

Revisión de diferentes Herramientas Estadísticas usadas en
Metanálisis en el ámbito de la Salud Pública y la
Epidemiología.

Documento de Trabajo presentado al 8^{vo} CLATSE

Autor: Ramón Álvarez*

Montevideo Octubre 7-10 de 2008

* IESTA Instituto de Estadística, Facultad de Ciencias Económicas, UDELAR

Índice

1. Metanálisis en Epidemiología y Salud Pública	3
2. Diferentes tipos de modelos usados	6
2.1. Modelos de efectos fijos	6
2.2. Modelos de efectos aleatorios	7
3. Perspectiva Frecuentista vs Bayesiana	7
4. Herramientas Estadísticas	8
5. Métodos de medición en estudios individuales	10
5.1. Peto Odds Ratio	10
5.2. Odds Ratio	10
5.3. Diferencias de Riesgos	10
5.4. Riesgo Relativo	11
5.5. Estimaciones Individuales para respuesta a escala continua	11
5.6. Curvas diagnósticas	11
6. Estimación para Estudios Individuales: Método Genérico	12
7. Métodos de Medición para estudios combinados (metanálisis)	12
7.1. Métodos de Mantel-Haenszel	12
7.2. Tests de Heterogeneidad	13
7.3. Modelos de Efectos Aleatorios	14
7.4. Intervalos de Confianza	15
7.5. Cálculo de Indices Globales para Curvas Diagnósticas	15
8. Herramientas Gáficas en Metanálisis	16
8.1. Gráfico Simple (Forest plot)	16
8.2. Gráfico para análisis acumulativo	17
8.3. Gráfico Radial (Galbraith plot)	19
8.4. Gráfico De Embudo(Funnel plot)	19
8.5. Gráfico de L'Abbée	22
8.6. Gráficos de Sesgos	22
8.7. Tests de Heterogeneidad en Curvas Diagnósticas	23
9. Tratamiento del Sesgo en metanálisis	24
10. Metaregresión	25

11.Revisión de diferentes programas de Computación disponibles para el metanálisis	27
12.Revman	27
13.Librerías en R	28
13.1. librería <i>rmeta()</i>	28
13.2. librería <i>meta()</i>	34
13.3. librería <i>copas()</i>	37

Resumen

Las aproximaciones clásicas para resumir la evidencia incluyen revisiones narrativas, revisiones sistemáticas y meta-análisis. En una revisión tradicional de la literatura clínica y epidemiológica, el experto decide cuales son los resultados más relevantes sobre un tema de interés y destaca sus hallazgos en términos de resultados, incidiendo en menor medida en los aspectos metodológicos. Entre las limitaciones de este tipo de revisiones se encuentra la posibilidad de incurrir en sesgos que provengan (a) de los estudios originales, (b) del investigador causados por la inclusión subjetiva de los estudios, (c) de la calidad deficiente de los estudios, o (d) de una interpretación errónea de la respuesta., generándose por lo tanto, sesgo de selección, el sesgo de realización (performance bias), el sesgo de desgaste o pérdida (attrition bias) y el sesgo de detección.

El insumo utilizado por estos métodos son las medidas de resumen que surgen cada estudio individual y que aparecen en los informes publicados de estudios individuales, lo que hace que se puede hablar de que el metanálisis es un análisis de un análisis estadístico. En esta revisión se presentan las diferentes fuentes de variabilidad entre estudios y los diferentes tipos de modelos usados (Modelos de efectos fijos, Modelos de efectos aleatorios), así como las ventajas y debilidades de los mismos. A su vez ir construyendo y modificando en forma secuencial y sistemática la incertidumbre sobre la respuesta que se pretende evaluar en las intervenciones de los estudios, usando el enfoque bayesiano.

Se presenta como trabajar mediante el uso de herramientas gráficas con los sesgos y la metaregresión, usando para eso 4 librerías implementadas en R y por otro lado el software RevMan.

1. Metanálisis en Epidemiología y Salud Pública

Las aproximaciones clásicas para resumir la evidencia incluyen revisiones narrativas, revisiones sistemáticas y meta-análisis. En una revisión tradicional de la literatura clínica y epidemiológica, el experto decide cuales son los resultados más relevantes sobre un tema de interés y destaca sus hallazgos en términos de resultados, incidiendo en menor

medida en los aspectos metodológicos. Entre las limitaciones de este tipo de revisiones se encuentra la posibilidad de incurrir en sesgos que provengan (a) de los estudios originales, (b) del investigador causados por la inclusión subjetiva de los estudios, (c) de la calidad deficiente de los estudios, o (d) de una interpretación errónea de la respuesta.

El término meta-análisis fue introducido por primera vez en 1976 [10] dentro del campo de las ciencias de la educación, para designar 'todo análisis estadístico de un gran colección de resultados de la literatura individual, con el propósito de integrar los resultados'. El insumo utilizado por este método (los datos) son las medidas de resumen (estadísticas) que surgen en cada estudio individual y que aparecen en los informes publicados de estudios individuales, lo que hace que se pueda hablar de que el metanálisis es *un análisis de un análisis estadístico*. Esta forma de sintetizar la información, hace que en las revisiones sistemáticas se termine luego usando Meta-análisis para poder generalizar resultados con una mayor validez externa (al incluir estudios de diferentes poblaciones). Al integrar los estudios existe un efecto secundario que en principio es muy favorable, al aumentar el tamaño de muestra, con el consiguiente cambio en la *potencia estadística, precisión mayor* al tener *intervalos de confianza* mas chicos (el error estandar de la variable de respuesta, efecto del tratamiento, es función del tamaño muestral, con una relación cuadrática inversa), aunque deben de ser tenidos en cuenta algunos aspectos que se presentan en (3). A su vez esta forma de investigar es menos costosa y mas sencillo de implementar al no tener que hacer estudios con tamaños muestrales tan grandes.

Hay una serie de aspectos que hay que tener en cuenta cuando se intenta evaluar la información recopilada para llevar adelante un metanálisis que tiene que ver con

- Métodos para resumir la validez de los estudios
- Tipos de sesgos
- Calidad de los estudios individuales

Los revisores que luego harán el *metanálisis*(MA) deben de seguir una serie de pautas para poder descartar problemas y poder manejar la heterogeneidad de los resultados. Existe un manual de revisiones desarrollado por *The Cochrane Collaboration* [1] [2] que en la actualidad esta en la **versión 4.2.6** actualizada a setiembre de 2006.

Debemos tratar de lograr tener *Validez* en el contexto de una revisión sistemática, como paso previo a un metanálisis, entendiéndola como la capacidad de evitar los errores sistemáticos o sesgos mediante el diseño. (Moher 1995).

Los estudios más rigurosos tienen mayor probabilidad de dar resultados más próximos a la *verdad*, mientras que los estudios menos rigurosos pueden estar sesgados hacia la sobreestimación de la eficacia de una intervención, dando lugar a conclusiones 'positivas

falsas' (concluyendo erróneamente que una intervención es eficaz).o por el contrario se puede llegar a conclusiones 'negativas falsas' (concluyendo errneamente que no tienen ningn efecto) si los estudios menos rigurosos están sesgados de manera que subestiman el efecto de una intervención (Detsky 1992).

Hay varias maneras de evaluar la validez. Una consiste en asignar los valores *se cumple*, *no se cumple*, o *'es poco claro'* a los criterios individuales y al mismo tiempo, usar los criterios individuales, como el grado de ocultamiento de las asignaciones, en los análisis de sensibilidad , (es decir hasta donde los resultados encontrados son sensibles a cambios en la forma en la que fue hecha la revisión y el (MA)). Se pueden manejar *escalas y listas de 'calidad'* (checklists) que aparecen en los libros de texto sobre ensayos clínicos(Moher 1995, Moher 1996b) , incorporando sistema de puntuaciones (se puntúa cada estudio) para luego poderarlo al hacer el análisis estadístico , como por ejemplo la metaregresión (10)

Los revisores deben de manejar varios tipos de sesgo o errores sistemáticos: el sesgo de selección, el sesgo de realización (performance bias), el sesgo de desgaste o pérdida (attrition bias) y el sesgo de detección. Lamentablemente, no hay evidencia empírica salida de una relación entre los resultados de los estudios y los criterios o conjuntos de criterios específicos usados para evaluar el riesgo de estos sesgos (Moher 1995, Moher 1996b). Sin embargo, hay fundamentos lógicos para creer que esta relación existe y por lo tanto, es razonable tener en cuenta estos cuatro sesgos potenciales en el momento de evaluar los estudios para una revisin (Feinstein 1985).

