cuando los investigadores deciden no publicar un estudio con resultado no significativo.

El método en si consiste en tres fases: “recortar” (eliminar) los estudios más pequeños que provocan la asimetría del gráfico de embudo, utilizar el gráfico recortado para la estimación del “centro” verdadero del embudo, y reemplazar los estudios omitidos y sus “contrapartes” perdidas alrededor del centro (rellenar). Al tiempo que proporciona una estimación del número de estudios que faltan. Incluso podría mostrar el efecto de la intervención ajustado al realizar un metaanálisis con la inclusión de los estudios rellenos.

Visto en formulación estadística los tres pasos serian:

Paso 1. Estimar $\hat{\Delta}^{(1)}$ usando el estimador RE

$$Y_i^{(1)} = Y_i - \hat{\Delta}^{(1)}, \quad i = 1, \dots, n,$$

y estima $\hat{k}_0^{(1)}$ usando uno de los tres estimadores

$$(\mathbf{R}_0^+ = \max\{0, R_0\},$$

$$\mathbf{L}_0^+ = \left[\max \left\{ 0, L_0 + \frac{1}{2} \right\} \right],$$

$$\mathbf{Q}_0^+ = \left[\max \left\{ 0, Q_0 + \frac{1}{2} \right\} \right]).$$

Paso 2. Basándonos en la simetría, eliminar los valores $k_0^{(1)}$ del extremo derecho de los valores de Y_i , y estimar $\hat{\Delta}^{(2)}$ basado en la simetría del conjunto $n - \hat{k}_0^{(1)}$. El siguiente set de valores centrados quedaría

$$Y_i^{(2)} = Y_i - \hat{\Delta}^{(2)}, \quad i = 1, \dots, n,$$

y estima $\hat{k}_0^{(2)}$.

Paso 3. Eliminar $\hat{k}_0^{(2)}$ valores del lado final derecho, reestimando $\hat{\Delta}^{(3)}$ basado en el conjunto dela izquierda de $n - \hat{k}_0^{(2)}$. El siguiente set construido es

$$Y_i^{(3)} = Y_i - \hat{\Delta}^{(3)}, \quad i = 1, \dots, n$$

y estima $\hat{k}_0^{(3)}$ de $Y_i^{(3)}$.

Paso 4. Continuar hasta que la J iteración, donde $\hat{k}_0^{(J)} = \hat{k}_0^{(J-1)-1} := \hat{k}_0$, y en ese punto tendremos $\hat{\Delta}^{(J)} = \hat{\Delta}^{(J-1)}$.

Paso 5. Rellenar el Funnel Plot usando los valores simétricos obtenidos

$$Y_j^* = 2 \hat{\Delta}^{(j)} - Y_{n-j+1}, \quad j = 1, \dots, \hat{k}_0$$

asociado al error estándar obtenido

$$\sigma_j^* = \sigma_{n-j+1}, \quad j = 1, \dots, \hat{k}_0.$$

Así ya podríamos estimar el tamaño del efecto final $\hat{\Delta}^{(F)}$.

Como ventaja, además de ser visualmente fácil de interpretar, el método de recortar y rellenar no necesita hacer suposiciones acerca de los mecanismos que dan lugar al sesgo de publicación. Pero como desventaja, el mismo se basa en la suposición sólida de que debe haber un gráfico de embudo simétrico. Esto no garantiza que el efecto de la intervención sea igual que el que se observaría en ausencia de sesgo de publicación, ya que no es posible conocer el verdadero mecanismo del sesgo de publicación.

8. Influencia de la heterogeneidad en el sesgo de publicación

Dentro del metaanálisis, una de las principales limitaciones que se investiga es la heterogeneidad entre estudios. Siguiendo la disciplina Cochrane Q, se intenta determinar si el tamaño del efecto resultante del metaanálisis puede ser interpretado frente a la variabilidad de los estudios sobre los que se ha calculado (Cochran, 1954).

Existen diferentes estadísticos, los cuales intentan determinar dentro del metaanálisis, el grado de heterogeneidad entre estudios. El estadístico Q de Cochran es un test que mide la heterogeneidad. La prueba Q de Cochran se calcula:

$$Q = \sum_{i=1}^k W_i(Y_i - M)^2,$$

donde W_i es el peso del estudio, Y_i es el tamaño del efecto del estudio, y M es el efecto global y k es el número de estudios (Borenstein et al., 2009).

Los valores de Q se obtienen comparando estadísticamente χ^2 con una distribución $k - 1$ grados de libertad. Este es un planteamiento clásico de cómo cuantificar el tamaño del efecto en un determinado nivel de confianza (Higgins et al., 2003).

La hipótesis nula sería que todos los estudios tienen un tamaño del efecto común y por lo tanto no habría diferencia significativa entre ellos. Cuando hay diferencia significativa entre ellos, no se cumple la hipótesis nula existiendo heterogeneidad.

Debido a las limitaciones del test Q, como son la falta de potencia, se desarrolló otro estadístico que es más actual que la Q de Cochran. Se conoce por valor I^2 el cual determina en qué porcentaje existe heterogeneidad entre los estudios analizados en el metaanálisis. Según sus autores, la prueba Q de Cochran es un test con poco poder predictivo de heterogeneidad y con valor significativo vago. Por lo tanto, I^2 es un estadístico más fiable y robusto (Higgins et al., 2003).

El valor I^2 describe el porcentaje de variación total entre los estudios que se debe a la heterogeneidad no relacionada con el azar. I^2 puede calcularse fácilmente a partir de los resultados básicos obtenidos a partir de un metaanálisis típico según:

$$I^2 = \frac{Q - df}{Q} \times 100 \%,$$

donde Q es el estadístico de heterogeneidad de Cochran y df los grados de libertad.

Los valores negativos de I^2 se ponen igual a cero, de manera que I^2 está entre 0% y 100%. Un valor de 0% indica que no hay heterogeneidad observada, y los valores mayores muestran una heterogeneidad creciente.

Siendo estos dos métodos los más populares, no hay que olvidar que el índice I^2 en su origen es creado a partir de otro índice conocido por H^2 . Se obtiene de utilizar el Chi-cuadrado de Cochran donde se divide Q por los grados de libertad (df) (Biotext, 1999):

$$H^2 = \frac{df}{Q},$$

y dado que $E[Q] = df$ en ausencia de heterogeneidad, $H^2 = 1$ es ausencia de heterogeneidad entre estudios. Por lo tanto, valores mayores que 1 indican existencia de heterogeneidad (Higgins & Thompson, 2002).

Hay que tener cautela en los resultados obtenidos pues un resultado de presencia de heterogeneidad (rechazo de la H_0 de homogeneidad) es concluyente. En cambio aceptar la H_0 de homogeneidad no supone ausencia de heterogeneidad.

Se debe considerar seriamente la posibilidad que el sesgo sea la causa de la heterogeneidad en las revisiones. Sin embargo, la falta de heterogeneidad no indica ausencia de sesgo, ya que es posible que exista un sesgo consistente en todos los estudios. ¿Es posible predecir la magnitud y la dirección del sesgo? (Higgins & Green, 2011).

Existen otros métodos para la detección de heterogeneidad de carácter gráfico, que visualmente pueden ser más agradables.

Hay autores que mencionan la importancia de la heterogeneidad en la valoración del sesgo de publicación. La relación entre estos dos elementos existe y continúa su investigación en la actualidad. Se han planteado diferentes escenarios para investigar esta relación tal y como se observa en la **Figura 10**. Este autor describe una forma en la que hay que trabajar para conocer la relación existente entre heterogeneidad y sesgo de publicación, intentando conectar de una manera lógica y ordenada las diferentes hipótesis de trabajo posibles (Ioannidis, 2008).

¿Cuánto influye la heterogeneidad en el sesgo de publicación, con qué consecuencias y cuáles son los diferentes escenarios que se pueden encontrar?

La heterogeneidad es común en el metaanálisis y se pueden encontrar ejemplos en las que determinadas decisiones en la toma de muestras por investigadores individuales conducen a la asimetría del diagrama de embudo. Incluso cuando los metaanálisis se centran en una hipótesis clínica específica, puede haber heterogeneidad considerable, y los requisitos de tamaño de la muestra pueden variar entre los estudios (Engels et al., 2000) (Terrin et al., 2003).

Table 1 Testing for heterogeneity and bias in meta-analyses: 10 scenarios of getting it wrong

Statistical inference	Clinical/pragmatic inference	Specific attribution of cause for heterogeneity or bias
False-positive	No further inferences made	Not done
False-positive	False conclusion that clinical/pragmatic heterogeneity or bias exists	Not done
False-positive	False conclusion that clinical/pragmatic heterogeneity or bias exists	False attribution to specific cause of heterogeneity or bias
False-negative	No further inferences made	Not done
False-negative	False conclusion that there is no overall clinical/pragmatic heterogeneity or bias	Not done
False-negative	False conclusion that there is no overall clinical/pragmatic heterogeneity or bias	False exclusion of specific cause(s) of heterogeneity or bias
True-positive	False conclusion that clinical/pragmatic heterogeneity or bias exists*	Not done
True-positive	False conclusion that clinical/pragmatic heterogeneity or bias exists*	False attribution to specific cause of heterogeneity or bias
True-negative	False conclusion that there is no overall clinical/pragmatic heterogeneity or bias*	Not done
True-negative	False conclusion that there is no overall clinical/pragmatic heterogeneity or bias*	False exclusion of specific cause(s) of heterogeneity or bias

*When statistical heterogeneity/bias is not concordant with clinical/pragmatic heterogeneity/bias.

952 © 2008 The Author. Journal compilation © 2008 Blackwell Publishing Ltd

Figura 10. Escenarios para evaluar heterogeneidad y sesgo de publicación. Extraído de : Ioannidis, J. (2008). Interpretation of tests of heterogeneity and bias in meta-analysis. *Journal of Evaluation in Clinical Practice*, 14(5), 951–957. <https://doi.org/10.1111/j.1365-2753.2008.00986.x>.

En presencia de heterogeneidad, un metaanálisis de efectos aleatorios pondera los estudios de forma relativamente más equitativa que un análisis de efectos fijos. Por lo tanto, en presencia de efectos de los estudios pequeños, según los estudios de estos autores, en la cual el efecto de la intervención es más beneficioso en los estudios más pequeños, la estimación de efectos aleatorios del efecto de la intervención será más beneficiosa que la de efectos fijos (Higgins & Green, 2011) .

El problema de la heterogeneidad es que las ponderaciones en el modelo de efectos aleatorios están influidas por la variabilidad del estimador (en el que influye el tamaño muestral) y la variabilidad entre los estudios. Si esta variabilidad es grande, entonces predomina sobre la variabilidad del estimador lo que produce que todos los estudios tengan ponderaciones muy parecidas y por lo tanto el efecto del tamaño muestral no se muestre tanto.

Como ejemplo a lo anterior se encuentran las estimaciones del efecto del magnesio intravenoso sobre la mortalidad después del infarto de miocardio. En el análisis de efectos fijos el ensayo ISIS-4 obtiene el 90% del peso sin evidencia de un efecto de intervención beneficiosa. Pero en el análisis de efectos aleatorios que dominan los pequeños estudios, y con pruebas claras de efectos beneficiosos de la intervención. Se hace un juicio sobre la probable validez de la evidencia combinada de los estudios más pequeños, en concreto en este ensayo (Higgins & Green, 2011).

Se recomienda que cuando los autores de la revisión estén preocupados por la influencia de los efectos de los resultados de un metaanálisis en el que hay evidencia de heterogeneidad entre estudios ($I^2 > 0$), comparar las estimaciones de efectos fijos y aleatorios del efecto de la intervención (Higgins & Green, 2011).

Se podría pensar que eliminando los estudios no significativos, en algunos casos, es suficiente para eliminar el sesgo de publicación. Esto no es así de fácil, pues en estudios con presencia de heterogeneidad puede surgir la necesidad de utilizar un modelo de sesgo de publicación capaz de analizar conjuntamente los diferentes niveles característicos de los estudios (Coburn & Vevea, 2015).

En relación a cómo utilizar los estudios seleccionados para la realización del metaanálisis, se han estudiado como afecta la heterogeneidad a los métodos de detección del sesgo de publicación. El tamaño de la distribución metaanalítica y el grado de heterogeneidad son limitaciones potenciales del resultado final del metaanálisis. Observando como puede influir este tamaño de la distribución y el grado de heterogeneidad, se concluye que el método de detección de sesgo de publicación Trim and Fill puede dar resultados erróneos en presencia de heterogeneidad, dentro de unas condiciones concretas simuladas (Terrin et al., 2003).

En métodos de regresión donde se puede evaluar la heterogeneidad de los estudios, no solo se investiga la presencia de heterogeneidad y su relación con el sesgo de publicación sino que se ahonda en la investigación hasta descubrir la relación entre heterogeneidad residual y sesgo de publicación (Thompson & Sharp, 1999).

La influencia de la heterogeneidad debe ser reconocida como una limitación tanto para el metaanálisis básico como para el análisis de sesgo de publicación. Con niveles "altos" de heterogeneidad, la interpretación de los resultados puede indicar que los resultados de métodos de detección de sesgo de publicación del tipo "modelo de selección" y "metaanálisis acumulado" son los indicados para esta situación, al ser estos dos métodos más robustos frente a la heterogeneidad (Terrin et al., 2003; Vevea & Woods, 2005).

En general, es importante utilizar múltiples métodos para evaluar el sesgo de publicación y hacerlo a través de los niveles característicos del estudio sospechoso. Así permitiría a los metaanalistas investigar las diferencias significativas sospechosas y evaluar si estas diferencias son un artefacto del sesgo de publicación (Coburn & Vevea, 2015).

Como ejemplo de cómo puede afectar la heterogeneidad al sesgo de publicación se puede leer un estudio donde una serie de investigadores médicos participaron en una encuesta para evaluar los diferentes puntos de vista frente a la heterogeneidad y sesgo de publicación sobre diferentes Funnel Plots. Un 50% de las respuestas fueron acertadas, de promedio, sobre los gráficos presentados. Este resultado fue así ya que algunos de los gráficos presentados eran inusuales en el mundo real, recibiendo mejor resultado los gráficos con representación más habitual. Al influir el sesgo de publicación y la heterogeneidad en la forma del gráfico, los datos se calcularon en cuatro categorías, preguntando por el nivel de asimetría, en un número igual de gráficos con y sin sesgo de publicación.

Aunque no es sorprendente que los participantes no pudieran distinguir entre gráficos de forma similar con y sin sesgo de publicación, los resultados son importantes porque muchos autores continúan evaluando el sesgo de publicación con el diagrama de embudo sin tener en cuenta como influye la heterogeneidad en el sesgo de publicación. La relación heterogeneidad y presencia de sesgo de publicación es evidente, hay que evaluar su impacto de alguna manera y determinar en que niveles de heterogeneidad no aparecería el sesgo de publicación para conseguir un resultado robusto en el metaanálisis (Terrin et al., 2005).

La heterogeneidad influye en los métodos de detección de sesgo de publicación debido a la baja calidad que pueden tener algunos de los estudios con resultados beneficiosamente exagerados. Ocurre con estudios pequeños donde los pacientes son de alto riesgo, para los cuales la intervención puede ser más beneficiosa. No tienen capacidad de reclutar a muchos sujetos y con características similares. En cualquier caso, se produciría una relación inversa entre el beneficio del tratamiento y la precisión, incluso cuando no hay sesgo de publicación. Por este motivo, los modelos de selección que incorporan co-variables podrían ajustarse mejor a criterios de calidad y al riesgo estándar (Terrin et al., 2003)

Los distintos autores interesados en la relación existente entre sesgo de publicación y heterogeneidad han ido evaluando diferentes escenarios donde encontrar esta clara relación. Hay un autor que evalúa el grado de sesgo de publicación en cuatro conjuntos de

datos metaanalíticos previamente publicados de diversos campos de estudio en las ciencias de la organización. De este conjunto de datos, uno no se ve afectado por el sesgo de publicación, mientras que los otros están notablemente influenciados por este sesgo (Kepes & Banks, 2012).

De esos estudios analizados, encuentran que en Hermelin et al. (2007) el sesgo de publicación afectó la estimación metaanalítica; en Quiñones et al. (1995) el sesgo de publicación afectó la distribución de los tamaños del efecto entre la experiencia y las evaluaciones subjetivas del desempeño, con un efecto de este sesgo incierto; en Eagly et al. (2003) la estimación metaanalítica no es robusta al sesgo de publicación y por último en McNatt (2000) la estimación metaanalítica de la muestra civil (primer ajuste) podría verse moderadamente afectada por el sesgo de publicación y la distribución desde los escenarios militares (segundo ajuste) era bastante robusta a la influencia del sesgo de publicación.

Los grados de heterogeneidad existentes en la distribución de los estudios no eran pequeños, excepto para Hermelin et al. (2007). Sus autores creen que si la distribución metaanalítica es lo suficientemente grande para realizar un metaanálisis básico, debe ser lo suficientemente grande como para realizar análisis de sesgo de publicación.

Aunque sea evidente que la heterogeneidad influye en el sesgo de publicación como se ha visto en el ejemplo anterior, otros autores creen que no deberían ser invalidados los resultados de una forma precipitada. Según Kepes & Bank (2012) los niveles potencialmente altos de heterogeneidad no debería invalidar automáticamente los resultados de los análisis de sensibilidad, como los análisis de sesgo de publicación (Kepes & Banks, 2012). Estos autores concluyen que los estudios de simulación pueden ser un forma primaria de evaluar la influencia de la heterogeneidad, incluyendo diferentes tipos de heterogeneidad, y factores relacionados (por ejemplo, tamaño de la distribución metaanalítica) sobre la detección del sesgo de publicación y los métodos de evaluación.

La importancia de investigar la heterogeneidad dentro de cualquier metaanálisis es primordial, no solo cómo ajustarla potencialmente, sino considerar las causas subyacentes de la variación entre estudios. Heterogeneidad y sesgo de publicación no son mutuamente excluyentes y el autor nos encamina al análisis de aquella heterogeneidad residual anteriormente explicada que puede alterar el sesgo de publicación (Malley, 2009).

Hay que hacer una evaluación crítica de los posibles enfoques para abordar el sesgo de publicación en la presencia de heterogeneidad explicable y residual entre estudios, siendo estudiadas ambas limitaciones conjuntamente (Peters et al., 2010).

9. Estudio de simulación para el análisis de la eficiencia de los métodos de detección del sesgo de publicación sin y con heterogeneidad.

Se muestran los resultados de los procesos de simulación para ver como se comportan los métodos más utilizados en la búsqueda del sesgo de publicación. Se parte de tres supuestos: modelo de efectos fijos, modelo de efectos fijos con un factor de heterogeneidad y modelo de efectos aleatorios.

En cada uno de estos supuesto se han determinado diferentes escenarios en relación a la magnitud del efecto, el número de artículos integrados y la magnitud del sesgo de publicación.

- **Modelo de efectos fijos.** Se generan 18 metaanálisis a partir de los siguientes supuestos: número de estudios integrados (10,100), tamaño del efecto poblacional que se pretende estimar (0,25;0,4;0,8) y diferente impacto del sesgo de publicación. La simulación de existencia de sesgo de publicación se realiza eliminando los estudios no significativos. Se identificara como 0%, 50% y 100% la presencia de sesgo de publicación.

a) **Presencia de sesgo de publicación (0%)**

En los análisis gráficos, **Figura 11**, existe una asimetría con un espacio en el lado inferior del Funnel Plot de metaanálisis de 10 estudios donde la muestra tiene un tamaño del efecto pequeño y medio. Por lo tanto, existe un posible sesgo de publicación sin realmente estar forzado en esta simulación.

En esa misma, **Figura 11**, el Funnel Plot de metaanálisis de 100 estudios no presenta asimetría detectando la ausencia real de sesgo de publicación.

El Método Trim And Fill, **Figura 11**, de metaanálisis de 10 estudios y tamaño del efecto medio se evidencia sesgo de publicación sin estar realmente forzado.

El Método Trim And Fill, **Figura 11**, de metaanálisis de 100 estudios detecta la ausencia real de sesgo de publicación.

El Funnel Plot y el Método Trim And Fill detectan erróneamente la existencia de sesgo de publicación en metaanálisis pequeños y con tamaño del efecto pequeño. Pues realmente este último está basado en los grados de asimetría del primero.

El Método Contour-enhanced Funnel Plot, **Figura 14**, en el metaanálisis con 10 estudios detecta una asimetría en los que tienen tamaño del efecto pequeño en relación a los estudios significativos frente a los que lo son menos.

El Método Contour-enhanced Funnel Plot, **Figura 14**, en el metaanálisis con 100 estudios detecta la ausencia de sesgo de publicación en tamaño del efecto pequeño y medio. Especial cuidado con tamaños del efecto grande, pues existe una asimetría entre estudios pero no hay presencia de sesgo de publicación. Esto es debido a que cuanto mayor es el tamaño del efecto mayor es el número de estudios significativos.

Los métodos estadísticos para la detección de sesgo de publicación, **Figura 15**, se comportan bien excepto Macaskill en metaanálisis de 100 estudios donde encuentra sesgo de publicación sin existir realmente. La disminución de la potencia de este método es dado porque la probabilidad de que un estudio sea censurado también disminuye a medida que el tratamiento o el tamaño muestral se hace más grande (Macaskill et al., 2001).

Fail-safe N con metaanálisis de 10 estudios con tamaño del efecto pequeño y medio detecta la presencia del sesgo de publicación sin realmente existir. Con tamaño del efecto grande si detecta la ausencia del sesgo de publicación.

En el Método Fail-safe N con metaanálisis de 100 estudios detecta la ausencia del sesgo de publicación.

b) Presencia de sesgo de publicación (50%)

En los análisis gráficos, **Figura 12**, existe una asimetría con un espacio en el lado inferior derecho del Funnel Plot de metaanálisis de 10 estudios donde la muestra tiene un tamaño del efecto pequeño y medio. Por lo tanto, detecta el sesgo de publicación forzado en esta simulación.

En los análisis gráficos, **Figura 12**, existe una asimetría con un espacio en el lado inferior derecho del Funnel Plot de metaanálisis de 100 estudios donde la muestra tiene un tamaño del efecto medio. Por lo tanto, detecta el sesgo de publicación forzado en esta simulación.

El Método Trim And Fill, **Figura 12**, de metaanálisis de 10 estudios detecta el sesgo de publicación pero sugiriendo menos estudios de los que realmente se han retirado.

El Método Trim And Fill, **Figura 12**, de metaanálisis de 100 estudios detecta el sesgo de publicación pero sugiriendo menos estudios de los que realmente se han retirado.

A medida que disminuyen el número de estudios integrados, se hace más compleja la detección del sesgo de publicación en aquellos metaanálisis de pocos estudios. Como en el escenario de sesgo de publicación 0%, Funnel Plot y Trim And Fill están relacionados detectando este último el sesgo en función del Funnel Plot. De hecho, es difícil detectar la asimetría en el Funnel Plot de metaanálisis de 10 estudios con tamaño del efecto grande y Trim And Fill tampoco lo detecta.

El Método Contour-enhanced Funnel Plot, **Figura 14**, en el metaanálisis con 10 estudios detecta la asimetría en los que tienen tamaño del efecto pequeño y grande. En el supuesto de tamaño del efecto pequeño, la asimetría es por existencia de más estudios de menor significancia frente a los de mayor. Esto reduce la posibilidad de que el sesgo de publicación sea la causa subyacente de esta asimetría del embudo (Higgins & Green, 2011).

El Método Contour-enhanced Funnel Plot, **Figura 14**, en el metaanálisis con 100 estudios detecta asimetría bien en tamaño del efecto grande.

Los métodos estadísticos, **Figura 16**, a medida que forzamos la presencia del sesgo de publicación, se acercan a la detección pero sin confirmarla. El Método de correlación de rango no detecta el sesgo de publicación en ninguno de los casos.

En el Método de Egger, **Figura 16**, con metaanálisis de 10 estudios con tamaño del efecto medio detecta el sesgo de publicación, pero no con tamaño del efecto pequeño y grande. Con metaanálisis de 100 estudios no detecta sesgo de publicación.

En el Método de Macaskill en metaanálisis de 10 estudios no detecta el sesgo de publicación en ninguno de ellos, pero en el metaanálisis de 100 estudios encuentra realmente el sesgo de publicación forzado.

En el Método Fail-safe N con metaanálisis de 10 estudios con tamaño del efecto pequeño y medio detecta la presencia del sesgo de publicación, con estudios necesarios por debajo del nivel de tolerancia pero lejos de lo real, sobre todo en el caso de tamaño del efecto medio. Con tamaño del efecto grande no detecta el sesgo de publicación a pesar de existir.

En el Método Fail-safe N con metaanálisis de 100 estudios no detecta el sesgo de publicación a pesar de existir.

c) Presencia de sesgo de publicación (100%)

En los análisis gráficos, **Figura 14**, el Funnel Plot en metaanálisis de 10 estudios es difícil encontrar una asimetría con ese espacio sin ocupar al quedarnos con pocos estudios significativos. De hecho, solo podemos analizar metaanálisis de tamaño del efecto medio y grande.

En los análisis gráficos, **Figura 14**, el Funnel Plot en los metaanálisis de 100 estudios detecta la presencia de sesgo de publicación presentando una asimetría en el lado inferior izquierdo.

El Método Trim And Fill, **Figura 13**, de metaanálisis de 10 estudios solo en tamaño del efecto medio detecta el sesgo de publicación.

El Método Trim And Fill, **Figura 13**, de metaanálisis de 100 estudios detecta el sesgo de publicación pero sugiriendo menos estudios de los que realmente se han retirado.

El Método Contour-enhanced Funnel Plot, **Figura 13**, en el metaanálisis con 10 estudios detecta la asimetría en los que tienen tamaño del efecto medio y grande, en relación a los estudios significativos frente a los que lo son menos.

El Método Contour-enhanced Funnel Plot, **Figura 13**, en el metaanálisis con 100 estudios detecta bien la presencia de sesgo de publicación.

Los métodos estadísticos, **Figura 17**, el Método de correlación de rango solo en metaanálisis de 100 estudios detecta la presencia del sesgo de publicación con resultados significativos siendo mejores en tamaños del efecto pequeño.

En el Método de Egger, **Figura 17**, solo en metaanálisis de 100 estudios detecta la presencia del sesgo de publicación.

En el Método de Macaskill no funciona bien en metaanálisis de 10 estudios, pero si detecta la presencia del sesgo en metaanálisis de 100 estudios.

El Fail-safe N con metaanálisis de 10 estudios con tamaño del efecto medio detecta la presencia de sesgo de publicación. Con un metaanálisis de 10 estudios de tamaño del efecto grande y metaanálisis de 100 estudios se supera el nivel de tolerancia, no detectando la presencia de sesgo publicación cuando realmente si esta forzada.

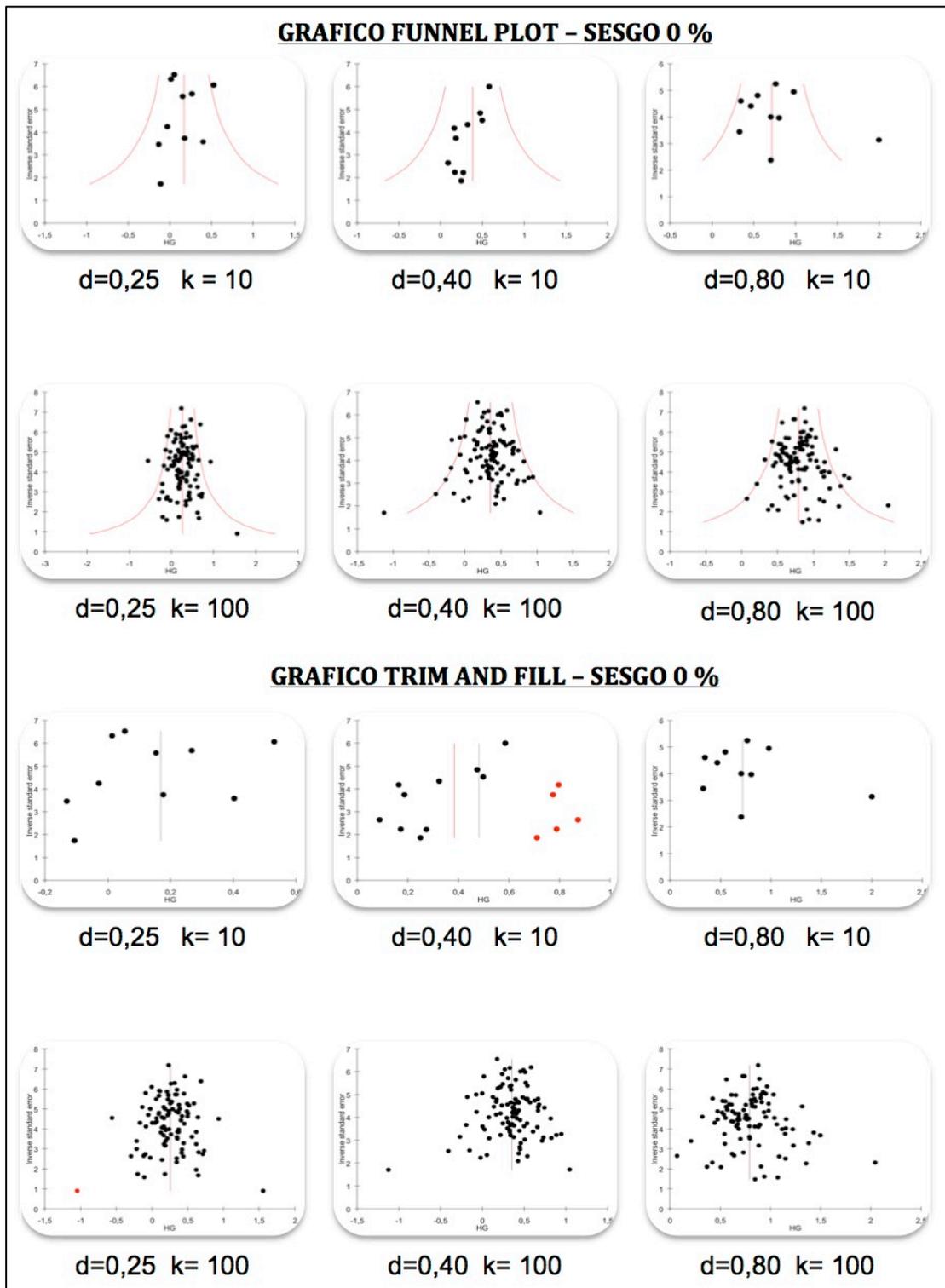


Figura 11. Funnel Plot y Trim And Fill bajo modelo efectos fijos, con cálculo del tamaño del efecto de Hedges, sin variabilidad entre estudios, mediante estimación del inverso de la varianza y con sesgo de publicación (0%). Nota. Tamaño del efecto = "d" y número de estudios de cada metaanálisis = "k".

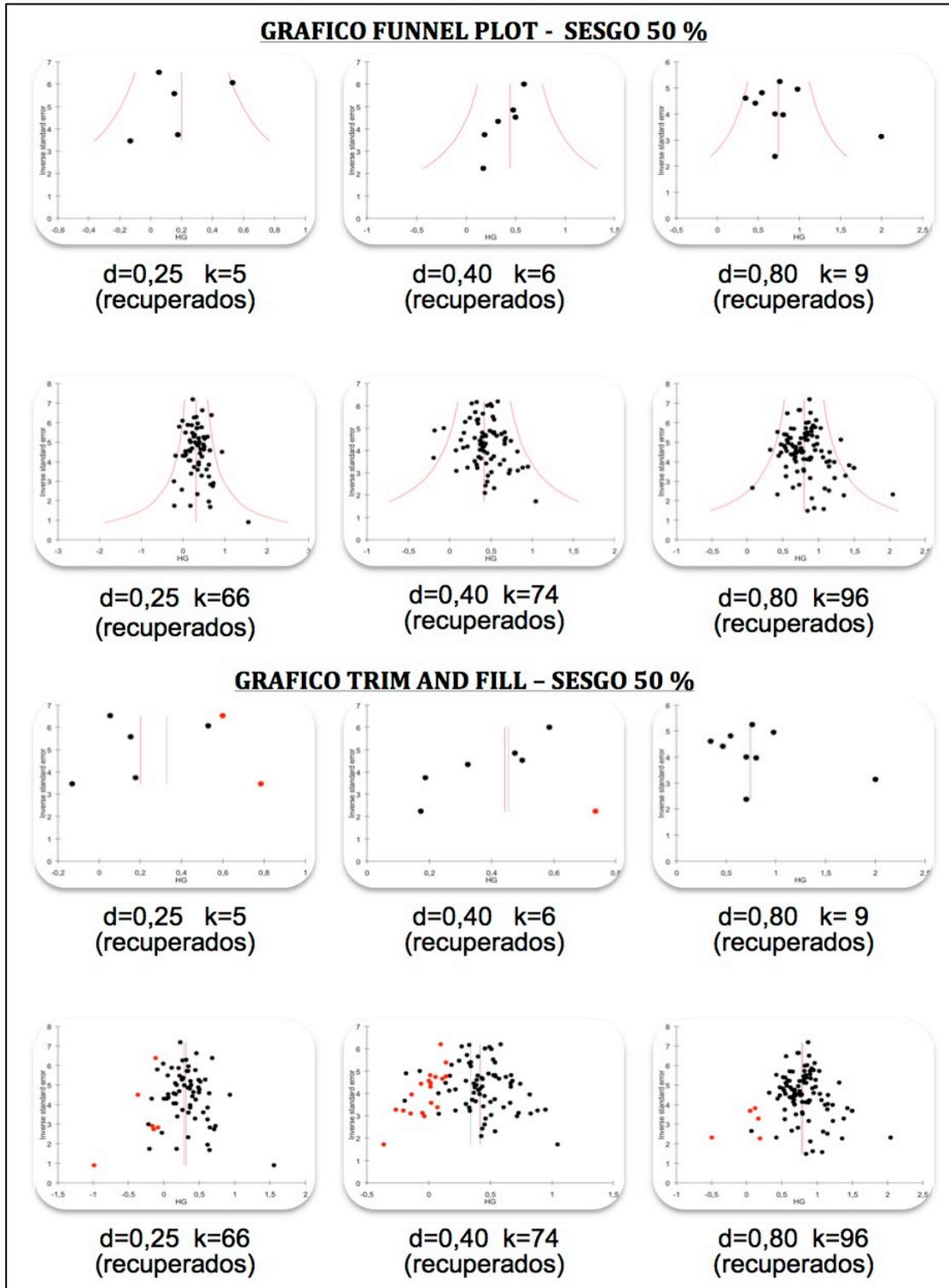
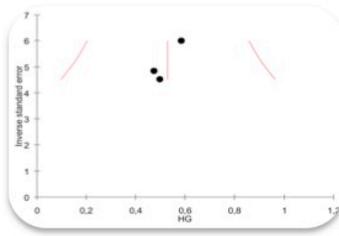


Figura 12. Funnel Plot y Trim And Fill bajo modelo efectos fijos, con cálculo del tamaño del efecto de Hedges, sin variabilidad entre estudios, mediante estimación del inverso de la varianza y con sesgo de publicación (50%). Nota. Tamaño del efecto = “d” y número de estudios de cada metaanálisis = “k”.