Sesgo de selección Uno de los factores ms importantes que puede sesgar y distorsionar las comparaciones entre tratamientos est relacionado con el proceso de la formación de los grupos a comparar (Kunz 1998). En el diseño de ensayos clínicos es muy importante utilizar un método apropiado para prevenir que se conozca de antemano la asignación de los tratamientos (técnicas de cegamiento) ,usando luego como parte de los criterios para evaluar la validez.

Sesgo de realización Con 'sesgo de realización' (performance bias) se hace referencia a las diferencias sistemáticas en la atención sanitaria proporcionada a los participantes en los grupos de comparación, a parte de la intervención bajo estudio. Para proteger contra las diferencias en la atención y contra el efecto placebo, puede 'cegar' a aquellos que dan y reciben la atención para que no conozcan el grupo al están asignados quienes la reciben; también se usa la evaluación de este tipo de sesgo como criterio de validez.

Sesgo de desgaste El Sesgo de desgaste (attrition bias) se refiere a las diferencias sistemáticas entre los grupos de comparación en cuanto a la pérdida de participantes durante el estudio.Aca los revisores deben ser prudentes al utilizar la información de seguimiento como un criterio de validez ya que la información sobre como se tratan las pérdidas

de participantes (por abandonos, bajas, desviaciones de protocolo, etc.) es insuficiente en algunos estudios y muy heterogénea pudiendo ocasionar efectos distorsivos al querer incorporarlo a la validez.

El *sesgo de detección* se refiere a las diferencias sistemáticas entre los grupos de comparación en la evaluación de resultados al tener un reporte selectivo de resultados. Este tipo de sesgo también se puede modelar y medir, por ejemplo a través de (9).

2. Diferentes tipos de modelos usados

A su vez existen diferentes alternativas de como situarse desde el punto de vista estadístico en la evaluación de los resultados al depender de que tipo de información se dispone, como ha sido relevada, como ha sido informada (ya que no siempre coincide) y por lo tanto que procedimientos a seguir tanto para la estimación como para la interpretación de los resultados.

2.1. Modelos de efectos fijos

En este primer tipo de modelo se asume un supuesto que termina siendo un poco simplificador y es que los estudios son homogéneos; es decir que los criterios de selección de pacientes son los mismos siempre, que las poblaciones que estos representan son las mismas, que las intervenciones se aplican siempre de la misma manera. Este aspecto en cierta forma implica un criterio de *intercambiabilidad* y desde el punto de vista estadístico se debe de considerar lo siguiente:

Dada una muestra *grande* de estudios individuales, se puede medir la variable de efecto que para ser un estadístico *suficiente* se tiene que $Y_i \sim N(\mu, \sigma_i^2)$

$$\sum_i^I \frac{1}{\sigma_i^2} Y_i \quad (1)$$

$$\hat{\mu} \sim N \left(\mu, \left(\sum_i^I \frac{1}{\sigma_i^2} \right)^{-1} \right) \quad (2)$$

Al suponer que existe una única fuente de variabilidad y que es dentro de los estudios (*within-study variation*) para poder comparar los efectos de los estudios tenemos

$$Q = \sum_i^I \frac{1}{\sigma_i^2} (Y_i - \hat{\mu})^2 \sim \chi_{I-1}^2 \quad (3)$$

Existen una serie de tests que pueden evaluar esta situación pero si se usan los modelos de efectos fijos (MEF) se termina por subestimar la variabilidad y por lo tanto las conclusiones que se hagan a nivel global (cuando se hace el análisis acumulado) sean erróneas.

2.2. Modelos de efectos aleatorios

En los modelos de efectos aleatorios (MEA)[18] , se descarta el supuesto restrictivo de homogeneidad, teniendo por lo tanto que modelar un *efecto aleatorio* θ_i . En ese caso se supone $Y_i \sim N(\theta_i, \sigma_i^2)$, donde cada estudio individual para poder ser intercambiable es tal que

$$\theta_i \sim N(\mu_\theta, \tau^2); \hat{\mu}_\theta = \frac{\sum_1^I (\sigma_i^2 + \tau^2)^{-1} y_i}{\sum_1^I (\sigma_i^2 + \tau^2)^{-1}} \quad (4)$$

en este caso τ^2 es la *variabilidad entre estudios* (between-study variability)

3. Perspectiva Frecuentista vs Bayesiana

A pesar de trabajar con modelos de efectos aleatorios, si se trabaja desde una *perspectiva frecuentista* toda las estimaciones que están basadas en la *teoría de la máxima verosimilitud y teoría asintótica* que requieren tamaños muestrales muy grandes. Por otra parte los tests clásicos para medir la homogeneidad entre estudios tiene potencia estadística baja. Una solución de compromiso para este problema de situaciones que cambian mucho entre un estudio y otro es pensar en 2 cosas

- luego de observar la respuesta en un estudio(el primero) x_1 se tiene una distribución a *posteriori*

$$\underbrace{f(\theta|x_1)}_{\text{posteriori}} \propto \underbrace{f(x_1|\theta)}_{\text{verosimilitud}} \underbrace{f(\theta)}_{\text{priori}}$$

- El segundo estudio va a brindar información, la que se va a usar para ajustar la creencia a *priori* $f(\theta|x_1)$ para tener

$$\underbrace{f(\theta|x_1, x_2)}_{\text{posteriori}} \propto \underbrace{f(x_2|\theta)}_{\text{verosimilitud}} \underbrace{f(\theta|x_1)}_{\text{priori}}$$

Stangl [18] En el *enfoque Bayesiano* , se realizan *inferencias* usando lo que se conoce como la *distribución a posteriori* de la cantidades de interés. La distribución a posteriori es un compromiso entre el conocimiento a priori del problema, el cual está reflejado en una *distribución a priori* y la información contenida en los datos a través de la función de *Verosimilitud*. Gelman ~ Andrew [9]

$$p(\theta, y) = p(\theta)p(y|\theta) \quad (5)$$

donde $p(\theta)$ es la distribución a priori y $p(y|\theta)$ es la verosimilitud. Utilizando esta igualdad, la regla de Bayes permite concluir que

$$p(\theta|y) = \frac{p(\theta, y)}{p(y)} = \frac{p(\theta)(y|\theta)}{\int p(\theta)p(y|\theta)\partial\theta} \quad (6)$$

De esta manera se puede ir construyendo y modificando en forma secuencial y sistemática la incertidumbre sobre la respuesta que se pretende evaluar en las intervenciones de los estudios. Esa forma de razonar y evaluar en la perspectiva bayesiana se puede desarrollar combinándola con un esquema en niveles o jerarquías dando lugar a un *modelo jerárquico bayesiano* para metanálisis

1. Nivel I $Y_i \sim N(\theta_i, \sigma_i^2)$
2. Nivel II $\theta_i \sim N(\mu, \tau^2)$
3. Nivel III $\mu \sim \pi(\mu), \tau^2 \sim \pi(\tau^2)$

Acá Y_i es el efecto del i-ésimo estudio con su correspondiente (DE)(desvío estándar). Con este modelo jerárquico se trabaja en forma conjunta con el efecto de cada estudio y un efecto medio en la población, para tener una media a posteriori para θ_i que no es otra cosa que un promedio ponderado.