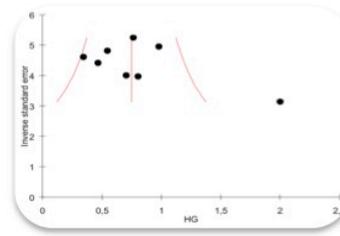
GRAFICO FUNNEL PLOT - SESGO 100 %



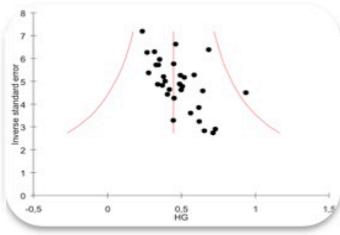
d=0,25 k= (n° significativos < 2)



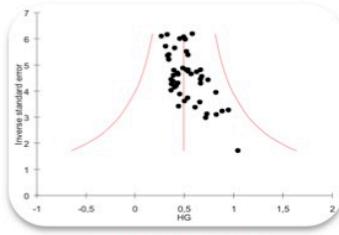
d=0,40 k=3 (recuperados)



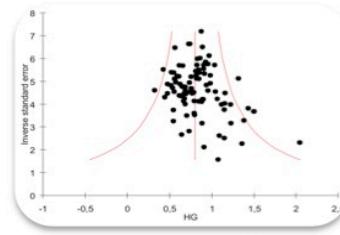
d=0,80 k=8 (recuperados)



d=0,25 k=32 (recuperados)



d=0,40 k=48 (recuperados)

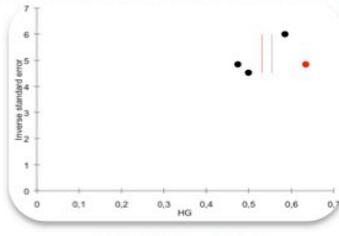


d=0,80 k=92 (recuperados)

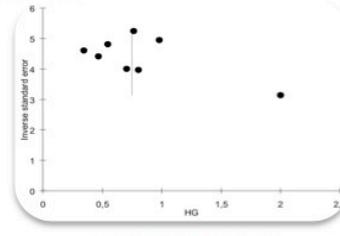
GRAFICO TRIM AND FILL - SESGO 100 %



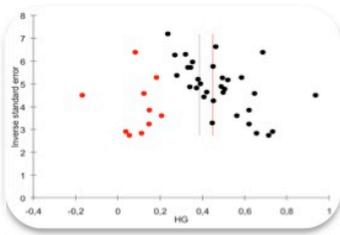
d=0,25 k= (n° significativos < 2)



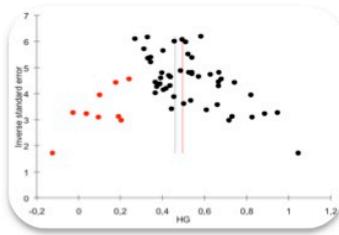
d=0,40 k=3 (recuperados)



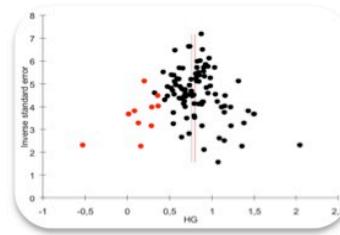
d=0,80 k=8 (recuperados)



d=0,25 k=32 (recuperados)



d=0,40 k=48 (recuperados)



d=0,80 k=92 (recuperados)

Figura 13. Funnel Plot y Trim And Fill bajo modelo efectos fijos, con cálculo del tamaño del efecto de Hedges, sin variabilidad entre estudios, mediante estimación del inverso de la varianza y con sesgo de publicación (100%). Nota. Tamaño del efecto = "d" y número de estudios de cada metaanálisis = "k".

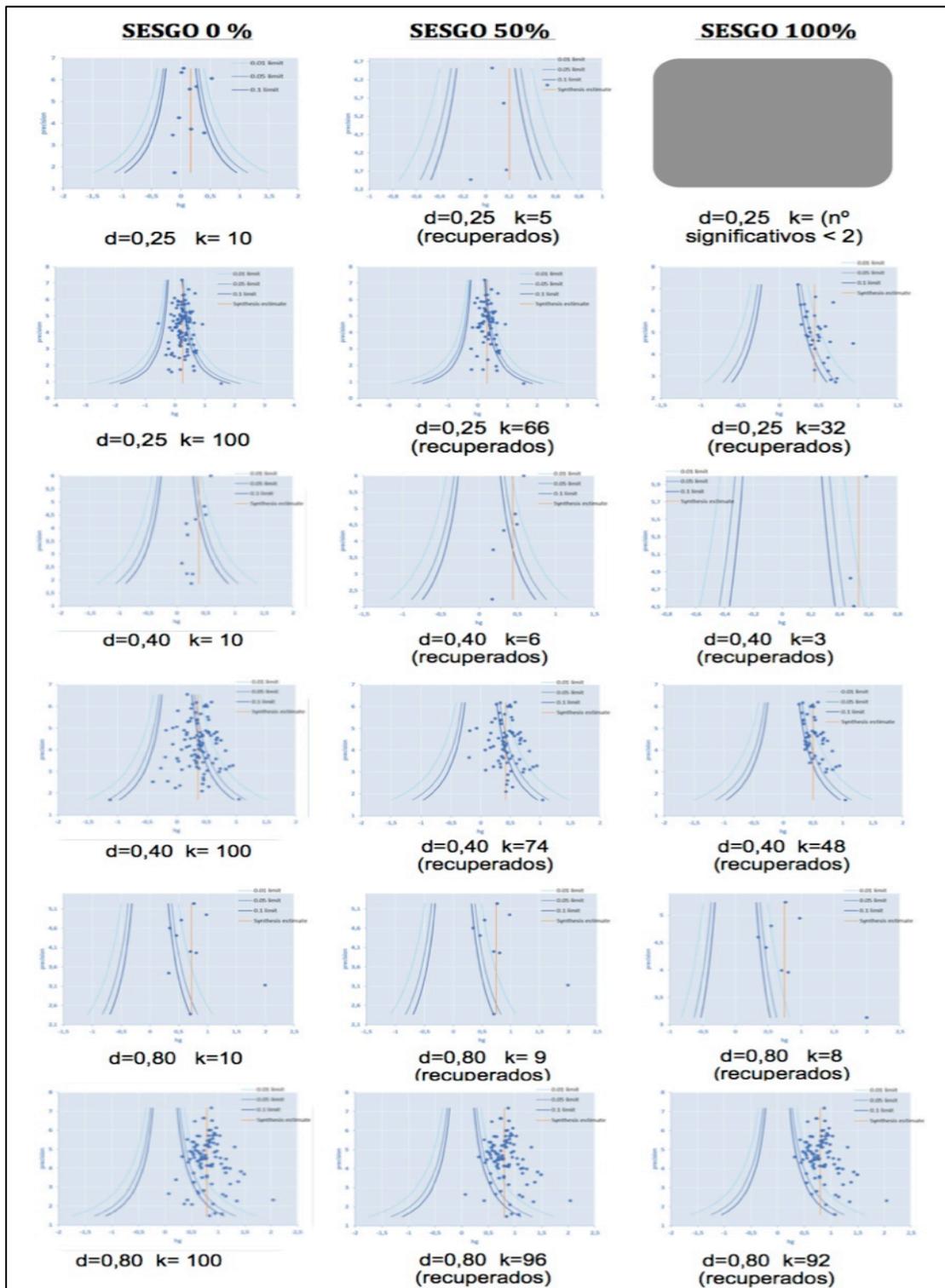


Figura 14. Contour-enhanced Funnel Plot bajo modelo efectos fijos, con cálculo del tamaño del efecto de Hedges, sin variabilidad entre estudios, mediante estimación del inverso de la varianza y diferentes grados de sesgo de publicación. Nota. Tamaño del efecto = "d" y número de estudios de cada metaanálisis = "k".

General	d-k		d-k		d-k		d-k		d-k	
	0,25-10	0,40-10	0,80-10	0,25-100	0,40-100	0,80-100	0,25-100	0,40-100	0,80-100	
Resultado Original Meta análisis	0,1696	0,3833	0,7186	0,258	0,3544	0,7927	0,258	0,3544	0,7927	
95% CI inferior	0,0436	0,2233	0,5701	0,2137	0,3088	0,7498	0,2137	0,3088	0,7498	
95% CI superior	0,2956	0,5433	0,867	0,3024	0,4	0,8356	0,3024	0,4	0,8356	
Evaluación del efecto										
Correlación rango ajustado										
Tau-Kendall (continuity corrected)	0	-0,2667	0,0444	0,0034	0,0099	0,0723	0,0034	0,0099	0,0723	
Z	0	-1,0733	0,1789	0,0507	0,1465	1,0652	0,0507	0,1465	1,0652	
p-valor (dos colas)	1	0,2831	0,858	0,9596	0,8835	0,2868	0,9596	0,8835	0,2868	
Método de regresión	Egger	Egger	Egger	Egger	Egger	Egger	Egger	Egger	Egger	
Intercepto	-0,6003	-1,3165	1,9952	-0,0134	-0,0781	0,4688	-0,0134	-0,0781	0,4688	
95% CI inferior	-3,1497	-2,5137	-4,0345	-0,7318	-0,8716	-0,3423	-0,7318	-0,8716	-0,3423	
95% CI superior	1,9491	-0,1193	8,0249	0,705	0,7155	1,2799	0,705	0,7155	1,2799	
p-valor (dos colas)	0,6019	0,0349	0,4673	0,9705	0,8456	0,2542	0,9705	0,8456	0,2542	
Método de regresión	Macaskill	Macaskill	Macaskill	Macaskill	Macaskill	Macaskill	Macaskill	Macaskill	Macaskill	
Pendiente	0,0013	0,0031	-0,0009	-0,0002	0,0006	-0,0003	-0,0002	0,0006	-0,0003	
95% CI inferior	-0,0017	0,0012	-0,0147	-0,0015	-0,001	-0,0017	-0,0015	-0,001	-0,0017	
95% CI superior	0,0043	0,005	0,0128	0,0012	0,0022	0,001	0,0012	0,0022	0,001	
p-valor (dos colas)	0,3389(0,7178)	0,0055(0,0016)	0,8809(0,9756)	< 0,0001*	< 0,0001*	< 0,0001*	< 0,0001*	< 0,0001*	< 0,0001*	
Análisis de sensibilidad										
Fail-safe N	5	31	228	3007	5524	32220	3007	5524	32220	
Nivel de Tolerancia	60	60	60	510	510	510	510	510	510	
Método Trim And Fill (automático)	L0	L0	L0	L0	L0	L0	L0	L0	L0	
Número de estudios	0	5	0	1	0	0	1	0	0	
Resultado Meta análisis	0,1696	0,4811	0,7186	0,2575	0,3544	0,7927	0,2575	0,3544	0,7927	
95% CI inferior	0,0436	0,3414	0,5701	0,2132	0,3088	0,7498	0,2132	0,3088	0,7498	
95% CI superior	0,2956	0,6208	0,867	0,3018	0,4	0,8356	0,3018	0,4	0,8356	

Figura 15. Datos estadísticos bajo modelo efectos fijos, con cálculo del tamaño del efecto de Hedges, sin variabilidad entre estudios, mediante estimación del inverso de la varianza y con sesgo de publicación (0%). Nota. Tamaño del efecto = "d" y número de estudios de cada metaanálisis = "k". Método Macaskill entre paréntesis valor FIV (ponderada inversa de la varianza). (*) No hay cambios en relación al método no ponderado.

General	d-k(recuperados)						d-k(recuperados) 0,80-10(9)	d-k(recuperados) 0,25-100(66)	d-k(recuperados) 0,40-100(74)	d-k(recuperados) 0,80-100(96)
	0,25-10(6)	0,40-10(6)	0,80-10(9)	0,25-100(66)	0,40-100(74)	0,80-100(96)				
Resultado Original Meta análisis	0,2023	0,4418	0,7467	0,3157	0,4171	0,7981				
95% CI inferior	0,0343	0,2608	0,593	0,2637	0,3649	0,7549				
95% CI superior	0,3704	0,6228	0,9005	0,3677	0,4692	0,8413				
Evaluación del efecto										
Correlación rango ajustado										
Tau-Kendall (continuity corrected)	-0,1	-0,6667	0,1389	0,0876	0,1037	0,1174				
Z	-0,2449	-1,8787	0,5213	1,0404	1,3076	1,6951				
p-valor (dos colas)	0,8065	0,0603	0,6022	0,2981	0,191	0,0901				
Método de regresión	Egger	Egger	Egger	Egger	Egger	Egger				
Intercepto	-1,6837	-1,9597	2,9421	0,3845	0,6581	0,7972				
95% CI inferior	-9,7747	-3,9555	-3,4566	-0,4835	-0,2266	-0,0595				
95% CI superior	6,4073	0,0361	9,3408	1,2524	1,5427	1,6539				
p-valor (dos colas)	0,5552	0,0526	0,3129	0,3795	0,1425	0,0678				
Método de regresión	Macaskill	Macaskill	Macaskill	Macaskill	Macaskill	Macaskill				
Pendiente	0,0021	0,0036	-0,0031	-0,001	-0,0011	-0,0009				
95% CI inferior	-0,0048	0,0012	-0,0181	-0,0025	-0,0026	-0,0023				
95% CI superior	0,0091	0,006	0,0118	0,0006	0,0004	0,0004				
p-valor (dos colas)	0,4032(0,6199)	0,0136(0,0187)	0,6368(0,7526)	<0,0001*	<0,0001*	<0,0001*				
Análisis de sensibilidad										
Fail-safe N	1	23	211	2274	4585	31448				
Nivel de Tolerancia	35	40	55	340	380	490				
Método Trim And Fill (automático)	L0	L0	L0	L0	L0	L0				
Número de estudios	2	1	0	6	19	6				
Resultado Meta análisis	0,3279	0,4537	0,7467	0,2875	0,3417	0,7761				
95% CI inferior	0,1859	0,2764	0,593	0,237	0,2947	0,7335				
95% CI superior	0,47	0,631	0,9005	0,3379	0,3887	0,8186				

Figura 16. Datos estadísticos bajo modelo efectos fijos, con cálculo del tamaño del efecto de Hedges, sin variabilidad entre estudios, mediante estimación del inverso de la varianza y con sesgo de publicación (50%). Nota. Tamaño del efecto = "d" y número de estudios de cada metaanálisis = "k". Método Macaskill entre paréntesis valor FIV (ponderado inversa de la varianza).

(*) No hay cambios en relación al método no ponderado.

- **Modelo efectos fijos con un solo factor.** Se generan 9 metaanálisis a partir de los siguientes supuestos: número de estudios integrados de 100, tamaño del efecto poblacional que se pretende estimar (0,25;0,4;0,8) y donde la magnitud del efecto cambia en aproximadamente un tercio de los estudios en función de un factor de tres categorías que se reparte -0,2, 0 y 0,2. La simulación de existencia de sesgo de publicación se realiza eliminando los estudios no significativos. Se identificara como 0%, 50% y 100% la presencia de sesgo de publicación.

a) Presencia de sesgo de publicación (0%)

En los análisis gráficos Funnel Plot, **Figura 18**, no se identifica la presencia de sesgo de publicación.

El Método Trim And Fill si evidencia la presencia de sesgo de publicación en la muestra con tamaño del efecto pequeño sin estar forzado.