4. Herramientas Estadísticas

Debemos considerar las diferentes formas de medir los resultados que se obtienen en los estudios que luego se pretenden combinar; y es en función de estos que luego podremos evaluar y hallar efectos globales y el error asociado. Para eso tenemos una forma genérica de resumir los resultados de cada estudio con formato tabular como aparece en

Cuando tenemos respuestas que provienen de estudios medidos a *escala binaria* nos queda

Estudio _i	Evento	No Evento	Total
Experimental	a_i	b_i	n_{ei}
Control	c_i	d_i	n_{ci}

Cuadro 2: Datos Binarios (Estudios Individuales)

Si las variables de respuesta están medidas a *escala continua* nos queda

Existen algunos estudios donde los resultados presentados por los investigadores son de tal naturaleza (en los estudios de seguimiento, en el análisis de sobrevida) que la evaluación se puede hacer comparando valores observados con valores esperados.

Estudio _i	Tamaño del grupo	Respuesta Media	D.E
Experimental	n_{ei}	μ_{ei}	σ_{ei}
Control	n_{ci}	μ_{ci}	σ_{ci}

Cuadro 3: Variable de respuesta a escala continua (Estudios Individuales)

Estudio _i	$O - E$	Varianza de $O - E$	Grupo Exp.	Grupo Control
	Z_i	σ_i	n_{ei}	n_{ci}

Cuadro 4: Diferencia de respuestas (Estudios Individuales)

Existen también mucho estudios donde la variable de respuesta ya no es una única variable sino una combinación de esta necesaria en la evaluación de *pruebas diagnósticas*. Recordemos que en las pruebas diagnósticas para comparar y evaluar los tratamientos se deben manejar 2 índices, la *sensibilidad* y la *especificidad*, que son los que describen la validez de la prueba. Si se quieren entonces combinar los resultados de pruebas diagnósticas de diferentes estudios debemos pensar en una tabla para reunir la información

	Prueba referencia		
Prueba evaluación	Positivo	Negativo	Total
Positivo	a	b	P=a+b
negativo	c	d	N=c+d
Total	D=a+c	ND=b+d	T

Cuadro 5: Prueba diagnósticas (Estudio Individual)

- a: número de personas enfermas y con resultado positivo: Verdaderos positivos (VP).
- b: número de personas no enfermas y con resultado positivo: Falsos positivos (FP).
- c: número de personas enfermas y con resultado negativo: Falsos negativos (FN).
- d: número de personas no enfermas y con resultado negativo: Verdaderos negativos (VN).
- P: número total de personas con resultado positivo en la prueba.
- N: número total de personas con resultado negativo en la prueba.
- D: número total de personas enfermas.
- ND: número total de personas no enfermas.
- T: número total de personas en el estudio.

5. Métodos de medición en estudios individuales

Antes de presentar para cada una de las diferentes formas de medir la respuesta ,es necesario tener en cuenta que para muchos de los cálculos (tasas y logaritmos) es necesario hacer modificaciones cuando se presentan casos donde algunas de las celdas de la *tabla tetracórica* tiene frecuencia nula ; en ese caso se le agrega 0,5 lo que permite seguir con los cálculos.[4]

5.1. Peto Odds Ratio

Tomando en cuenta los totales para cada grupo, control y experimental (n_{ei}, n_{ci}) se pueden calcular los *odds ratios* para el

$$OR_{Peto,i} = \exp \left\{ \frac{Z_i}{V_i} \right\} \quad (7)$$

donde Z_i evalua la diferencia $O_i - E_i$ tal comose presentó en 4 donde $Z_i = a_i - E[a_i]$ y la esperanza es

$$E[a_i] = \frac{n_i(a_i + c_i)}{N_i} \quad (8)$$

y la varianza es

$$V_i = \frac{n_{ei}n_{ci}(a_i + c_i)(b_i + d_i)}{N_i^2(N_i - 1)} \quad (9)$$

bajo supuesto de distribución *hipergeométrica*.Para la construcción de los se usa el

$$SE \{ \ln(OR_{peto,i}) \} = \sqrt{\frac{1}{V_i}} \quad (10)$$

5.2. Odds Ratio

Generalmente el método de Peto no es el mas habitual sino que las respuestas se evalan de acuerdo al los *OR* convencionales entre 2 subpoblaciones con

$$OR_i = \frac{a_i d_i}{b_i c_i}, SE \{ \ln(OR_i) \} = \sqrt{\frac{1}{a_i} + \frac{1}{b_i} + \frac{1}{c_i} + \frac{1}{d_i}} \quad (11)$$

y su correspondiente medida del error.

5.3. Diferencias de Riesgos

Cuando o que interesa es evaluar la difrencia de riesgos tenemos

$$DR_i = \frac{a_i}{n_{ei}} - \frac{c_i}{n_c}, SE \{ DR_i \} = \sqrt{\frac{a_i b_i}{n_{ei}^3} + \frac{c_i d_i}{n_{ci}^3}} \quad (12)$$

5.4. Riesgo Relativo

El riesgo relativo de cada estudio se puede evaluar como

$$RR_i = \frac{a_i/c_{ei}}{c_i/c_{ci}}, SE \{\ln(RR_i)\} = \sqrt{\frac{1}{a_i} + \frac{1}{c_i} + \frac{1}{nc_{ei}} + \frac{1}{nc_i}} \quad (13)$$

Tanto para la estimación de la varianza para el (*OR*) como para el (*RR*) se usa la aproximación de la varianza del logaritmo de una VA con la suma de los inversos de las frecuencias que aparecen en 5.2 y 5.4.

5.5. Estimaciones Individuales para respuesta a escala continua

En este caso a partir de la información resumida como en (3) y tenemos $N_i = n_{ei} + n_{ci}$ donde evaluamos la media μ_{ei}, μ_{ci} y un *pool* de varianzas

$$s_i = \sqrt{\frac{(n_{ei} - 1)\sigma_{ei}^2 + (n_{ci} - 1)\sigma_{ci}^2}{N_i - 2}} \quad (14)$$

Para las Diferencia de medias (*mean difference*) hay que evaluar es estadístico que da cuenta de $DM_i \mu_{ei} - \mu_{ci}$ con su correspondiente medida del error a través de

$$DM_i = \sqrt{\frac{\sigma_{ei}^2}{n_{ei}} + \frac{\sigma_{ci}^2}{n_{ci}}} \quad (15)$$

Para el caso de Diferencias de medias estandarizada se puede evaluar a través de el *estadístico g de Hedges* (que se usa en REVMAN 12) similar al *estadístico d de Cohen*. El *g* incluye un ajuste para el sesgo por tamaño insuficiente de muestra

$$SDM_i = \frac{\mu_{ei} - \mu_{ci}}{s_i} \left(1 - \frac{3}{4N_i - 9}\right) \quad (16)$$

$$SE \{SDM_i\} = \sqrt{\frac{N_i}{n_{ei}n_{ci}} + \frac{SDM_i^2}{2(N_i - 3,94)}} \quad (17)$$

5.6. Curvas diagnósticas

A partir de los valores presentados en la tabla de evaluación ??se determina la *validez de las pruebas* de acuerdo a

- Sensibilidad $\frac{a}{D}$
- Especificidad $\frac{d}{ND}$
- cocientes de probabilidad $CP+ = \frac{Sen}{1-Sen}, CP- = \frac{1-Spe}{Spe}$

Estos índices para los estudios individuales en *curvas diagnósticas* permiten luego construir la curva de respuesta característica, (*ROC*), y los *cocientes de probabilidad*, evaluar la relación entre los resultados comparando la frecuencia con los de la prueba de referencia.

6. Estimación para Estudios Individuales: Método Genérico

Como complemento de los métodos de evaluación individual antes presentados puede interesar medir tasas, índices, *tiempo transcurridos hasta el evento* o los (NNT) *número necesario a tratar* que es otra forma de expresar los resultados de un estudio de tipo longitudinal o de seguimiento, donde se evalúa el número de personas que se beneficiarán dentro de cierto período de tiempo, y debiendo de tener cuidado de solo combinar ensayos con periodos de seguimiento similares [17].