El Método Contour-enhanced Funnel Plot, **Figura 19**, presenta asimetría en tamaño del efecto pequeño hacia menos estudios significativos frente a los de mayor significancia. Algo similar a lo ocurrido en metaanálisis de 10 estudios con tamaño del efecto pequeño y con presencia de sesgo de publicación del 50%. Esto reduce la posibilidad de que el sesgo de publicación sea la causa subyacente de esta asimetría del embudo (Higgins & Green, 2011). Con tamaño del efecto medio se comporta bien, pero con tamaño del efecto grande puede sugerir presencia de sesgo de publicación sin realmente existir, al tener el metaanálisis mayor cantidad de estudios altamente significativos.

Se puede concluir que los métodos tienen una relación directa en el grado de asimetría del Funnel Plot, tanto Método Trim And Fill como Método Contour-enhanced Funnel Plot los cuales fallan en tamaño del efecto grande con mayor número de estudios significativos.

Los métodos estadísticos, **Figura 20**, se comportan bien todos los métodos excepto Macaskill al repetirse un metaanálisis de 100 estudios donde detecta presencia de sesgo de publicación en tamaño del efecto medio y grande.

El Fail-safe N presenta resultados que superan el nivel de tolerancia, detectando la presencia de sesgo publicación que en este supuesto está presente.

b) Presencia de sesgo de publicación (50%)

En los análisis gráficos, **Figura 18**, el Funnel Plot no se identifica la presencia sesgo de publicación.

El Método Trim And Fill evidencia un mínimo sesgo de un artículo perdido en el metaanálisis con tamaño medio del efecto, cuando realmente se han eliminado bastantes más.

El Método Contour-enhanced Funnel Plot, **Figura 19**, detecta la presencia del sesgo de publicación mayormente en tamaño del efecto grande. En tamaño del efecto pequeño y medio no detecta la presencia de sesgo.

Los métodos estadísticos, **Figura 20**, a medida que se fuerza la presencia del sesgo de publicación, los métodos se acercan a la detección del sesgo de publicación pero sin llegar a confirmarla. El Método de correlación de rango no detecta el sesgo de publicación.

En el Método de Egger no detecta el sesgo de publicación.

En el Método de Macaskill presenta resultados similares a la simulación anterior, no detectando el sesgo de publicación en el caso de tamaño del efecto pequeño.

El Fail-safe N presenta resultados que superan el nivel de tolerancia, no detectando la presencia de sesgo publicación que en este supuesto no está presente.

c) Presencia de sesgo de publicación (100%)

En los análisis gráficos, **Figura 18**, el Funnel Plot se identifica la presencia de sesgo de publicación, siendo menor la asimetría según aumenta el tamaño del efecto.

El Método Trim And Fill evidencia sesgo de publicación solo en tamaño del efecto pequeño, sugiriendo un número de artículos recuperados inferior a los que realmente se han eliminado. Nuevamente con un mayor grado de asimetría del Funnel Plot, en este caso de tamaño del efecto pequeño, se condiciona la detección de sesgo de publicación.

El Método Contour-enhanced Funnel Plot, **Figura 19**, presenta asimetría detectando la presencia de sesgo de publicación.

Los métodos estadísticos, **Figura 20**, el Método de correlación de rango solo en metaanálisis de tamaño del efecto pequeño detecta el sesgo de publicación.

En el Método de Egger ocurre exactamente lo mismo que con el Método de correlación de rango.

Al igual que con el Método Trim And Fill, el Método de correlación de rango y el Método de Egger, dependientes del grado de asimetría del Funnel Plot detectan la presencia del sesgo de publicación cuando mayor es la presencia visual de esta asimetría (Kepes & Banks, 2012).

En el Método de Macaskill ocurre algo similar al caso anterior, pero ahora detecta sesgo de publicación en todos los casos de tamaño del efecto.

El Fail-safe N presenta resultados que superan el nivel de tolerancia, no detectando la presencia de sesgo publicación que en este supuesto no está presente. Todos los resultados en sus distintos escenarios son similares, por tanto, no deberían de ser muy valorables en presencia de metaanálisis con número de estudios elevado.

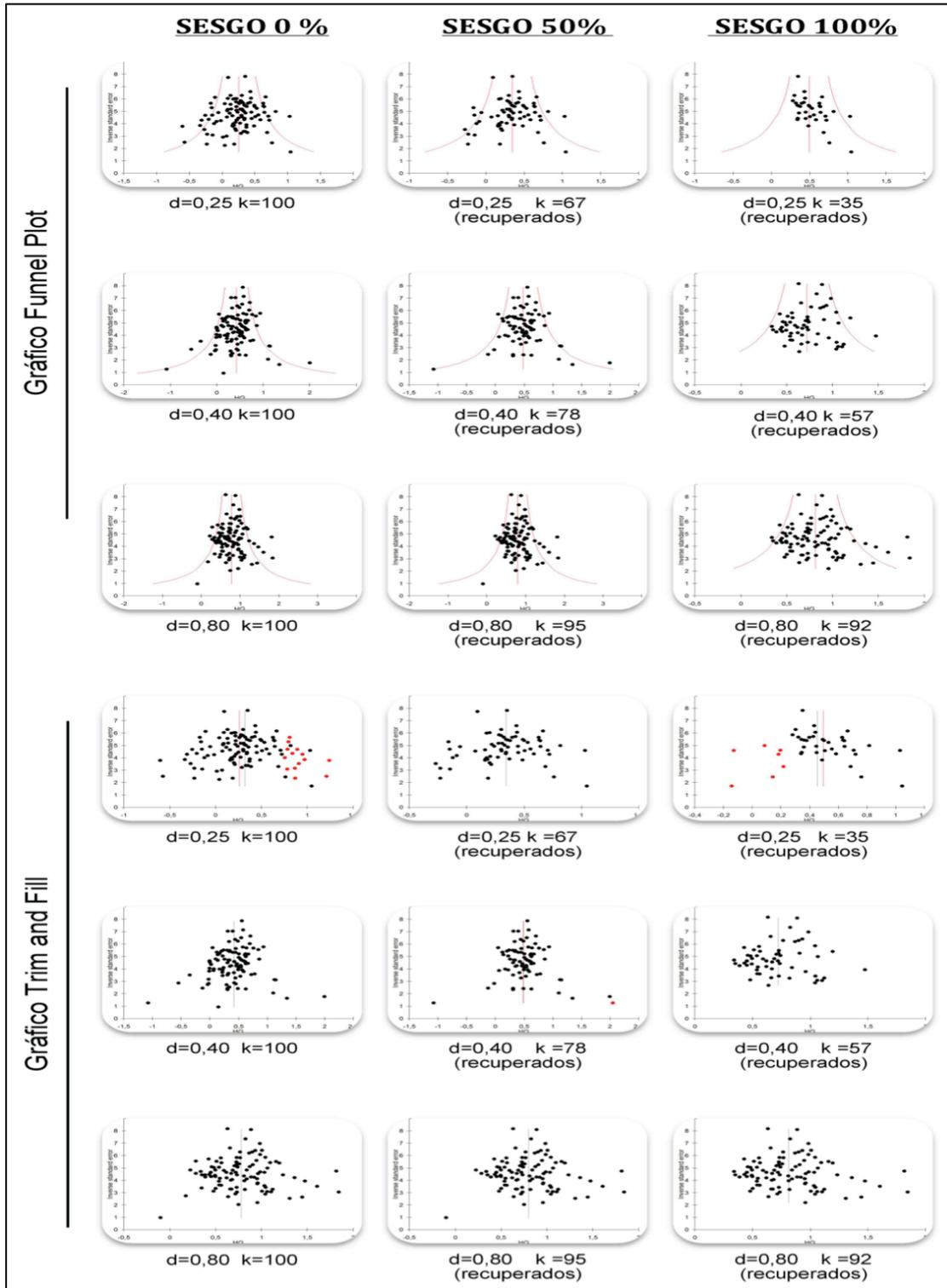


Figura 18. Funnel Plot y Trim And Fill bajo modelo efectos aleatorios, con cálculo del tamaño del efecto de Hedges, con variabilidad entre grupos de estudios con $(\tau^2 = [-0,2; 0; 0,2])$ según pertenencia, mediante estimación del DerSimonian-Laird y diferentes grados de sesgo de publicación. Nota. Tamaño del efecto = "d" y número de estudios de cada metaanálisis = "k".

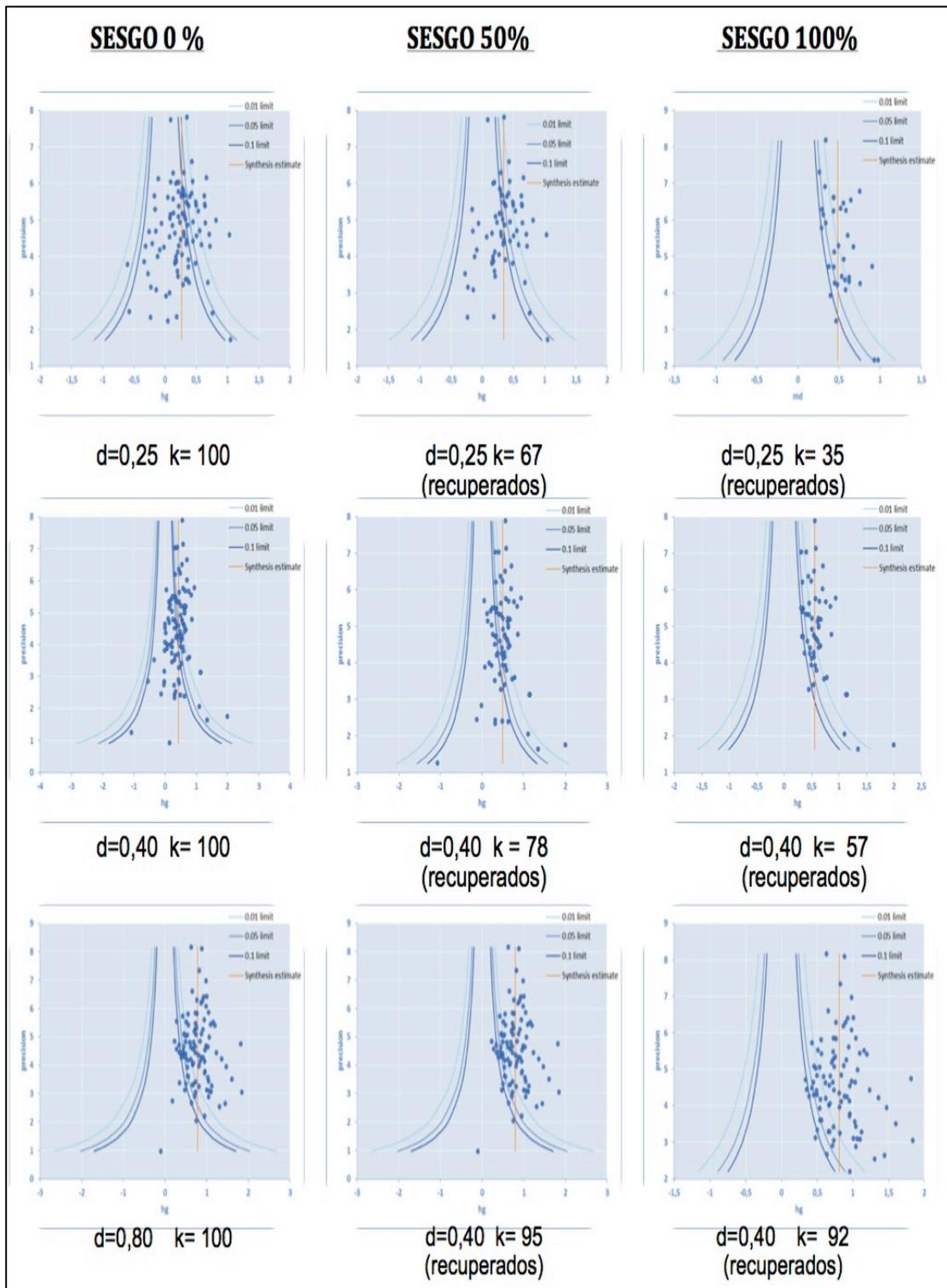


Figura 19. Contour-enhanced Funnel Plot bajo modelo efectos aleatorios, con cálculo del tamaño del efecto de Hedges, con variabilidad entre grupos de estudios con $(\tau^2 = [-0,2; 0; 0,2])$ según pertenencia, mediante estimación del DerSimonian-Laird y diferentes grados de sesgo de publicación. Nota. Tamaño del efecto = “d” y número de estudios de cada metaanálisis = “k”.

General	0% d = 0,25 k = 100		0% d = 0,4 k = 100		0% d = 0,8 k = 100		50% d = 0,25 k = 67 (recuperados)		50% d = 0,4 k = 78 (recuperados)		50% d = 0,8 k = 95 (recuperados)		100% d = 0,25 k = 55 (recuperados)		100% d = 0,4 k = 57 (recuperados)		100% d = 0,8 k = 92 (recuperados)	
	Resultado Original Meta analisis	0,2575	0,4222	0,781	0,3436	0,4977	0,7977	0,4935	0,721	0,8106	0,4291	0,6616	0,7535	0,4935	0,721	0,8106	0,4291	0,6616
95% CI inferior	0,2045	0,369	0,7224	0,2843	0,4365	0,7398	0,2843	0,4365	0,7398	0,2843	0,4365	0,7398	0,2843	0,4365	0,7398	0,2843	0,4365	0,7398
95% CI superior	0,3106	0,4753	0,8395	0,4029	0,539	0,8557	0,4029	0,539	0,8557	0,4029	0,539	0,8557	0,4029	0,539	0,8557	0,4029	0,539	0,8557
Evaluación del efecto																		
Correlación rango ajustado																		
Tau-Kendall (continuity corrected)	-0,103	-0,0984	0,0342	-0,1213	0,0173	0,0542	-0,1213	0,0173	0,0542	-0,1213	0,0173	0,0542	-0,1213	0,0173	0,0542	-0,1213	0,0173	0,0542
z	-1,5181	-1,451	0,5035	-1,4518	0,2245	0,7819	-1,4518	0,2245	0,7819	-1,4518	0,2245	0,7819	-1,4518	0,2245	0,7819	-1,4518	0,2245	0,7819
p-valor (dos colas)	0,129	0,1468	0,6146	0,1466	0,8224	0,4343	0,1466	0,8224	0,4343	0,1466	0,8224	0,4343	0,1466	0,8224	0,4343	0,1466	0,8224	0,4343
Método de regresión	Egger	Egger	Egger	Egger	Egger	Egger	Egger	Egger	Egger	Egger	Egger	Egger	Egger	Egger	Egger	Egger	Egger	Egger
Intercepto	-0,7093	-0,5617	0,124	-0,5037	0,1644	0,269	-0,5037	0,1644	0,269	-0,5037	0,1644	0,269	-0,5037	0,1644	0,269	-0,5037	0,1644	0,269
95% CI inferior	-1,7714	-1,3732	-0,8778	-1,7397	-0,7116	-0,7292	-1,7397	-0,7116	-0,7292	-1,7397	-0,7116	-0,7292	-1,7397	-0,7116	-0,7292	-1,7397	-0,7116	-0,7292
95% CI superior	0,3528	0,2498	1,1257	0,7324	1,0404	1,2672	0,7324	1,0404	1,2672	0,7324	1,0404	1,2672	0,7324	1,0404	1,2672	0,7324	1,0404	1,2672
p-valor (dos colas)	0,1882	0,1727	0,8065	0,4187	0,7096	0,5939	0,4187	0,7096	0,5939	0,4187	0,7096	0,5939	0,4187	0,7096	0,5939	0,4187	0,7096	0,5939
Método de regresión	Macaskill	Macaskill	Macaskill	Macaskill	Macaskill	Macaskill	Macaskill	Macaskill	Macaskill	Macaskill	Macaskill	Macaskill	Macaskill	Macaskill	Macaskill	Macaskill	Macaskill	Macaskill
Pendiente	0,0011	0,001	0,0005	0,0008	-0,0001	0,0003	0,0008	-0,0001	0,0003	-0,0001	0,0003	0,0003	-0,0001	0,0003	0,0003	-0,0001	0,0003	0,0003
95% CI inferior	-0,0002	-0,0005	-0,0008	-0,0007	-0,0017	-0,0009	-0,0007	-0,0017	-0,0009	-0,0017	-0,0009	-0,0017	-0,0007	-0,0017	-0,0009	-0,0007	-0,0017	-0,0009
95% CI superior	0,0025	0,0025	0,0018	0,0024	0,0016	0,0016	0,0024	0,0016	0,0016	0,0016	0,0016	0,0016	0,0016	0,0016	0,0016	0,0016	0,0016	0,0016
p-valor (dos colas)	0,1018(0,2221)	< 0,0001*	< 0,0001*	0,2787(0,6276)	< 0,0001*	< 0,0001*	0,2787(0,6276)	< 0,0001*	< 0,0001*	< 0,0001*	< 0,0001*	< 0,0001*	< 0,0001*	< 0,0001*	< 0,0001*	< 0,0001*	< 0,0001*	< 0,0001*
Análisis de sensibilidad																		
Fail-safe N	3586	8650	32280	2958	7957	31379	2958	7957	31379	2958	7957	31379	2958	7957	31379	2958	7957	31379
Nivel de Tolerancia	510	510	510	345	400	480	345	400	480	345	400	480	345	400	480	345	400	480
Método Trim And Fill (automático)	LO	LO	LO	LO	LO	LO	LO	LO	LO	LO	LO	LO	LO	LO	LO	LO	LO	LO
Número de estudios	14	0	0	0	1	0	0	1	0	1	0	0	1	0	0	1	0	0
Resultado Meta analisis	0,3216	0,4222	0,781	0,3436	0,4894	0,7977	0,3436	0,4894	0,7977	0,3436	0,4894	0,7977	0,3436	0,4894	0,7977	0,3436	0,4894	0,7977
95% CI inferior	0,263	0,369	0,7224	0,2843	0,4373	0,7398	0,2843	0,4373	0,7398	0,2843	0,4373	0,7398	0,2843	0,4373	0,7398	0,2843	0,4373	0,7398
95% CI superior	0,3803	0,4753	0,8395	0,4029	0,5415	0,8557	0,4029	0,5415	0,8557	0,4029	0,5415	0,8557	0,4029	0,5415	0,8557	0,4029	0,5415	0,8557

Figura 20. Datos estadísticos bajo modelo efectos aleatorios, con cálculo del tamaño del efecto de Hedges, con variabilidad entre grupos de estudios con $(\tau^2 = [-0,2; 0; 0,2])$ según pertenencia, mediante estimación del DerSimonian-Laird y diferentes grados de sesgo de publicación.

Nota. Tamaño del efecto = "d" y número de estudios de cada metaanálisis = "k".

Método Macaskill entre paréntesis valor FIV (ponderado inversa de la varianza).

(*) No hay cambios en relación al método no ponderado.

- **Modelo de efectos aleatorios.** Se generan 27 metaanálisis a partir de los siguientes supuestos: número de estudios integrados de 100, tamaño del efecto poblacional que se pretende estimar (0,25;0,4;0,8) y donde la magnitud del efecto cambia (0,10;0,20;0,40) en cada uno de los tamaños del efecto. La simulación de existencia de sesgo de publicación se realiza eliminando los estudios no significativos. Se identificará como 0%, 50% y 100% la presencia de sesgo de publicación.

a) Presencia de sesgo de publicación (0%)

En los análisis gráficos de los Funnel Plots, **Figura 21**, presenta en los distintos tamaños del efecto viéndose incrementada su magnitud, aparece una pequeña asimetría en el lado inferior izquierdo que sugiere presencia de sesgo de publicación, siendo mayor en tamaño del efecto pequeño.

El Método Trim And Fill, **Figura 21**, establece un límite con un tamaño del efecto de 0,40 y magnitud del efecto de 0,40, en donde por debajo detecta la presencia del sesgo de publicación y por encima no.

Se relaciona el grado de asimetría que el Funnel Plot tiene en tamaño del efecto pequeño en sus distintas modalidades de magnitud del efecto con la detección de presencia del sesgo de publicación del Método Trim And Fill.

El Método Contour-enhanced Funnel Plot, **Figura 24 y 25**, detecta la ausencia de sesgo de publicación en tamaños del efecto pequeño y medio en sus distintos supuesto de magnitud del efecto. En la **Figura 25**, en el caso de tamaño del efecto grande con magnitud del efecto 0,10 y 0,20 hay que tener especial cuidado. Existe una asimetría entre estudios pero realmente no hay presencia de sesgo de publicación. Esto es debido a que cuanto mayor es el tamaño del efecto mayor es el número de estudios con valores altos significativamente.

Los métodos estadísticos, **Figura 27-28-29**, en el caso del Método de correlación de rango y el Método de Egger a partir de tamaños del efecto de 0,80 y una magnitud del efecto de 0,20 indica presencia del sesgo de publicación sin estar realmente presente.

El Método Macaskill indica sesgo de publicación a partir de un tamaño del efecto de 0,40 y una magnitud del efecto de 0,10 sin estar realmente presente.

El Fail-safe N presenta resultados que superan el nivel de tolerancia, no detectando la presencia de sesgo publicación que en este supuesto está presente.

b) Presencia de sesgo de publicación (50%)

En los análisis gráficos de los Funnel Plots, **Figura 22**, en el escenario de tamaño del efecto de 0,40 según se incrementa la magnitud del efecto, la asimetría en el espacio inferior cambia de lado yendo del derecho al izquierdo.

El resto de tamaños del efecto, tienen una asimetría dudosa siendo más evidente en tamaño del efecto de 0,80 con magnitud del efecto 0,20 y 0,40.

El Método Trim And Fill, **Figura 22**, solo detecta la presencia del sesgo de publicación con tamaño del efecto 0,40 con magnitud del efecto 0,10 y tamaño del efecto 0,80 con magnitud del efecto 0,10 y 0,20.

El Método Contour-enhanced Funnel Plot, **Figura 24 y 25**, no detecta la presencia de sesgo de publicación en tamaños del efecto pequeño y medio en sus distintos supuestos de magnitud del efecto. En la Figura 25, en el caso de tamaño del efecto grande si detecta la presencia del sesgo de publicación.

Los métodos estadísticos, **Figura 27-28-29**, detecta la presencia del sesgo de publicación con el Método de correlación de rango y el Método de Egger a partir de tamaños del efecto de 0,80 y magnitud del efecto de 0,20.

El Método Macaskill indica presencia del sesgo de publicación a partir de un tamaño del efecto de 0,40 y magnitud del efecto de 0,1.

El Fail-safe N presenta resultados que superan el nivel de tolerancia, no detectando la presencia de sesgo publicación que en este supuesto si está presente.

c) Presencia de sesgo de publicación (100%)

En los análisis gráficos de los Funnel Plots, **Figura 23**, detecta la presencia del sesgo de publicación incrementando la asimetría en el lado inferior izquierdo según aumenta el tamaño del efecto y la magnitud del efecto,

El Método Trim And Fill detecta la presencia del sesgo de publicación forzado en todos los tamaños del efecto salvo con tamaño del efecto 0,25 con magnitud del efecto 0,20 y tamaño del efecto 0,80 con magnitud del efecto 0,20 y 0,40.

El Método Contour-enhanced Funnel Plot, **Figura 24-25-26**, detecta la presencia de sesgo de en sus distintos supuestos de magnitud del efecto.

Los métodos estadísticos, **Figura 27-28-29**, en el caso del Método de correlación de rango y el Método de Egger identifica el sesgo de publicación en todos los casos excepto con tamaño del efecto 0,80 y magnitud del efecto de 0,10. Este supuesto concreto presenta una imagen de Funnel Plot que evidencia un menor grado de asimetría y el número de estudios ausentes es menor que en el resto de magnitudes del efecto para un tamaño del efecto de 0,80.

El Método Macaskill indica la presencia del sesgo de publicación en todos los casos.

El Fail-safe N presenta resultados que superan el nivel de tolerancia, no detectando la presencia de sesgo publicación que en este supuesto si está presente.

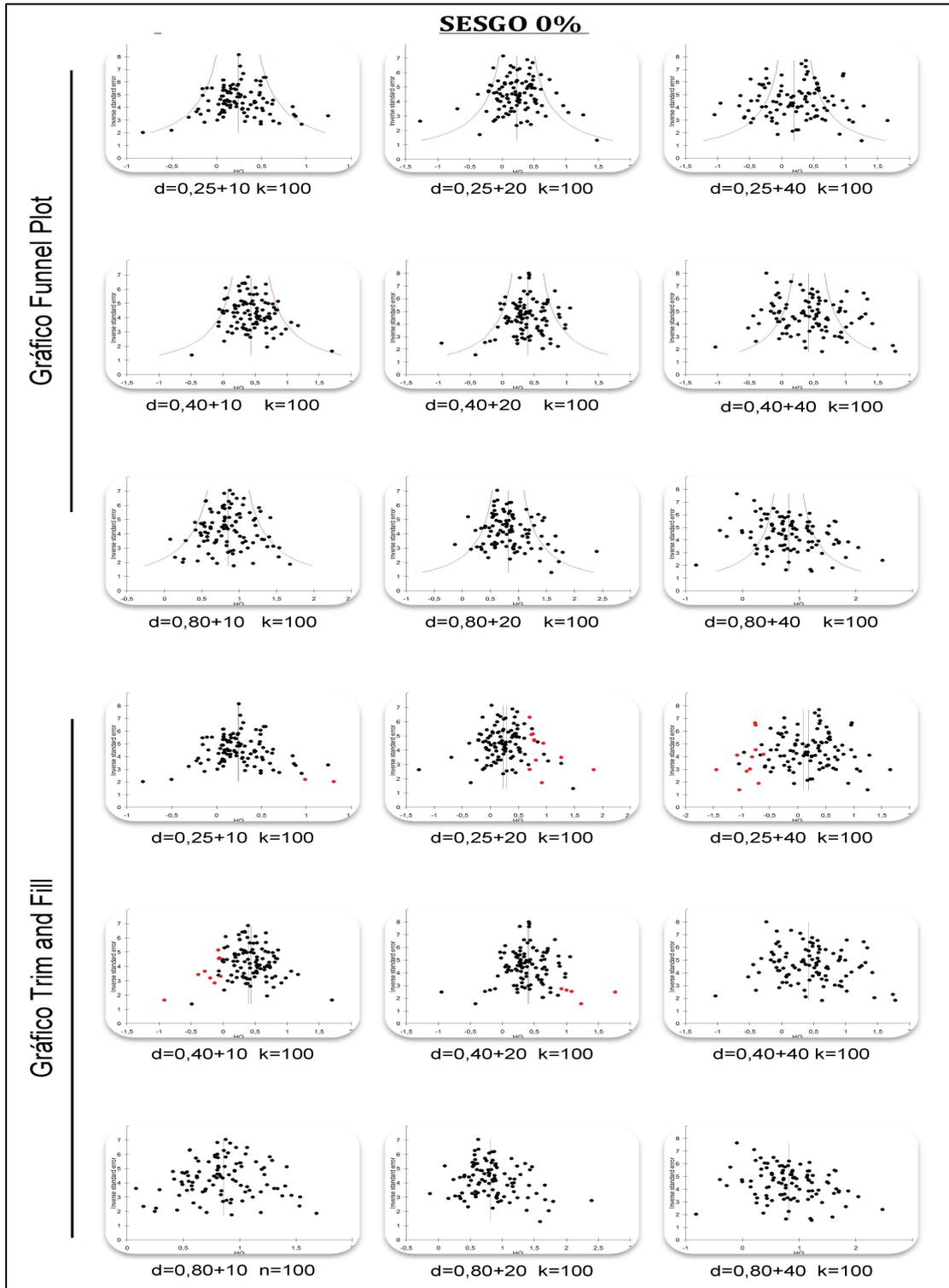


Figura 21. Funnel Plot y Trim And Fill bajo modelo efectos aleatorios, con cálculo del tamaño del efecto de Hedges, con variabilidad entre estudios ($\tau^2 = [0,10; 0,20; 0,40]$), mediante estimación del DerSimonian-Laird y diferentes grados de sesgo de publicación. Nota. Tamaño del efecto = “d” y número de estudios de cada metaanálisis = “k”.

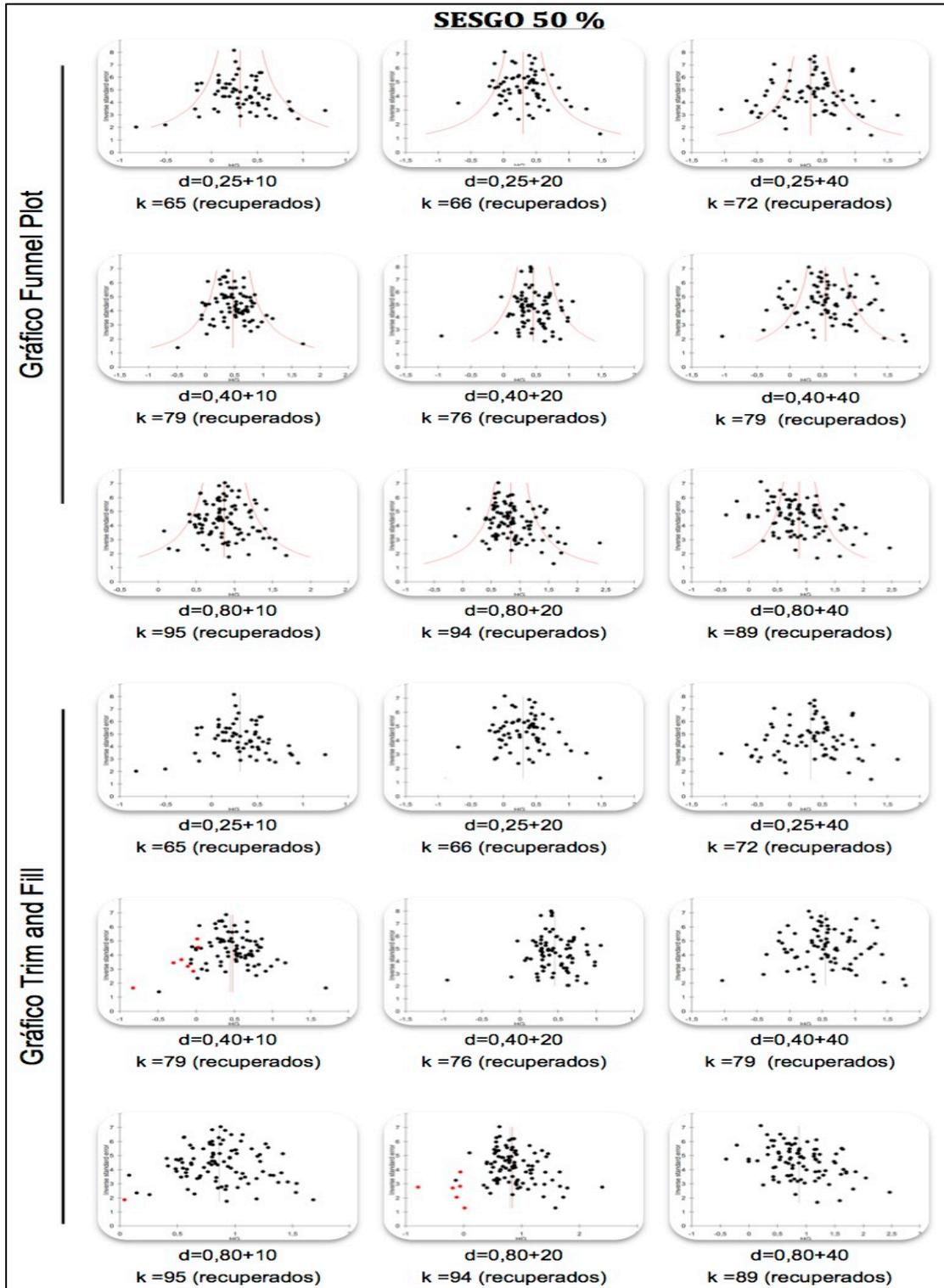


Figura 22. Funnel Plot y Trim And Fill bajo modelo efectos aleatorios, con cálculo del tamaño del efecto de Hedges, con variabilidad entre estudios ($\tau^2 = [0,10; 0,20; 0,40]$), mediante estimación de DerSimonian-Laird y diferentes grados de sesgo de publicación. Nota. Tamaño del efecto = “d” y número de estudios de cada metaanálisis = “k”.