7. Métodos de Medición para estudios combinados (metanálisis)

Acá debemos considerar todos los estudios que cumplen los criterios de inclusión de nuestro (MA). Una vez recolectados los estudios con algunos de los formatos antes presentados, de acuerdo a la variable de respuesta podemos usar diferentes modelos que combinan los resultados

7.1. Métodos de Mantel-Haenszel

Estos métodos son lo que se conocen como de efecto fijo que ya se presentaron en 2.1

$$OR_{MH} = \frac{\sum W_{MH,i} OR_i}{\sum W_{MH,i}} \quad (18)$$

donde cada estudio tiene un peso de

$$W_{MH,i} = \frac{b_i c_i}{N_i}, \quad SE \{ \ln (OR_{MH}) \} = \sqrt{\frac{1}{2} \left(\frac{E}{R^2} + \frac{F+G}{RS} + \frac{H}{S^2} \right)} \quad (19)$$

$$\begin{cases} R = \sum \frac{a_i d_i}{N_i}; S = \sum \frac{b_i c_i}{N_i}; E = \sum \frac{(a_i+d_i)a_i d_i}{N_i^2}; F = \sum \frac{(a_i+d_i)b_i c_i}{N_i^2} \\ G = \sum \frac{(b_i+c_i)a_i d_i}{N_i^2}; H = \sum \frac{(b_i+c_i)b_i c_i}{N_i^2} \end{cases}$$

Para evaluar el riesgo relativo combinado tenemos

$$RR_{MH} = \frac{\sum W_{MH,i} RR_i}{\sum W_{MH,i}}; W_{MH,i} = \frac{c_i(a_i + b_i)}{N_i} \quad (20)$$

En este caso el error se mide a través de

$$\begin{cases} SE \{ \ln (RR_{MH}) \} = \sqrt{\frac{P}{RS}}, \text{ con } P = \sum \frac{n_{ei}n_{ci}(a_i+c_i)-a_i c_i N_i}{N_i^2} \\ R = \sum \frac{a_i n_{ci}}{N_i}; S = \sum \frac{c_i n_{ci}}{N_i} \end{cases}$$

Si en vez de querer comparar los riesgos entre tratamientos y controles en ratios lo hacemos en diferencia tenemos una diferencia ponderada

$$DR_{MH} = \frac{\sum W_{MH,i} DR_i}{\sum W_{MH,i}}; W_{MH,i} = \frac{n_{ei}(n_{ci})}{N_i} \quad (21)$$

con un error estandar de

$$SE \{ DR_{MH} \} = \sqrt{\frac{J}{K^2}}; J = \sum \frac{a_i b_i n_{ci}^3 + c_i d_i n_{ei}^3}{n_{ei} n_{ci} N_i^2}; K = \frac{n_{ei}(n_{ci})}{N_i} \quad (22)$$

Otra forma de ponderar lo estudios , tanto los que tienen respuesta binaria,(2) como los de respuesta en escala continua (3) es a través de considerar un *tamaño del efecto* que se afecta con el inverso de la varianza (esta varianza será el cuadrado del error estandar de cada estudio en particular según sea la estimación usada.ver fórmulas (5.1,5.2,5.3,5.4,5.5). Por último resta precisar que para evaluar el ratio combinado según el método de PETO tenemos que

$$OR_{Peto} = \exp \frac{\sum V_i \ln (OR_{Peto,i})}{\sum V_i}; SE \{ \ln (OR_{Peto}) \} = \frac{1}{\sum V_i} \quad (23)$$

Con una modificación del método de Peto para la respuesta de tipo *tiempo hasta el evento* el parámetro a estimar es

$$\hat{\theta} = \exp \left\{ \frac{\sum V_i \hat{\theta}_i}{\sum V_i} \right\}; SE \{ \hat{\theta} \} = \frac{1}{\sum V_i}; \quad (24)$$

7.2. Tests de Heterogeneidad

Existen muchas formas de contrastar la *heterogeneidad* para los modelos de efecto fijo de los cuales presentamos los que aparecen en 12,??, en 13 como ser

- Test de Cochran
- Test de Wolf
- Test de Mantel-Haenszel

■ Test de Inversa de Varianza

En general los estadísticos construídos para evaluar heterogeneidad están basados en suma de cuadrados (diferencia entre cantidades esperadas y observadas) que hace generalmente se distribuyan χ^2 . El test de Cochran es $Q = \sum \frac{(O_i - E_i)^2}{V_i} - \left(\sum (O_i - E_i)^2 \right) / \sum V_i$

EL test de (MH) es $Q_{MH} = \sum w_i \left(\hat{\theta}_i - \hat{\theta}_{MH} \right)^2$ θ puede ser el logaritmo del (OR),(RR),(DR) y los w_i son los pesos calculados como $\frac{1}{SE\{\hat{\theta}_i\}^2}$ mas que los pesos usados en 7.1. $Q_{MH} \sim \chi_{k-1}^2$
 A su vez se puede evaluar un índice de *inconsistencia*

$$I^2 = \text{máx} \left\{ 100\% \frac{Q_{MH} - (k-1)}{Q_{MH}}, 0 \right\} \quad (25)$$

Este índice evalúa la incosistencia a lo largo de los estudios y se interpreta como la proporción de la variación total estimada y que se debe a la heterogeneidad mas que al *error de muestreo*.

La heterogeneidad con métodos de inversa de varianza (Inverse-variance methods) se mide en forma similar a MantelHaenszel

$$Q_{IV} = \sum w_i \left(\hat{\theta}_i - \hat{\theta}_{IV} \right)^2 \sim \chi_{k-1}^2 \quad (26)$$

$$I^2 = \text{máx} \left\{ 100\% \frac{Q_{IV} - (k-1)}{Q_{IV}}, 0 \right\} \quad (27)$$

La heterogeneidad para el caso del método de Peto es

$$Q_{Peto} = \sum V_i \left\{ \ln (OR_{Peto,i})^2 - \ln (OR_{Peto})^2 \right\} \sim \chi_{k-1}^2 \quad (28)$$

La heterogeneidad para *tiempo hasta el evento* es

$$Q_{Peto} = \sum V_i \left\{ \hat{\theta}_i^2 - \hat{\theta}^2 \right\} \sim \chi_{k-1}^2 \quad (29)$$

7.3. Modelos de Efectos Aleatorios

Para el caso de los modelos de efectos aleatorios tal como se preentaron en (2.2) vamos a presentar los mas conocidos que son los de *DerSimonian and Laird* [5] donde se supone para el *tamaño del efecto* un comportamiento aleatorio que se recoge en la siguiente distribución:

$$\theta_i \sim N(\theta, \tau^2); \hat{\tau}^2 = \text{máx} \left\{ \frac{\hat{Q} - (k-1)}{\sum w_i - (\sum w_i^2) / \sum w_i}, 0 \right\} \quad (30)$$

con

$$w_i = \frac{1}{SE\{\hat{\theta}_i\}^2} \quad (31)$$

que vale para cualquiera de los tipos de respuestas de estudios individuales (5.2,5.3,5.4,5.5).