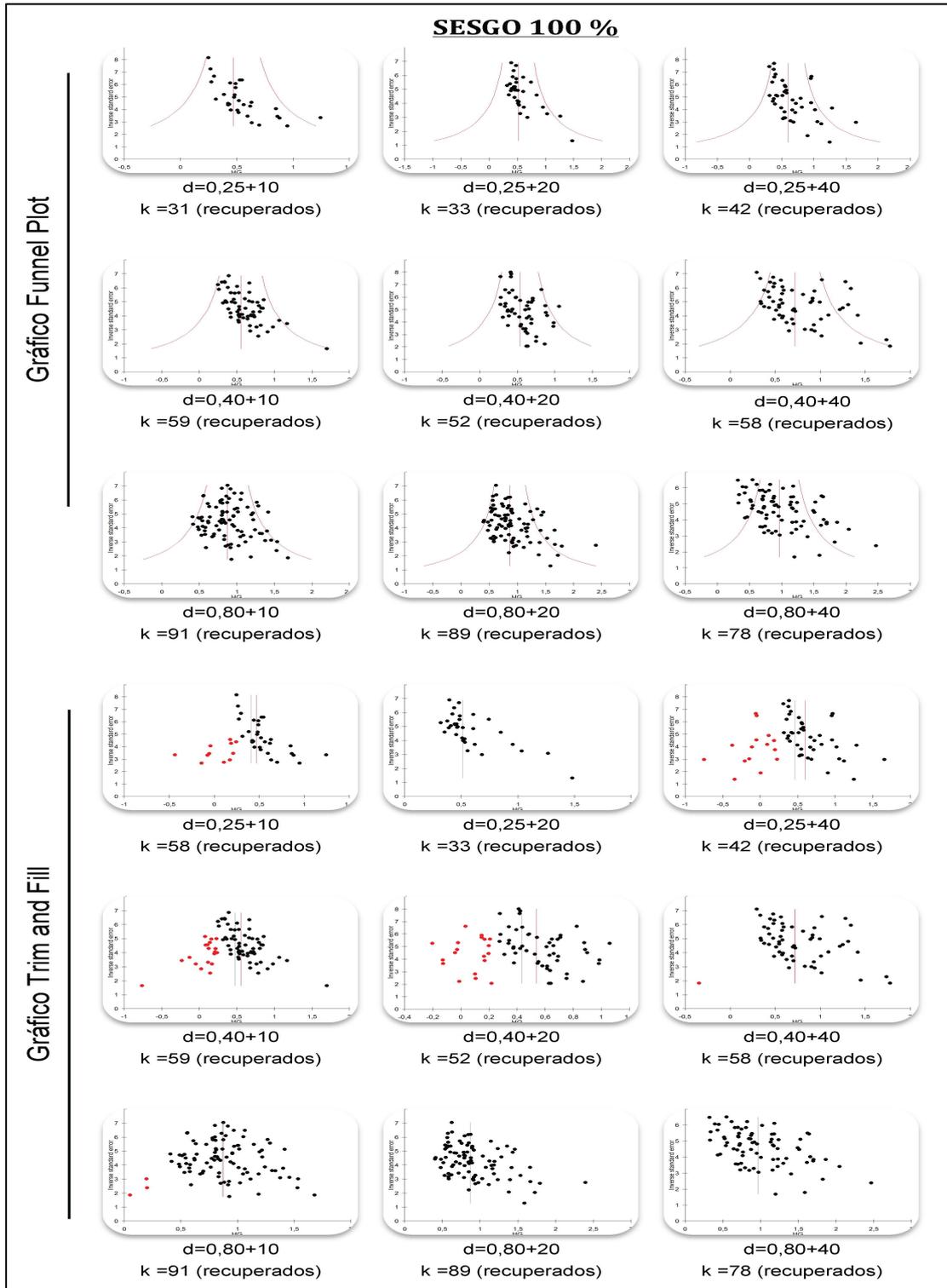


Figura 23. Funnel Plot y Trim And Fill bajo modelo efectos aleatorios, con cálculo del tamaño del efecto de Hedges, con variabilidad entre estudios ($\tau^2 = [0,10; 0,20; 0,40]$), mediante estimación del DerSimonian-Laird y diferentes grados de sesgo de publicación. Nota. Tamaño del efecto = "d" y número de estudios de cada metaanálisis = "k".

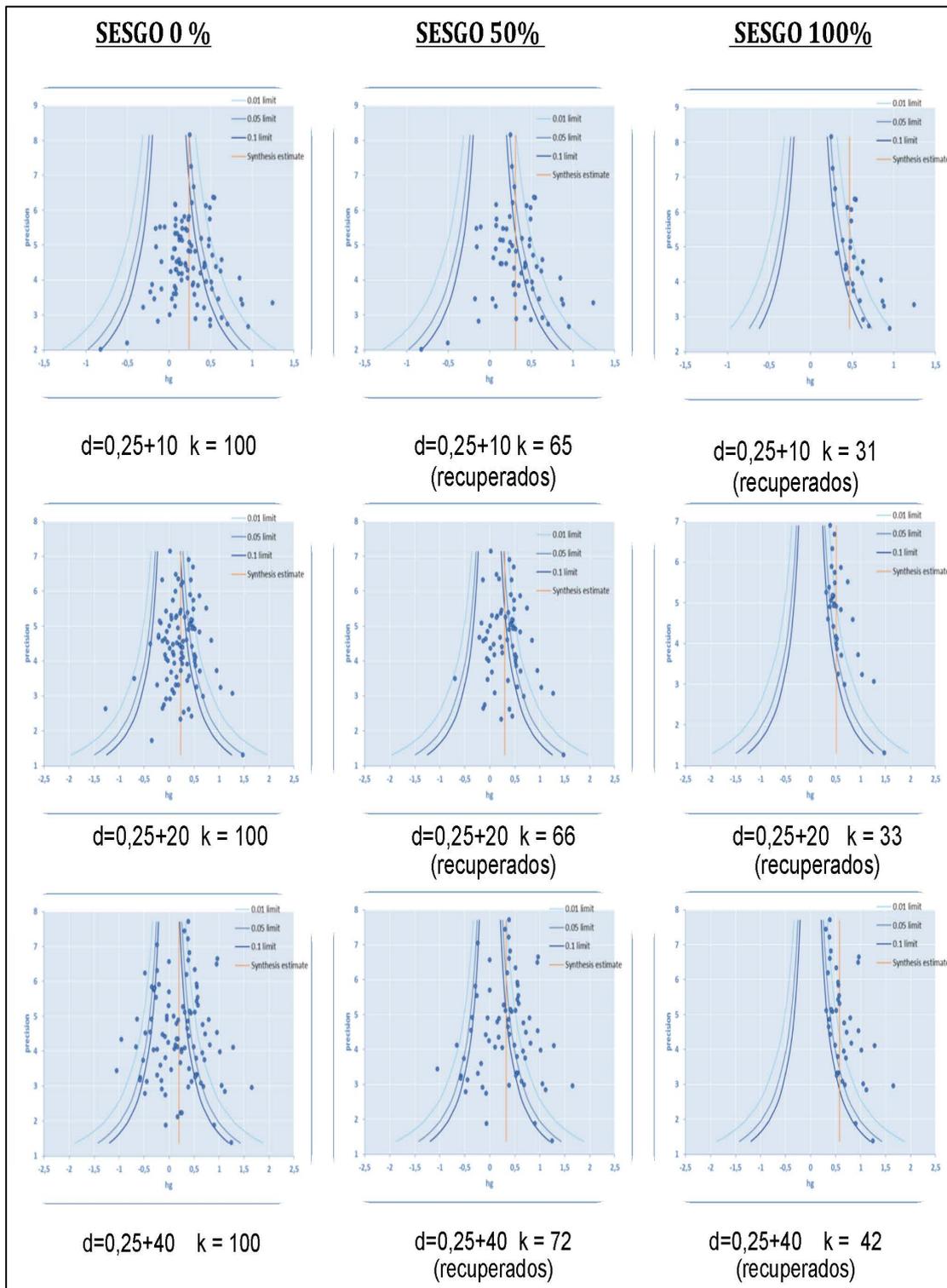


Figura 24. Contour-enhanced Funnel bajo modelo efectos aleatorios, con cálculo del tamaño del efecto de Hedges, con variabilidad entre estudios ($\tau^2 = [0,10; 0,20; 0,40]$), tamaño del efecto ($d = 0,25$) mediante estimación del DerSimonian-Laird y diferentes grados de sesgo de publicación. Nota. Tamaño del efecto = “d” y número de estudios de cada metaanálisis = “k”.

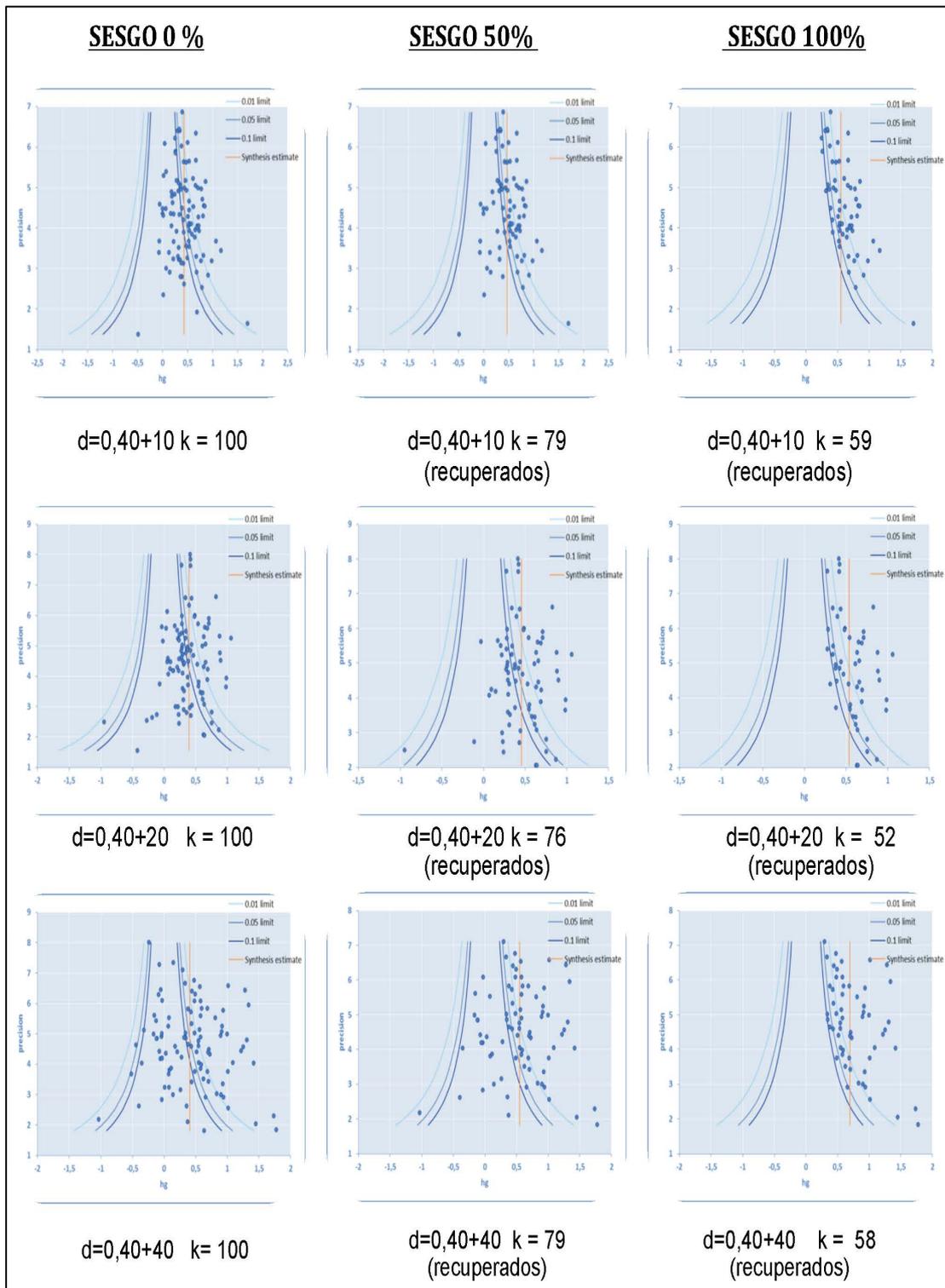


Figura 25. Contour-enhanced Funnel bajo modelo efectos aleatorios, con cálculo del tamaño del efecto de Hedges, con variabilidad entre estudios ($\tau^2 = [0,10; 0,20; 0,40]$), tamaño del efecto ($d = 0,40$) mediante estimación del DerSimonian-Laird y diferentes grados de sesgo de publicación. Nota. Tamaño del efecto = “d” y número de estudios de cada metaanálisis = “k”.

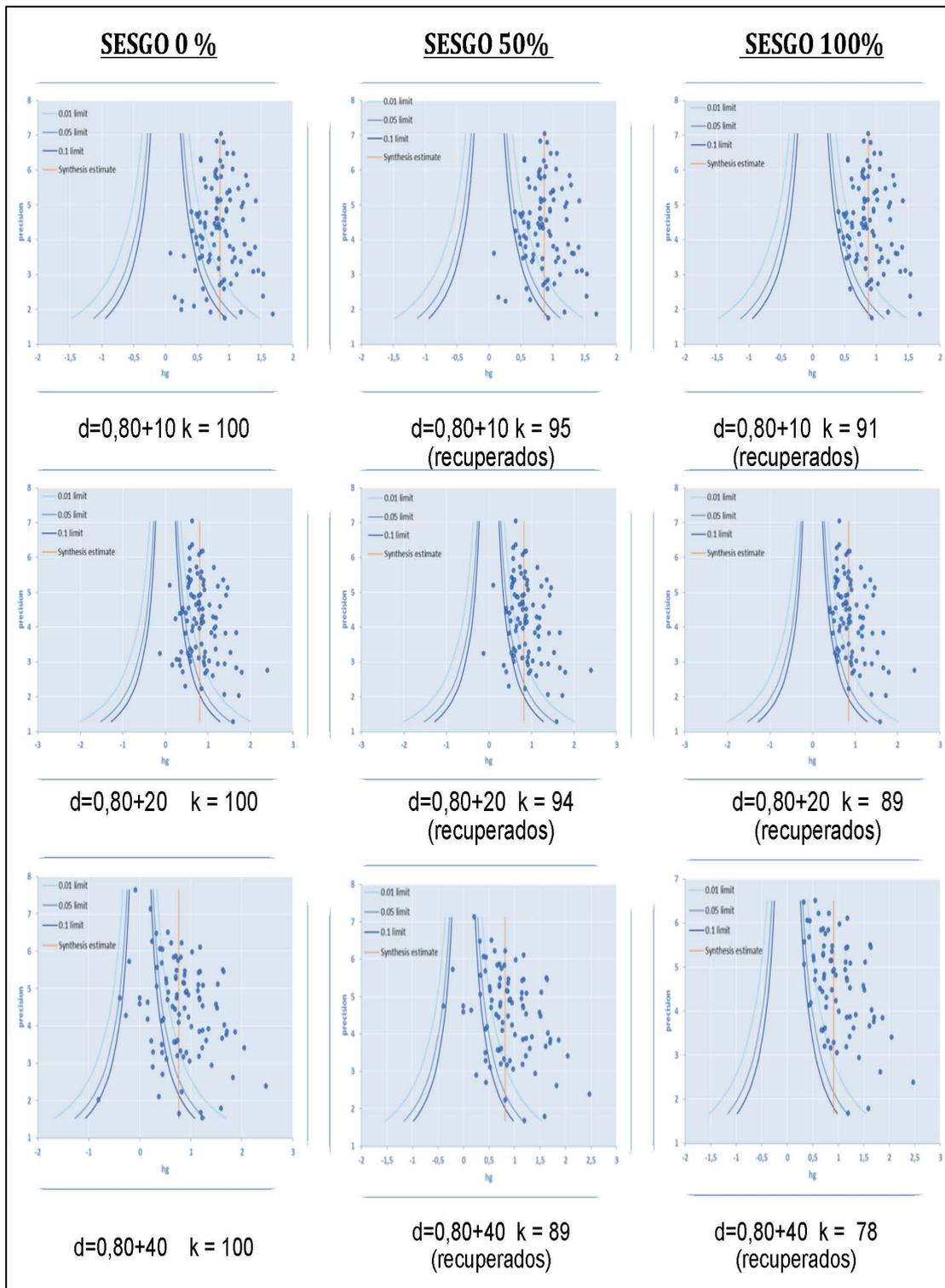


Figura 26. Contour-enhanced Funnel bajo modelo efectos aleatorios, con cálculo del tamaño del efecto de Hedges, con variabilidad entre estudios ($\tau^2 = [0,10; 0,20; 0,40]$), tamaño del efecto ($d = 0,80$) mediante estimación del DerSimonian-Laird y diferentes grados de sesgo de publicación. Nota. Tamaño del efecto = “d” y número de estudios de cada metaanálisis = “k”.