A partir de la estimación de τ tenemos nuevos pesos para cada estudio individual

$$w'_i = \frac{1}{SE\{\hat{\theta}_i\}^2 + \hat{\tau}^2}; \quad (32)$$

$$\hat{\theta}_{DL} = \frac{\sum w'_i \hat{\theta}_i}{\sum w'_i} \quad (33)$$

$$SE\{\hat{\theta}_{DL}\} = \frac{1}{\sum w'_i} \quad (34)$$

7.4. Intervalos de Confianza

Para cualquiera de los parámetros estimados ($\hat{\theta}$) en todos los tipos de respuesta del tratamiento bajo estudio tenemos un intervalo genérico de $100(1 - \alpha) \%$

$$\left[\hat{\theta} - SE\{\hat{\theta}\} \phi(1 - \alpha/2); \hat{\theta} + SE\{\hat{\theta}\} \phi(1 - \alpha/2) \right] \quad (35)$$

7.5. Cálculo de Índices Globales para Curvas Diagnósticas

Se deben calcular índices globales que no son otra cosas que promedios ponderados de los índices individuales vistos en 5. De la misma manera podemos luego trabajar con (MEF) o (MEA) [21]

$$Sen_T = \frac{\sum a_i}{\sum D_i}; Spe_T = \frac{\sum d_i}{\sum ND_i} \quad (36)$$

$$\theta_{MH}^T = \frac{\sum W_{MH,i} \theta_i}{\sum W_{MH,i}}; \ln(\theta_{DL}^T) = \frac{\sum W_{DL,i} \ln(\theta)_i}{\sum W_{DL,i}} \quad (37)$$

De esta manera para los (ORD)(Odds Ratios Diagnósticos) y para los (CP+) y (CP-) tenemos

$$\underbrace{ORD : W_{MH,i} = \frac{b_i c_i}{T_i}; \quad CP+ : W_{MH,i} = \frac{b_i D_i}{T_i}; \quad CP- : W_{MH,i} = \frac{d_i D_i}{T_i}}_{\text{Pesos de Mantel-Haenszel}} \quad (38)$$

$$\underbrace{W_{DL,i} = \frac{1}{SE(\ln \theta_i)^2 + \tau^2}, \tau^2 = \begin{cases} \frac{Q - (k-1)}{\sum w_i - (\sum w_i^2) / \sum w_i} & Q > k - 1 \\ 0 & Q < k - 1 \end{cases}}_{\text{Pesos de DerSimonian y Laird}}$$

Para el caso de (DerSimonian- Laird) [5], Q es el estadístico de Cochran para medir la heterogeneidad y w_i pesos con el método de inversa de la varianza.

8. Herramientas Gráficas en Metanálisis

Para presentar los diferentes tipos de gráficos es importante no perder de vista los aspectos que desde el punto de vista estadístico deben de tener estos. Siguiendo principios establecidos por Cleveland (Cleveland 1994) los gráficos deben de cumplir lo siguiente [2]

- exactitud (accuracy)
- simplicidad (simplicity)
- claridad (clarity)
- apariencia (appearance)
- estructura bien definida (well defined structure)

Recordando estos aspectos es muy importante que el epidemiólogo o salubrista recurra a ellos para poder desentrañar lo que muchas veces la información no muestra detrás de conjuntos de datos numerosos. Sin embargo se debe apoyar en los gráficos para luego verificar mediante tests estadísticos adecuados lo que la evidencia empírica parece mostrar.

Muchos de los gráficos que se presentan a continuación fueron generados con las diferentes funciones de las librerías de R (13) pero se pueden hacer con muchos de los programas reseñados en (11). Para cada gráfico presentado se presenta la referencia de la función con la que fue generado.

8.1. Gráfico Simple (Forest plot)

Este tipo de gráfico presenta la estimación del efecto de la intervención para cada estudio individual, donde puede aparecer las diferentes medidas presentadas en (5). A su vez cada valor estimado se acompaña de un intervalo de confianza y obviamente puede tener incluido el metanálisis, el que aparece generalmente en la parte inferior del gráfico. Las medidas de los efectos del tratamiento se pueden presentar en escala logarítmica, pudiendo superponer a su vez una línea punteada de otro color para indicar claramente el efecto global (*pooled effect*) estimado. y a su vez el punto que representa el valor de este efecto global con otro formato, como ser un rombo invertido, de otro color. Presentamos un ejemplo

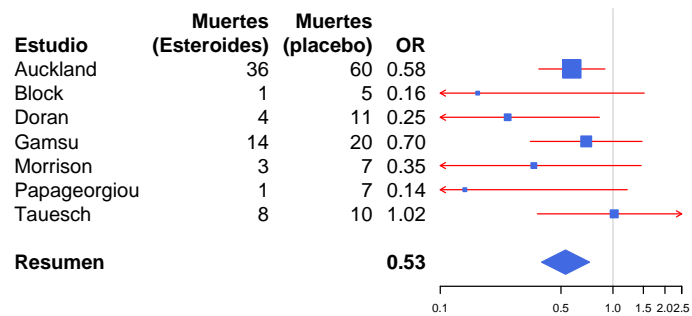


Figura 1: Funnel plot para datos(catheter) MEF (MH)

8.2. Gráfico para análisis acumulativo

Podemos ir viendo como queda graficamente cuando vamos integrando en forma sistemática un estudio por vez haciendo un gráfico acumulativo.

```
{
> head(Olkin95)
      author year event.e n.e event.c n.c
1   Fletcher 1959      1  12      4  11
2     Dewar  1963      4  21      7  21
3 Lippschutz 1965      6  43      7  41
4 European 1 1969     20  83     15  84
5 European 2 1971     69 373     94 357
6 Heikinheimo 1971     22 219     17 207

metal <- metabin(event.e, n.e, event.c, n.c,
                 data=Olkin95, subset=c(41,47,51,59),
```

8.2 Gráfico para análisis acumulativo

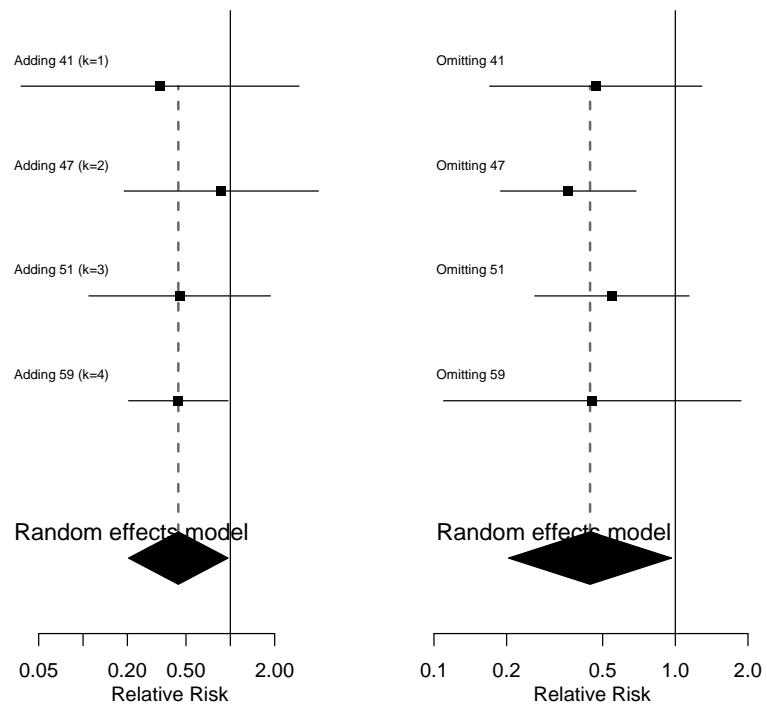


Figura 2: Gráfico acumulativo

8.3 Gráfico Radial (Galbraith plot)

```
sm="RR", meth="I")
> metacum(metal1, pooled="random")

Cumulative meta-analysis (Random effects model)

              RR              95%-CI  p.value      I2
Adding 41 (k=1)  0.3333 [0.0380; 2.9253]  0.3215
Adding 47 (k=2)  0.8685 [0.1903; 3.9635]  0.8556      23.0%
Adding 51 (k=3)  0.4528 [0.1096; 1.8699]  0.2735      50.0%
Adding 59 (k=4)  0.4434 [0.2038; 0.9648]  0.0403      25.1%

Pooled estimate  0.4434 [0.2038; 0.9648]  0.0403      25.1%

Method: Inverse variance method
>
}
```