General	0% d = 0,25+0,1 k = 100	0% d = 0,25+0,2 k = 100	0% d = 0,25+0,4 k = 100	50% d = 0,25+0,1 k = 65 (recuperados)	50% d = 0,25+0,2 k = 66 (recuperados)	50% d = 0,25+0,4 k = 72 (recuperados)	100% d = 0,25+0,1 k = 31 (recuperados)	100% d = 0,25+0,2 k = 33 (recuperados)	100% d = 0,25+0,4 k = 42 (recuperados)
Resultado Original Meta analisis	0,2393	0,2312	0,1984	0,3127	0,3007	0,3171	0,4682	0,5165	0,6032
95% CI inferior	0,1918	0,1725	0,1056	0,2499	0,2282	0,2118	0,3981	0,446	0,5227
95% CI superior	0,2869	0,2899	0,2911	0,3755	0,3733	0,4224	0,5402	0,587	0,6636
Evaluación del efecto									
Correlación rango ajustado									
Tau-Kendall (continuity corrected)	0,0348	-0,0334	0,0106	0,1326	0,0351	-0,0158	0,5083	0,4603	0,4117
z	0,5133	-0,492	0,1566	1,5609	0,4168	-0,195	4,0263	3,7657	3,8417
p-valor (dos colas)	0,6077	0,6227	0,8756	0,1185	0,6768	0,8454	< 0,0001	0,0002	0,0001
Método de regresión	Egger	Egger	Egger	Egger	Egger	Egger	Egger	Egger	Egger
Intercepto	0,0831	-0,238	0,0104	0,5489	0,5278	-0,3216	2,4632	2,0672	1,9069
95% CI inferior	-0,835	-1,2856	-1,4779	-0,5746	-0,8233	-1,9964	1,6005	1,0574	0,7172
95% CI superior	1,0011	0,8096	1,4987	1,6724	1,8789	1,3432	3,3259	3,077	3,0966
p-valor (dos colas)	0,8579	0,653	0,999	0,3326	0,4381	0,7011	< 0,0001	0,0002	0,0024
Método de regresión	Macaskill	Macaskill	Macaskill	Macaskill	Macaskill	Macaskill	Macaskill	Macaskill	Macaskill
Pendiente	0,0002	0,0004	0	-0,0003	-0,0008	0,0006	-0,0025	-0,0036	-0,0023
95% CI inferior	-0,0012	-0,0013	-0,002	-0,002	-0,0029	-0,0017	-0,0036	-0,0055	-0,0037
95% CI superior	0,0016	0,0022	0,002	0,0015	0,0012	0,0029	-0,0013	-0,0018	-0,0008
p-valor (dos colas)	0,8108(0,7151)	0,627(0,4009)	0,9967(0,2956)	0,7583(< 0,0001)	0,411(0,8262)	0,6143(0,3523)	< 0,0001*	< 0,0001*	< 0,0001*
Análisis de sensibilidad									
Fail-safe N	2866	2624	1969	2203	2112	2704	1435	1803	3848
Nivel de Tolerancia	510	510	510	335	340	365	165	175	220
Método Trim And Fill (automático)	LO	LO	LO	LO	LO	LO	LO	LO	LO
Número de estudios	2	11	11	0	1	0	11	0	15
Resultado Meta analisis	0,244	0,2891	0,1043	0,3127	0,2882	0,3171	0,409	0,5165	0,4661
95% CI inferior	0,1955	0,2249	0,0041	0,2499	0,2253	0,2118	0,334	0,446	0,3705
95% CI superior	0,2925	0,3533	0,2044	0,3755	0,371	0,4224	0,484	0,587	0,5617

Figura 27. Datos estadísticos bajo modelo efectos aleatorios, con cálculo del tamaño del efecto de Hedges, con variabilidad entre estudios ($\tau^2 = [0,10; 0,20; 0,40]$), tamaño del efecto ($d = 0,25$) mediante estimación del DerSimonian-Laird y diferentes grados de sesgo de publicación.

Nota. Tamaño del efecto = "d" y número de estudios de cada metaanálisis = "k".
Método Macaskill entre paréntesis valor FIV (ponderado inversa de la varianza).
(* No hay cambios en relación al método no ponderado.

General	0% d = 0,40+0,1 k = 100	0% d = 0,40+0,2 k = 100	0% d = 0,40+0,4 k = 100	50% d = 0,40+0,1 k = 79 (recuperados)	50% d = 0,40+0,2 k = 76 (recuperados)	50% d = 0,40+0,4 k = 79 (recuperados)	100% d = 0,40+0,1 k = 59 (recuperados)	100% d = 0,40+0,2 k = 52 (recuperados)	100% d = 0,40+0,4 k = 58 (recuperados)
Resultado Original Meta análisis	0,3944	0,4315	0,4259	0,474	0,4576	0,5454	0,5518	0,5377	0,7224
95% CI inferior	0,3425	0,3802	0,3341	0,4162	0,4015	0,4498	0,4973	0,4834	0,6383
95% CI superior	0,4464	0,4827	0,5178	0,5317	0,5138	0,641	0,6064	0,592	0,8065
Evaluación del efecto									
Correlación rango ajustado									
Tau-Kendall (continuity corrected)	-0,0337	0,0903	0,0768	0,1162	0,04	0,0405	0,393	0,236	0,3061
z	-0,4961	1,3318	1,1327	1,5158	0,5107	0,5279	4,3976	2,4699	3,3941
p-valor (dos colas)	0,6198	0,1829	0,2573	0,1296	0,6095	0,5976	< 0,0001	0,0135	0,0007
Método de regresión	Egger	Egger	Egger	Egger	Egger	Egger	Egger	Egger	Egger
Intercepto	-0,4114	0,6499	0,8571	0,8136	0,0312	0,0459	2,5095	1,2886	1,9605
95% CI inferior	-1,2482	-0,234	-0,6967	-0,22	-0,9076	-1,5968	1,6681	0,346	0,4651
95% CI superior	0,4254	1,5338	2,4109	1,8472	0,97	1,6916	3,3509	2,2312	3,456
p-valor (dos colas)	0,3316	0,1478	0,2764	0,1211	0,9474	0,9558	< 0,0001	0,0084	0,0111
Método de regresión	Macaskill	Macaskill	Macaskill	Macaskill	Macaskill	Macaskill	Macaskill	Macaskill	Macaskill
Pendiente	0,0009	-0,0006	-0,0004	-0,0007	0,0003	0,0007	-0,0035	-0,0012	-0,0022
95% CI inferior	-0,0003	-0,0022	-0,0023	-0,0025	-0,001	-0,0017	-0,0048	-0,0021	-0,0041
95% CI superior	0,0021	0,0009	0,0016	0,0011	0,0015	0,003	-0,0022	-0,0003	-0,0004
p-valor (dos colas)	< 0,0001*	< 0,0001*	< 0,0001(0,6698)	< 0,0001*	< 0,0001*	0,5711(0,1742)	< 0,0001*	< 0,0001*	< 0,0001*
Análisis de sensibilidad									
Fail-safe N	7878	8963	9589	7062	6598	9823	6280	5130	10147
Nivel de Tolerancia	510	510	510	405	390	405	305	270	300
Método Trim And Fill (automático)	L0	L0	L0	L0	L0	L0	L0	L0	L0
Número de estudios	5	9	0	7	0	0	19	19	1
Resultado Meta análisis	0,4089	0,3943	0,4259	0,4403	0,4576	0,5454	0,4736	0,435	0,7173
95% CI inferior	0,3547	0,3389	0,3341	0,3788	0,4015	0,4498	0,4152	0,3696	0,6327
95% CI superior	0,4632	0,4496	0,5178	0,5019	0,5138	0,641	0,5321	0,5004	0,8019

Figura 28. Datos estadísticos bajo modelo efectos aleatorios, con cálculo del tamaño del efecto de Hedges, con variabilidad entre estudios ($\tau^2 = [0,10; 0,20; 0,40]$), tamaño del efecto ($d = 0,40$) mediante estimación del DerSimonian-Laird y diferentes grados de sesgo de publicación.
Nota. Tamaño del efecto = "d" y número de estudios de cada metaanálisis = "k".
Método Macaskill entre paréntesis valor FIV (ponderado inversa de la varianza).
(*) No hay cambios en relación al método no ponderado.

General	0% d = 0,80+0,1 k = 100	0% d = 0,80+0,2 k = 100	0% d = 0,80+0,4 k = 100	50% d = 0,80+0,1 k = 95 (recuperados)	50% d = 0,80+0,2 k = 94 (recuperados)	50% d = 0,80+0,4 k = 89 (recuperados)	100% d = 0,80+0,1 k = 91 (recuperados)	100% d = 0,80+0,2 k = 89 (recuperados)	100% d = 0,80+0,4 k = 78 (recuperados)
Resultado Original Meta análisis	0,855	0,8177	0,8137	0,8649	0,8448	0,8771	0,8762	0,8684	0,8659
95% CI inferior	0,8018	0,7492	0,7127	0,8117	0,7764	0,7774	0,8244	0,8026	0,8747
95% CI superior	0,9082	0,8861	0,9148	0,9182	0,9132	0,9767	0,9279	0,9341	1,057
Evaluación del efecto									
Correlación rango ajustado									
Tau-Kendall (continuity corrected)	-0,0277	0,1426	0,1577	0,0161	0,2187	0,2214	0,0783	0,2853	0,3104
z	-0,4081	2,102	2,3252	0,2317	3,1219	3,0719	1,0999	3,9599	4,0213
p-valor (dos colas)	0,6832	0,0355	0,0201	0,8168	0,0018	0,0021	0,2714	< 0,0001	< 0,0001
Método de regresión	Egger	Egger	Egger	Egger	Egger	Egger	Egger	Egger	Egger
Intercepto	-0,1719	1,0515	2,1159	0,1193	1,4633	2,9325	0,5398	1,9193	3,4342
95% CI inferior	-0,9849	-0,0275	0,4925	-0,7558	0,3946	1,1051	-0,3583	0,8869	1,7427
95% CI superior	0,6411	2,1305	3,7392	0,9843	2,5319	4,76	1,439	2,9517	5,1257
p-valor (dos colas)	0,6758	0,056	0,0112	0,7873	0,0078	0,002	0,2361	0,0004	0,0001
Método de regresión	Macaskill	Macaskill	Macaskill	Macaskill	Macaskill	Macaskill	Macaskill	Macaskill	Macaskill
Pendiente	0,0007	-0,0009	-0,0011	0,0002	-0,0016	-0,0022	-0,0005	-0,0025	-0,0032
95% CI inferior	-0,0005	-0,0028	-0,0034	-0,0011	-0,0035	-0,0047	-0,0017	-0,0043	-0,0056
95% CI superior	0,002	0,0011	0,0013	0,0015	0,0003	0,0004	0,0007	-0,0007	-0,0008
p-valor (dos colas)	< 0,0001*	< 0,0001*	< 0,0001*	< 0,0001*	< 0,0001*	< 0,0001*	< 0,0001*	< 0,0001*	< 0,0001*
Análisis de sensibilidad									
Fail-safe N	34776	30309	31974	33786	29251	31002	33286	28679	29903
Nivel de Tolerancia	510	510	510	485	480	455	465	455	400
Método Trim And Fill (automático)	LO	LO	LO	LO	LO	LO	LO	LO	LO
Número de estudios	0	0	0	1	6	0	3	0	0
Resultado Meta análisis	0,855	0,8177	0,8137	0,8629	0,8097	0,8771	0,8682	0,8684	0,8659
95% CI inferior	0,8018	0,7492	0,7127	0,8094	0,736	0,7774	0,8154	0,8026	0,8747
95% CI superior	0,9082	0,8861	0,9148	0,9165	0,8834	0,9767	0,9209	0,9341	1,057

Figura 29. Datos estadísticos bajo modelo efectos aleatorios, con cálculo del tamaño del efecto de Hedges, con variabilidad entre estudios ($\tau^2 = [0,10; 0,20; 0,40]$), tamaño del efecto ($d = 0,80$) mediante estimación del DerSimonian-Laird y diferentes grados de sesgo de publicación. Nota. Tamaño del efecto = "d" y número de estudios de cada metaanálisis = "k". Método Macaskill entre paréntesis valor FIV (ponderado inversa de la varianza). (*) No hay cambios en relación al método no ponderado.

10. Conclusiones

El sesgo de publicación está presente en casi todos los procesos de investigación siendo un problema conocido por todos los investigadores, los cuales deben valorar sus resultados en función de su presencia antes de la publicación de sus estudios.

Este trabajo evalúa el sesgo de publicación según el proceso de integración de estudios en los metaanálisis en función de la ausencia y/o presencia de los no significativos, siendo este tipo de sesgo el más estudiado en la bibliografía disponible.

Las conclusiones sobre la revisión sistemática realizada han sido:

1. Debe evaluarse el sesgo de publicación como un factor más en la valoración de la calidad del trabajo de investigación tanto en revisión sistemática como en metaanálisis.
2. El sesgo de publicación tiene diversas dimensiones y esferas como son la cantidad de información recuperada, el tamaño muestral del estudio, la calidad que aporta el estudio en cuestión,....
3. El número de métodos de detección y corrección del sesgo de publicación de los que se dispone en la actualidad son muchos y muy diversos, siendo los métodos gráficos de origen exploratorio y los basados en la asimetría que presente el Funnel Plot los más utilizados.
4. La heterogeneidad es otro factor que complementa la evaluación del sesgo de publicación y que influye, en mayor o menor medida, en los métodos para su detección sin existir una solución única y válida.

Las conclusiones sobre los análisis realizados con los distintos métodos sobre diferentes tipos de metaanálisis han sido:

1. El Método Fail-safe N solo ha funcionado con metaanálisis que integran pocos estudios en ausencia de heterogeneidad con tamaño del efecto pequeño y medio. Cuando se integran muchos estudios en el metaanálisis y además existe heterogeneidad de cualquier tipo entre estudios, no detecta la presencia del sesgo de publicación.
2. El Método Funnel Plot, se comporta generosamente bien con metaanálisis que integran muchos estudios, donde se visualiza mejor la asimetría en caso de sesgo de publicación. En metaanálisis que integran un número menor de estudios, no deja visualizar adecuadamente esta asimetría. Curioso es que en presencia de

heterogeneidad intrínseca se comporta mejor que en heterogeneidad de un solo factor.

3. El Método de correlación de rangos ha funcionado mejor en metaanálisis que integran muchos estudios y tamaño del efecto grande tanto en ausencia como en presencia intrínseca de heterogeneidad. Así se cumple la recomendación donde se pide cautela sobre resultados con ausencia de sesgo de publicación en metaanálisis que integran pocos estudios para este método.
4. El Método de Egger funciona mejor en metaanálisis que integran muchos estudios y tamaño del efecto grande en ausencia de heterogeneidad. Pero en presencia de heterogeneidad intrínseca se comporta bien hasta simulación con tamaño del efecto grande donde sus resultados ya no son tan fiables, provocando falsos positivos y negativos.
5. El Método de Macaskill solo ha funcionado en condiciones muy concretas donde los metaanálisis integran muchos estudios y tamaño del efecto pequeño en presencia de heterogeneidad intrínseca. En el resto, no se comporta bien generando falsos positivos y negativos.
6. El Método Trim And Fill se comporta de forma anárquica, sin seguir una tendencia en los diferentes escenarios. En metaanálisis que integran pocos estudios no funciona bien, y en presencia de heterogeneidad genera falsos negativos y positivos en los diferentes escenarios con sesgo de publicación.
7. El Método Contour-enhanced Funnel Plot se comporta, en general, bien detectando por asimetría el sesgo de publicación. Puede generar un falso positivo en ausencia de sesgo de publicación al presentar asimetría con tamaño del efecto grande, tanto exista heterogeneidad como si no, donde se desconocen las condiciones reales. La heterogeneidad, por si sola, no es la causa de que este método falle en esos escenarios. Parece estar más cercano a la detección del sesgo de publicación real, pero manteniendo el hecho de que con tamaño del efecto grande puede generar falsos positivos.