Se puede ver como la respuesta global va cambiando a medida que se van teniendo más estudios en el metanálisis y a su vez como cambia al ir sistemáticamente quitando un estudio en particular

8.3. Gráfico Radial (Galbraith plot)

En el gráfico de Galbraith Galbraith [8] o *gráfico radial* podemos ver la relación que existe entre *la precisión de cada estudio* (el inverso del error estándar) y el *efecto estandarizado*; Ambas magnitudes se representan en un par de ejes coordinados donde se puede a su vez evaluar la línea de regresión ajustada con esos puntos ;y se puede construir un intervalo de confianza. Los estudios (los puntos) que caen fuera de esta banda son los que más contribuyen a la heterogeneidad. Por otra parte, la posición de los estudios en el eje de abscisas permite identificar visualmente aquellos que tienen un mayor peso en el meta-análisis. En este tipo de representación podemos aprovechar la potencia de este tipo de gráficos de 2 variables donde los puntos se pueden colorear y representar con tamaño variable, de acuerdo por ejemplo tamaño del estudio y año de publicación. Se puede detectar también puntos (estudios) *anómalos* o *outliers*

8.4. Gráfico De Embudo(Funnel plot)

Los gráficos de embudo, introducidos por Light and Pillemer Light [12] y luego discutido en detalle por Egger Egger~M [6] son gráficos con un par de ejes coordinados donde

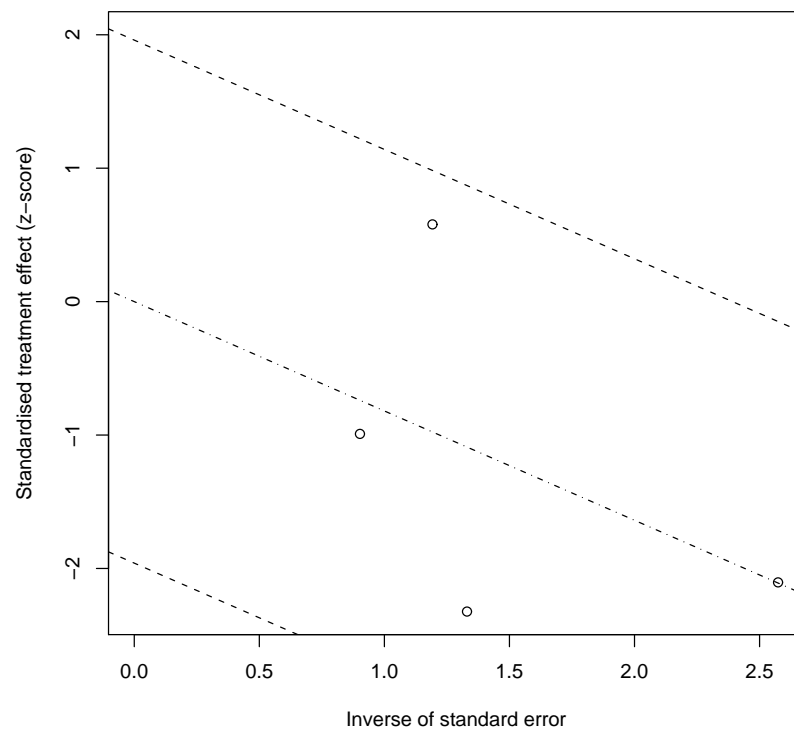


Figura 3: Gráfico de Galbraith (radial) para datos(Olkin95)

8.4 Gráfico De Embudo(Funnel plot)

se ve la relación que existe entre el efecto de la intervención estimado contra la medida del tamaño del estudio (existiendo varias formas de medirlo, tal como se presenta en (7), lo que permite evaluar sesgos, heterogeneidad sistemática, en la que la nube de puntos debiera de tener el formato de un embudo invertido bajo el supuesto de que no existe sesgo de publicación, ya que una asimetría en esa *nube* de puntos pone de manifiesto una asociación entre el efecto de la intervención y el tamaño del estudio, situación que el investigador desea descartar. Frente a esa situación que es un primer llamado de atención se debe de tener cuidado al tratar de hacer el análisis acumulativo.

```
smoking
```

```
      qt  tt qc  tc
Blondal89   37  92 24  90
Campbell91  21 107 21 105
Fagerstrom82 30  50 23  50
Fee82       23 180 15 172
Garcia89    21  68  5  38
Garvey00    75 405 17 203
```

```
> summary(smokingDSL)
```

```
Random effects ( DerSimonian-Laird ) meta-analysis
```

```
Call: meta.DSL(ntrt = smoking[["tt"]], nctrl = smoking[["tc"]], ptrt = smoking[["qt"],
      pctrl = smoking[["qc"]], names = rownames(smoking))
```

```
-----
              OR (lower 95% upper)
Blondal89    1.85    0.99    3.46
Campbell91   0.98    0.50    1.92
Fagerstrom82 1.76    0.80    3.89
Fee82        1.53    0.77    3.05
Garcia89     2.95    1.01    8.62
Garvey00     2.49    1.43    4.34
Gross95      2.62    1.03    6.71
.....
Niaura94     1.34    0.35    5.19
Pirie92      1.84    1.20    2.82
Puska79      1.46    0.78    2.75
Schneider85  1.71    0.52    5.62
Tonnesen88   2.12    0.93    4.86
Villa99      1.76    0.55    5.64
Zelman92     1.46    0.68    3.14
```

 SummaryOR= 1.75 95% CI (1.48,2.07)
 Test for heterogeneity: $X^2(25) = 34.87$ (p-value 0.0905)
 Estimated random effects variance: 0.05

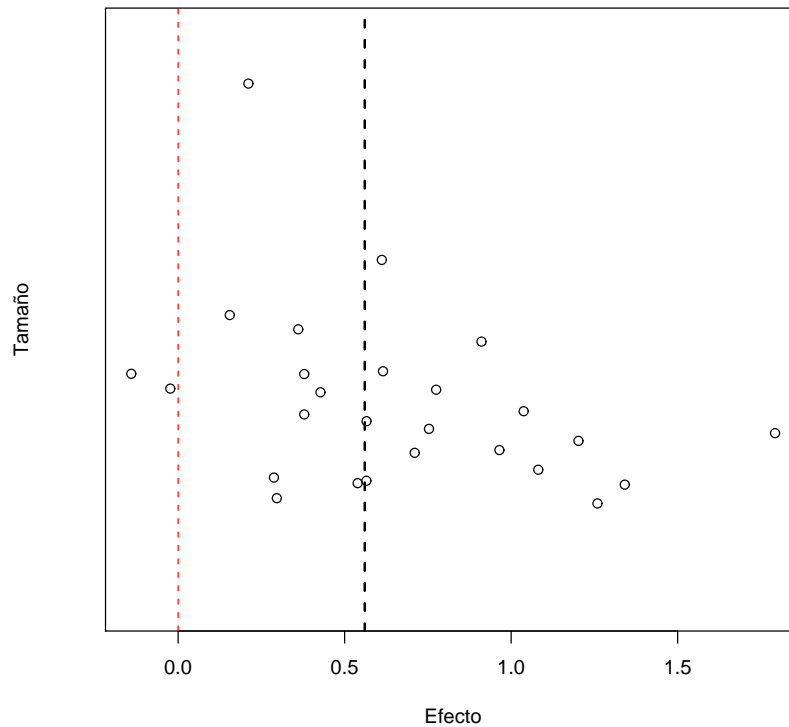


Figura 4: Gráfico de embudo (Funnel plot) para (MEA) con datos=smoking

8.5. Gráfico de L'Abbée

El gráfico de L'Abbée [11] representa la tasa de eventos del grupo tratamiento frente a la tasa del grupo control, por tanto, este gráfico solo se presenta cuando la variable respuesta es binaria. En este tipo de gráfico se puede ver la heterogeneidad o la dependencia que puede existir entre la relación de la tasa en ambos.

8.6. Gráficos de Sesgos

(librerías de R) Para evaluar graficamente los diferentes tipos de sesgos se pueden combinar algunos de los ya vistos manejando las respuestas (efectos de los tratamientos), con precisión de los estudios. Si los sesgos de publicación se quieren modelar para

luego modificar los (MA), con las herramientas que aparecen en (9) se pueden ver los resultados siguientes

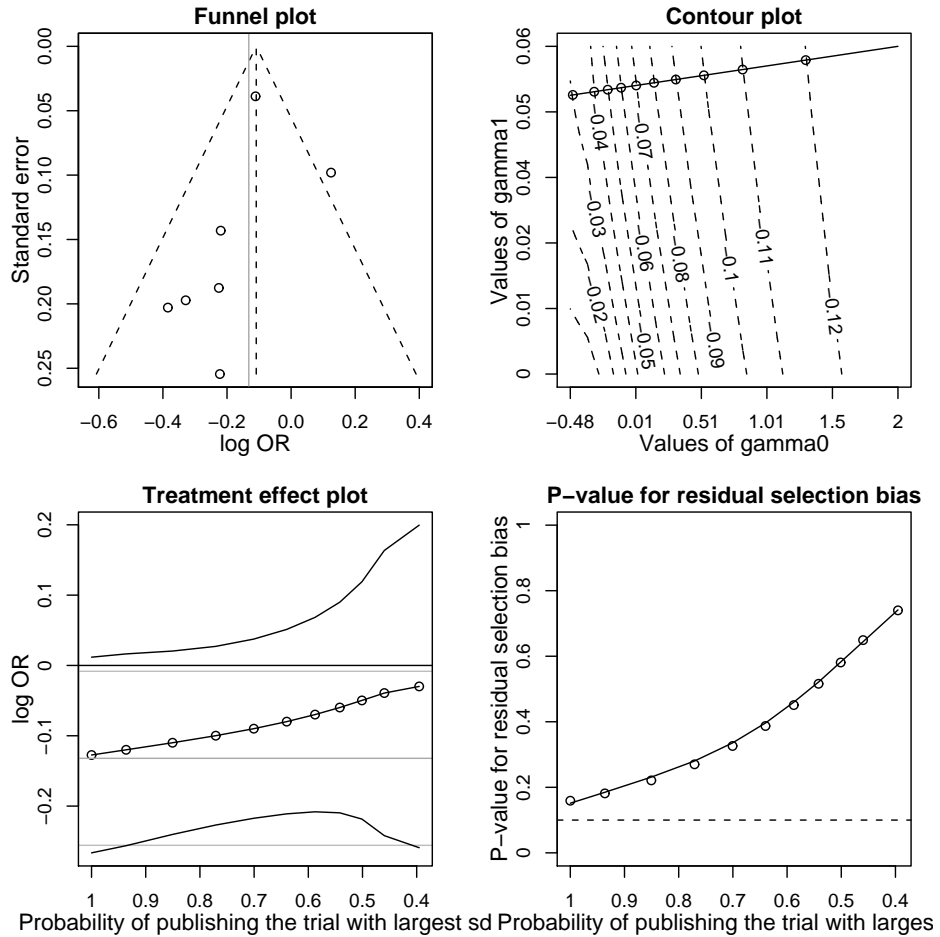


Figura 5: Sesgo de Publicación(MOdelo librería copas())de R)

8.7. Tests de Heterogeneidad en Curvas Diagnósticas

Tal como se presentaron para los otras tipos de variable de respuesta (2,3) se debe analizar una fuente **extra** de variabilidad entre estudios de (CD), pensando en que en los estudios incluidos se hayan manejados diferentes *umbrales* para la definición de casos positivos/negativos.

Para eso se pueden hacer *pruebas de homogeneidad* de sensibilidad y especificidad usando *test de razón de verosimilitud* con distribución χ^2 . Para el resto de los parámetros (CP+),(CP-),(ROD) nos podemos manejar con la prueba de Cochran ya vista en secciones precedentes. Si existe un *efecto umbral* este puede evaluarse mediante coeficiente de correlación de Spearman y se debe de construir una *curva Roc* para evaluar la (Sen) y la (Spe) al no poder construir índices globales Walter [20], Mitchell [14], Unidad de Bioes-

tadística Clínica del Hospital Ramón y Cajal [19].

9. Tratamiento del Sesgo en metanálisis

Ya se vieron en ?? diferentes tipos de *sesgo* y vamos a presentar como identificarlos y como tratarlos mediante modelos estadísticos.

- Sesgo de publicación τ de Kendall (pruebas de Begg basada en correlación de rangos entre tratamientos y varianza estimada del efecto del tratamiento; Test de Egger (es una regresión lineal ponderada del efecto del tratamiento contra el ES)) están implementadas en ??
 - Método 'rank', basado en τ de Kendall, test estadístico que sigue distribución normal estandar (Kendall Gibbons, 1990).13.2
 - Método 'linreg', basado en regresión lineal ponderada del efecto del tratamiento contra el ES (Egger et al., 1997). The test estadístico que sigue distribution t con número de estudios - 2 grados de libertad.13.2
 - Método 'mm', test basado en regresión lineal ponderada del efecto del tratamiento contra el ES estimado a través del método de los momentos (Thompson, Sharp, 1999). The test estadístico que sigue distribution t con número de estudios - 2 grados de libertad.13.2
 - Método 'count', basada en correlación de rangos entre frecuencia de celdas de la tabla e inversa de la varianza de la frecuencia; (Schwarzer et al., 2007).13.2
 - Método 'peters', basado en regresión lineal ponderada del efecto del tratamiento contra el inverso del tamaño de muestra usando la varianza de la tasa promedio de eventos (Peters et al., 2006).13.2

En particular y sin ser estos los únicos vamos a interesarnos en los sesgos de selección-publicación. Nos referimos a este tipo como *modelo de selección de COPAS* que está implementado en la librería *COPAS* (ver 13) que consiste en la aplicación de 2 modelos que se estiman en forma conjunta

- Modelos de efectos aleatorios (MEA) ya presentado en 2.2
- Modelo de selección donde un modelo i es seleccionado para ser publicado si $Z > 0$, donde

$$Z = \gamma_0 + \frac{\gamma_1}{SE(i)} + \delta_i \quad (39)$$

En el modelo 39 el término γ_0 es la constante del *modelo Probit* de selección del estudio-*i-ésimo*. **Así la distribución acumulativa normal de γ_0 es aproximadamente la probabilidad de que un estudio pequeño sea publicado**. Si tenemos valores altos hay mas probabilidad de que ese tipo de estudio sea publicado. En el modelo el parámetro γ_1 es el *coeficiente de precisión* en el *modelo probit de selección*. **Ese parámetro modula la tasa a la que la probabilidad de publicación se incrementa a medida que el error estandar del efecto del tratamiento reporta un decrecimiento**. Por último el error δ_1 está correlacionado con el error del (MEA) con correlación ρ_0 . **Si el $\rho_0 = 0$, entonces esta componente se anula y entonces equivale a estar usando el (MEA)**. **A medida que ρ_0 se mueve desde 0 a 1, un estudio con un efecto de tratamiento estimado tiene mas probabilidad de ser selecciondo/publicado**. Estos modelos que está implementados en la librería *COPAS* que permite a su vez evaluar graficamente y comienza dando valores a *gamma₀* y *gamma₁* usando los valores del estudio mas pequeño. A su vez también se puede evaluar el *número de estudios sin publicar* que el modelo estimado sugiere Copas [3].