11. Bibliografía

- Auger, C. (1989). *Information Sources in Grey Literature* (2nd editio). London: Bowker Saur.
- Becker, B. J. (2006). Failsafe N or File-Drawer Number. *Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments*, (1979), 111–125. <https://doi.org/10.1002/0470870168.ch7>
- Begg, C. B. (1985). A measure to aid in the interpretation of published clinical trials. *Statistics in Medicine*, 4(1), 1–9. <https://doi.org/10.1002/sim.4780040103>
- Begg, C. B., & Mazumdar, M. (1994). Operating Characteristics of a Rank Correlation Test for Publication Bias. *Biometrics*, 50(4), 1088. <https://doi.org/10.2307/2533446>
- Bekelman, J. E., & Gross, C. P. (2015). of Interest in Biomedical Research, 289(4).
- Bero, L. (2013). Industry Sponsorship and Research Outcome. *JAMA Internal Medicine*, 173(7), 580. <https://doi.org/10.1001/jamainternmed.2013.4190>
- Biotext (Ed.). (1999). *How to review the evidence: systematic identification and review of the scientific literature Handbook series on preparing clinical practice guidelines. Chemistry & ...* (Commonweal). Camberra: National Health and Medical Research Council. Retrieved from https://www.nhmrc.gov.au/_files_nhmrc/publications/attachments/cp65.pdf
- Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2009). *Introduction to meta-analysis*. Wiley. https://doi.org/10.1007/978-3-319-14908-0_2
- Bruns, S. B., & Ioannidis, J. P. A. (2016). P-curve and p-hacking in observational research. *PLoS ONE*, 11(2), 1–13. <https://doi.org/10.1371/journal.pone.0149144>
- Cobos-Carbó, A., & Augustovski, F. (2011). Declaración CONSORT 2010: actualización de la lista de comprobación para informar ensayos clínicos aleatorizados de grupos paralelos. *Medicina Clinica*, 137(5), 213–215. <https://doi.org/10.1016/j.medcli.2010.09.034>

- Coburn, K. M., & Vevea, J. L. (2015). Publication bias as a function of study characteristics. *Psychological Methods*, 20(3), 310–330. <https://doi.org/10.1037/met0000046>
- Cochran, W. G. (1954). The Combination of Estimates from Different Experiments. *Biometrics*, 10(1), 101–129. Retrieved from <http://www.jstor.org/stable/3001666> .
- Cook, D. J., Guyatt, G. H., Ryan, G., Clifton, J., Buckingham, L., Willan, A., ... Oxman, A. D. (1993). Should in Unpublished. *Jama*, 269(21), 2749.
- Cooper, H., & Hedges, L. V. (1994). *The Handbook Of Research Synthesis*. Russell Sage Foundation, New York, (Vol. 16). New York: John Wiley & Sons, Ltd. [https://doi.org/10.1002/\(SICI\)1097-0258\(19970330\)16:6<713::AID-SIM430>3.0.CO;2-4](https://doi.org/10.1002/(SICI)1097-0258(19970330)16:6<713::AID-SIM430>3.0.CO;2-4)
- Coursol, A., & Wagner, E. E. (1986). Effect of positive findings on submission and acceptance rates: A note on meta-analysis bias. *Professional Psychology: Research and Practice*, 17(2), 136–137. <https://doi.org/10.1037/0735-7028.17.2.136>
- Davey Smith, G., & Egger, M. (1994). Who benefits from medical interventions? Treating low risk patients can be a high risk strategy. *British Medical Journal*, 308, 72–74.
- Dickersin, K. (1990). The Existence of Publication Bias and Risk Factors for Its Occurrence. *JAMA Internal Medicine*, 263(10).
- Duvall, S., & Tweedie, R. (2000). A non parametric “trim and fill” method of assessing publication bias in meta-analysis. *Journal of the American Statistical Association*, 95, 89–98. Retrieved from <http://www.stat.colostate.edu/statresearch/stattechreports/Technical Reports/1998/98-18 Taylor Tweedie.pdf>
- Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, 315(7109), 629–634. <https://doi.org/10.1136/bmj.316.7129.469>

- Egger, M., Zellweger-Zähner, T., Schneider, M., Junker, C., Lengeler, C., & Antes, G. (1997). Language bias in randomised controlled trials published in English and German. *Lancet*, 350(9074), 326–329. [https://doi.org/10.1016/S0140-6736\(97\)02419-7](https://doi.org/10.1016/S0140-6736(97)02419-7)
- Engels, E. A., Schmid, C. H., Terrin, N., Olkin, I., & Lau, J. (2000). Heterogeneity and statistical significance in meta-analysis: an empirical study of 125 meta-analyses- Eric. *Statistics in Medicine*, 19(September 1999), 1707–1728.
- Ferguson, C. J., & Heene, M. (2012). A Vast Graveyard of Undead Theories. *Perspectives on Psychological Science*, 7(6), 555–561. <https://doi.org/10.1177/1745691612459059>
- Franco, A., Malhotra, N., & Simonovits, G. (2014). Publication bias in the social sciences: Unlocking the file drawer. *Science*, 345(6203), 1502–1505. <https://doi.org/10.1126/science.1255484>
- García Santiago, L. (1998). *Manual Básico de literatura gris: el lado oscuro de la documentación*.
- Gerber, A. S., Malhotra, N., Dowling, C. M., & Doherty, D. (2010). Publication Bias in Two Political Behavior Literatures. *American Politics Research*, 38(4), 591–613. <https://doi.org/10.1177/1532673X09350979>
- Givens, G. H., Smith, D. D., & Tweedie, R. L. (1997). Publication Bias in Meta-Analysis: A Bayesian Data-Augmentation Approach to Account for Issues Exemplified in the Passive Smoking Debate. *Statistical Science*, 12(4), 221–250. <https://doi.org/10.1214/ss/1030037958>
- Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *American Educational Research Association*, 5(10), 3–8. <https://doi.org/10.3102/0013189X005010003>
- Grégoire, G., Derderian, F., & Le Lorier, J. (1995). Selecting the language of the publications included in a meta-analysis: Is there a tower of babel bias? *Journal of Clinical Epidemiology*, 48(1), 159–163. [https://doi.org/10.1016/0895-4356\(94\)00098-B](https://doi.org/10.1016/0895-4356(94)00098-B)

- Guimaraes Almeida, M. do R. (1998). *SISTEMAS Y REDES EN EL ÁMBITO María do Rosário Gu ; marñes Almeida*.
- Hackshaw, A. K., Law, M. R., & Wald, N. J. (1997). The accumulated evidence on lung cancer and environmental tobacco smoke. *BMJ (Clinical Research Ed.)*, *315*(7114), 980–988. <https://doi.org/10.1136/bmj.315.7114.980>
- He, L., & Liu, M. (1997). *Preparing, maintaining and disseminating the evidence*. Amsterdam, The Netherlands.
- Hedges, L. V. (1981). Distribution Theory for Glass ' s Estimator of Effect Size and Related Estimators, *6*(2), 107–128.
- Hedges, L. V. (1984). Estimation of Effect Size under Nonrandom Sampling: The Effects of Censoring Studies Yielding Statistically Insignificant Mean Differences. *Journal of Educational Statistics*, *9*(1), 61–85. <https://doi.org/10.3102/10769986009001061>
- Higgins, J. P. T., & Green, S. (2011). Manual Cochrane de revisiones sistemáticas de intervenciones. *Cochrane*, (March), 1–639. <https://doi.org/Available from www.cochrane-handbook.org>.
- Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*(11), 1539–1558. <https://doi.org/10.1002/sim.1186>
- Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ: British Medical Journal*, *327*(7414), 557–560. <https://doi.org/10.1136/bmj.327.7414.557>
- Hunter, & Schmidt. (2004). *Methods of Meta-Analyses: Correcting Error and Bias in Research Findings* (Newbury Pa). London: Sage Publications.
- Ioannidis, J. (2008). Interpretation of tests of heterogeneity and bias in meta-analysis. *Journal of Evaluation in Clinical Practice*, *14*(5), 951–957. <https://doi.org/10.1111/j.1365-2753.2008.00986.x>
- Jackson, D. (2006). The implications of publication bias for meta-analysis ' other

- parameter. *Statistics in Medicine*, (October 2004), 2911–2921.
<https://doi.org/10.1002/sim.2293>
- Juni, P., Holestein, F., Sterne, J., Bartlett, C., & Egger, M. (2002). Direction and impact of language bias in meta-analyses of controlled trials: empirical study. *International Journal of Epidemiology*, 31(1), 115–123.
<https://doi.org/10.1093/ije/31.1.115>
- Kepes, S., & Banks, G. C. (2012). Avoiding Bias in Publication Bias Research : The Value of “ Null ” Findings. *J Bus Psychol*. <https://doi.org/10.1007/s10869-012-9279-0>
- Lau, J., Antman, E., Jimenez-Silva, J., Kupelnick, B., Mosteller, F., & Chalmers, T. (1992). Cumulative Meta-analysis of Therapeutic Trials for Myocardial Infarction. *The New England Journal of Medicine*, 248–254.
- Lau, J., Schmid, C., & Chalmers, T. (1995). Cumulative Meta Analysis of Clinical Trials Builds Evidence for Exemplary Medical Care. *Science*, 48(1), 59–60.
[https://doi.org/10.1016/0895-4356\(94\)00112-4](https://doi.org/10.1016/0895-4356(94)00112-4)
- Lexchin, J. (2012). Sponsorship bias in clinical research. *International Journal of Risk and Safety in Medicine*, 24(4), 233–242. <https://doi.org/10.3233/JRS-2012-0574>
- Lexchin, J., Bero, L. A., Djulbegovic, B., & Clark, O. (2003). Pharmaceutical industry sponsorship and research outcome and quality: systematic review, 326(May).
- Light, R., & Pillemer, D. (1984). *Summing UP. The Science of Reviewing Research. Contemporary Sociology* (Vol. 5).
- Macaskill, P., Walter, S. D., & Irwig, L. (2001). A comparison of methods to detect publication bias in meta-analysis. *Statistics in Medicine*, 20(4), 641–654.
<https://doi.org/10.1002/sim.698>
- Malley, P. (2009). *Sensitivity Methods for Publication Bias in a Meta-Analysis Department of Statistics*. Warwick.
- Martín-Vallejo, J. (1995). *Métodos estadísticos en meta-análisis. Departamento de Estadística y Matemática Aplicadas*.

- McAuley, L., Pham, B., Tugwell, P., & Moher, D. (2000). Does the inclusion of grey literature influence estimates of intervention effectiveness reported in meta-analyses? *The Lancet*, *356*(9237), 1228–1231. [https://doi.org/10.1016/S0140-6736\(00\)02786-0](https://doi.org/10.1016/S0140-6736(00)02786-0)
- Million, M., & Raoult, D. (2012). Publication biases in probiotics. *European Journal of Epidemiology*, *27*(11), 885–886. <https://doi.org/10.1007/s10654-012-9740-4>
- Misakian, A. L., & Bero, L. A. (1998). Publication bias and research on passive smoking: comparison of published and unpublished studies. *Jama*, *280*(3), 250–253. <https://doi.org/10.1001/jama.280.3.250>
- Moher, D., Fortin, P., Jadad, A. R., Jüni, P., Klassen, T., Le Lorier, J., ... Penna, A. (1996). Completeness of reporting of trials published in languages other than English: Implications for conduct and reporting of systematic reviews. *Lancet*, *347*(8998), 363–366. [https://doi.org/10.1016/S0140-6736\(96\)90538-3](https://doi.org/10.1016/S0140-6736(96)90538-3)
- Moher, D., Pham, Klassen, T. P., Schulz, K. F., Berlin, J. A., Jadad, A. R., & Liberati, A. (2000). What contributions do languages other than English make on the results of meta-analyses? *Journal of Clinical Epidemiology*, *53*(9), 964–972. [https://doi.org/10.1016/S0895-4356\(00\)00188-8](https://doi.org/10.1016/S0895-4356(00)00188-8)
- Naci, H., Dias, S., & Ades, A. E. (2014). Industry sponsorship bias in research findings: a network meta-analysis of LDL cholesterol reduction in randomised trials of statins. *Bmj*, *349*(oct03 1), g5741–g5741. <https://doi.org/10.1136/bmj.g5741>
- Niemeyer, H., Musch, J., & Pietrowsky, R. (2014). El sesgo de publicación en metanálisis sobre la eficacia de las intervenciones psicoterapéuticas para la Esquizofrenia. *Revista de Toxicomanias*, 32–48.
- Orwin, R. (1983). A Fail-Safe N for Effect Size in Meta-Analysis. *Journal of Educational Statistics*, *8*(2), 157–159.
- Palma Pérez, S., & Delgado Rodríguez, M. (2006). Consideraciones prácticas acerca de la detección del sesgo de publicación. *Gaceta Sanitaria*, *20*(Supl 3), 10–16. <https://doi.org/10.1157/13101085>

- Perez, A. (2009). *Heterogeneidad y Sesgo de Publicación en Metaanálisis*. Universidad de Jaén.
- Peters, J. L. (2006). Comparison of Two Methods to Detect Publication Bias in Meta-analysis. *Jama*, 295(6), 676. <https://doi.org/10.1001/jama.295.6.676>
- Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., & Rushton, L. (2008). Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *Journal of Clinical Epidemiology*, 61(10), 991–996. <https://doi.org/10.1016/j.jclinepi.2007.11.010>
- Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., Rushton, L., & Moreno, S. G. (2010). Assessing publication bias in meta-analyses in the presence of between-study heterogeneity. *Journal of the Royal Statistical Society. Series A: Statistics in Society*, 173(3), 575–591. <https://doi.org/10.1111/j.1467-985X.2009.00629.x>
- Pham, B., Klassen, T. P., Lawson, M. L., & Moher, D. (2005). Language of publication restrictions in systematic reviews gave different results depending on whether the intervention was conventional or complementary. *Journal of Clinical Epidemiology*, 58(8), 769–776. <https://doi.org/10.1016/j.jclinepi.2004.08.021>
- Pocock, S., Hughes, M., & Lee, R. (1987). Statistical problems in the reporting of clinical trials. *The New England Journal of Medicine*.
- Rennie, D. (1997). Thyroid storm. *Journal of the American Medical Association*, 277(15), 1238–1243.
- Rosa, M. C. (2016). Sesgo de publicación: ¿existe también en estudios de metanálisis? *CUIDEN*, 25, 7–8.
- Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, 86(3), 638–641. <https://doi.org/10.1037/0033-2909.86.3.638>
- Rothstein, H., Sutton, A., & Borenstein, M. (2005). *Publication Bias in Meta-Analysis*. (H. R. Rothstein, A. J. Sutton, & M. Borenstein, Eds.), *BMC Medical Research Methodology* (Vol. 1). Chichester, UK: John Wiley & Sons, Ltd.

<https://doi.org/10.1002/0470870168>

- Sackett, D., Rosenberg, W., Gray, M., Haynes, B., & Richardson, S. (1996). Evidence based medicine: what it is and what it isn't. *British Medical Journal*, *312*(7023), 71–72. <https://doi.org/10.1136/bmj.312.7023.71>
- Simes, R. J. (1987). Confronting publication bias: A cohort design for meta-analysis. *Statistics in Medicine*, *6*(1), 11–29. <https://doi.org/10.1002/sim.4780060104>
- Simonsohn, U., & Nelson, L. D. (2013). P- curve : A Key To The File Drawer. *Forthcoming, Journal of Experimental Psychology: General P-Curve.*, 1–40.
- Smith, M. L., & Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist*, *32*(9), 752–760. <https://doi.org/10.1037/0003-066x.32.9.752>
- Sterling, D. (1959). Publication Decisions and Their Possible Effects on Inferences Drawn from Tests of Significance--Or Vice Versa. *Journal of the American Statistical Association*, *54*(285), 30–34.
- Suarez-Almazor, M. E., Belseck, E., Homik, J., Dorgan, M., & Ramos-Remus, C. (2000). Identifying clinical trials in the medical literature with electronic Databases: MEDLINE alone is not enough. *Controlled Clinical Trials*, *21*(5), 476–487. [https://doi.org/10.1016/S0197-2456\(00\)00067-2](https://doi.org/10.1016/S0197-2456(00)00067-2)
- Sugita, M., Kanamori, M., Izuno, T., & Miyakawa, M. (1992). Estimating a summarized odds ratio whilst eliminating publication bias in meta-analysis. *Japanese Journal of Clinical Oncology*, *22*(5), 354–358. <https://doi.org/10.1093/oxfordjournals.jjco.a039577>
- Terrin, N., Schmid, C. H., & Lau, J. (2005). In an empirical evaluation of the funnel plot, researchers could not visually identify publication bias. *Journal of Clinical Epidemiology*, *58*(9), 894–901. <https://doi.org/10.1016/j.jclinepi.2005.01.006>
- Terrin, N., Schmid, C. H., Lau, J., & Olkin, I. (2003). Adjusting for publication bias in the presence of heterogeneity. *Statistics in Medicine*, *22*(13), 2113–2126. <https://doi.org/10.1002/sim.1461>

- Thompson, S. G., & Sharp, S. J. (1999). Explaining heterogeneity in meta-analysis: A comparison of methods. *Statistics in Medicine*, 18(20), 2693–2708. [https://doi.org/10.1002/\(SICI\)1097-0258\(19991030\)18:20<2693::AID-SIM235>3.0.CO;2-V](https://doi.org/10.1002/(SICI)1097-0258(19991030)18:20<2693::AID-SIM235>3.0.CO;2-V)
- Thornton, A., & Lee, P. (2000). Publication bias in meta-analysis: Its causes and consequences. *Journal of Clinical Epidemiology*, 53(2), 207–216. [https://doi.org/10.1016/S0895-4356\(99\)00161-4](https://doi.org/10.1016/S0895-4356(99)00161-4)
- van Assen, M. A. L. M., van Aert, R. C. M., & Wicherts, J. M. (2015). Meta-analysis using effect size distributions of only statistically significant studies. *Psychological Methods*, 20(3), 293–309. <https://doi.org/10.1037/met0000025>
- Van Lent, M., Overbeke, J., & Out, H. J. (2014). Role of editorial and peer review processes in publication bias: Analysis of drug trials submitted to eight medical journals. *PLoS ONE*, 9(8). <https://doi.org/10.1371/journal.pone.0104846>
- Vevea, J. L., & Woods, C. M. (2005). Publication Bias in Research Synthesis: Sensitivity Analysis Using A Priori Weight Functions. *Psychological Methods*, 10(4), 428–443. <https://doi.org/10.1037/1082-989X.10.4.428>
- Vickers, A., Goyal, N., Harland, R., & Rees, R. (1998). Do certain countries produce only positive results? A systematic review of controlled trials. *Controlled Clinical Trials*, 19(2), 159–166. [https://doi.org/10.1016/S0197-2456\(97\)00150-5](https://doi.org/10.1016/S0197-2456(97)00150-5)
- Viechtbauer, W. (2010). Journal of Statistical Software. *Journal of Statistical Software* August, 36(3).