10. Metaregresión

Muchas veces puede ser necesario al combinar resultados tener en cuenta no solamente los efectos de los tratamientos (sin importar en que escala esten medidos, sino otra covariables que pueden impactar en esos resultados, que pueden ir desde diferentes dosificaciones, duraciones de la exposición al tratamiento, etc). Primero se debe tratar de inspeccionar esa situación visualmente a través de un gráfico bivariado donde se vea si existe un patrón de comportamiento de la respuesta, asociada a algunas de esas covariables, pudiendo luego eso corroborarlo mediante modelos de regresión lineal o no lineal entre la variable de respuesta y las covariables. Everitt [7]

Por ejemplo en el caso de un (MA) sobre estudios en vacunas de BCG y tratamiento contra la tuberculosis, se ve que existe una asociación entre el efecto y la latitud, de donde se realiza el estudio.

```
summary(BCG_mod)
Call:
lm(formula = y ~ Latitude + Year, data = BCG, weights = studyweights)
Residuals:
    Min       1Q   Median       3Q      Max
-1.66012 -0.36910 -0.02937  0.31565  1.26040
Coefficients:
                Estimate Std. Error t value Pr(>|t|)
```

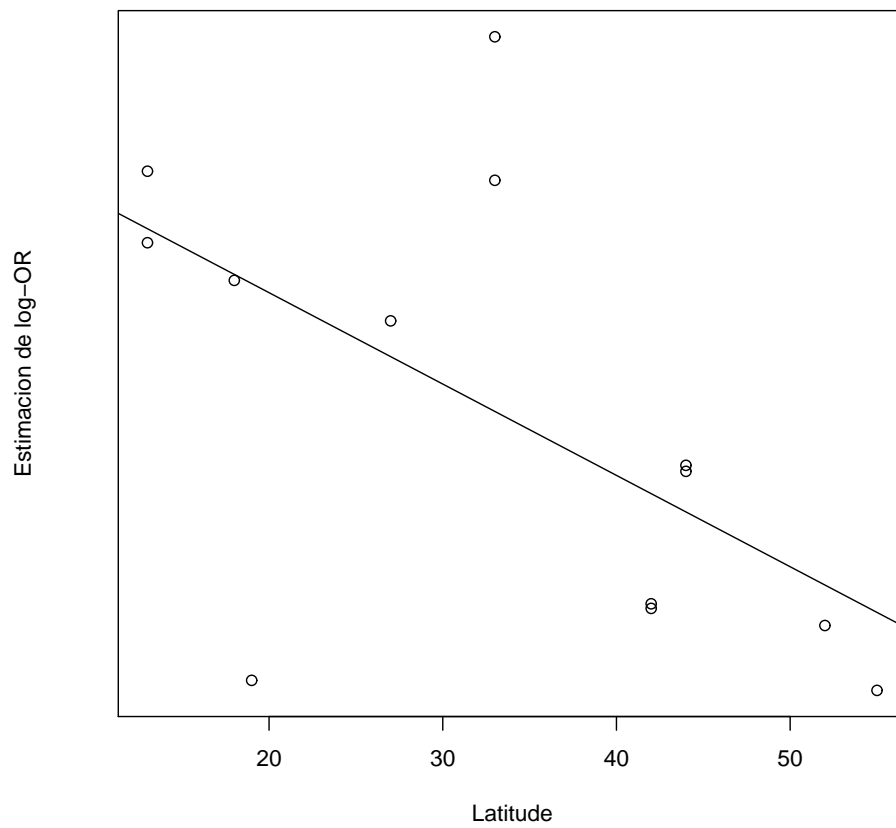


Figura 6: Ejemplo de asociación entre tratamiento y covariables

Revisión de diferentes programas de Computación disponibles para el metanálisis

```
(Intercept) -16.199115  37.605403  -0.431  0.6758
Latitude     -0.025808   0.013680  -1.887  0.0886 .
Year         0.008279   0.018972   0.436  0.6718
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 0.7992 on 10 degrees of freedom
Multiple R-squared:  0.4387,    Adjusted R-squared:  0.3265
F-statistic: 3.909 on 2 and 10 DF,  p-value: 0.05569
>
}
```

11. Revisión de diferentes programas de Computación disponibles para el metanálisis

Se presenta una nómina actualizada al día de hoy de varias herramientas de dominio público o comerciales que permiten evaluar varias de los métodos y efectuar gráficos (estos se presentan en ??).

12. Revman

Es un software desarrollado por la COCHRANE Collaboration para preparar y mantener *revisiones COCHRANE* , haciendo la preparación de *protocolos* y revisiones completas, que luego pueden ser analizadas mediante metanálisis, pudiendo presentar los resultados en forma de reportes y a su vez graficamente.

Actualmente está en la *versión 5* que es multiplataforma (al estar desarrollada en JAVA) y permite evaluar sesgos de inclusión (con una nueva herramienta *Risk of bias tool*, generación de reportes con *Resumen de hallazgos*(Summary of findings) que se pueden manejar con un herramienta complementaria *GRADE approach*. Maneja respuestas de tratamiento en escala continua, escala dicotómica ,Odds ratio,Riesgo relativo,Diferencia de riesgos,Diferencia estandarizada de medias y estima modelos de efectos fijos y aleatorios.Permite Ponderación de los resultados por el método de *inversa de varianza* y tests de heterogeneidad.Collaboration [2],Deeks [4]

13. Librerías en R

Al momento existen 4 librerías en \mathfrak{R} que tienen implementadas enre todas ellas los aspectos vistos en ?? y en 7,7.3 y 8.7 Se pueden leer datos creados con otros programas como ser REVMAN

```
Localiza archivo "FLEISS93.MTV" in subdirectorio de paquete "meta"
filename <- paste(searchpaths()[seq(along=search())[search()==
      "package:meta"]], "/data/FLEISS93.MTV", sep="")
fleiss93.cc <- read.mtv(filename)
```

```
%example(Fleiss93)
```

13.1. librería *rmeta()*

Lumley [13] Acá tenemos como queda un modelo de efectos fijos y de efectos aleatorios sobre los datos *catheter*

```
summary(a)
Fixed effects ( Mantel-Haenszel ) meta-analysis
Call: meta.MH(ntrt = n.trt, nctrl = n.ctrl, ptrt = col.trt, pctrl = col.ctrl,
  names = Name, data = catheter, subset = c(13, 6, 5, 3, 7,
    12, 4, 11, 1, 8, 10, 2))
```

```
-----
                OR (lower 95% upper)
Tennenberg 0.22    0.10    0.49
Maki        0.49    0.29    0.82
vanHeerden 0.27    0.07    1.00
Hannan     0.83    0.40    1.72
Bach(a)    0.00    0.00    NaN
Bach(b)    0.11    0.02    0.49
Heard      0.60    0.38    0.95
Collins    0.10    0.02    0.41
Ciresi     0.69    0.34    1.42
Ramsay     0.58    0.37    0.92
Trazzera   0.47    0.23    0.94
George     0.12    0.04    0.33
-----
```

```
Mantel-Haenszel OR =0.44 95% CI ( 0.36,0.54 )
```

```
Test for heterogeneity:  $X^2(10) = 25.36$  ( p-value 0.0047 )
```

13.1 librería rmeta()

```
summary(b)
Random effects ( DerSimonian-Laird ) meta-analysis
Call: meta.DSL(ntrt = n.trt, nctrl = n.ctrl,
prrt = col.trt, pctrl = col.ctrl,
  names = Name, data = catheter, subset = c(13, 6, 5, 3, 12,
  4, 11, 1, 8, 10, 2))
```

```
-----
              OR (lower 95% upper)
Tennenberg 0.22   0.10   0.49
Maki        0.49   0.29   0.82
vanHeerden 0.27   0.07   1.00
Hannan     0.83   0.40   1.72
Bach(b)    0.11   0.02   0.49
Heard      0.60   0.38   0.95
Collins    0.10   0.02   0.41
Ciresi     0.69   0.34   1.42
Ramsay     0.58   0.37   0.92
Trazzera   0.47   0.23   0.94
George     0.12   0.04   0.33
-----
```

```
SummaryOR= 0.4 95% CI ( 0.28,0.57 )
Test for heterogeneity:  $X^2(10) = 24.9$  ( p-value 0.0055 )
Estimated random effects variance: 0.19
```

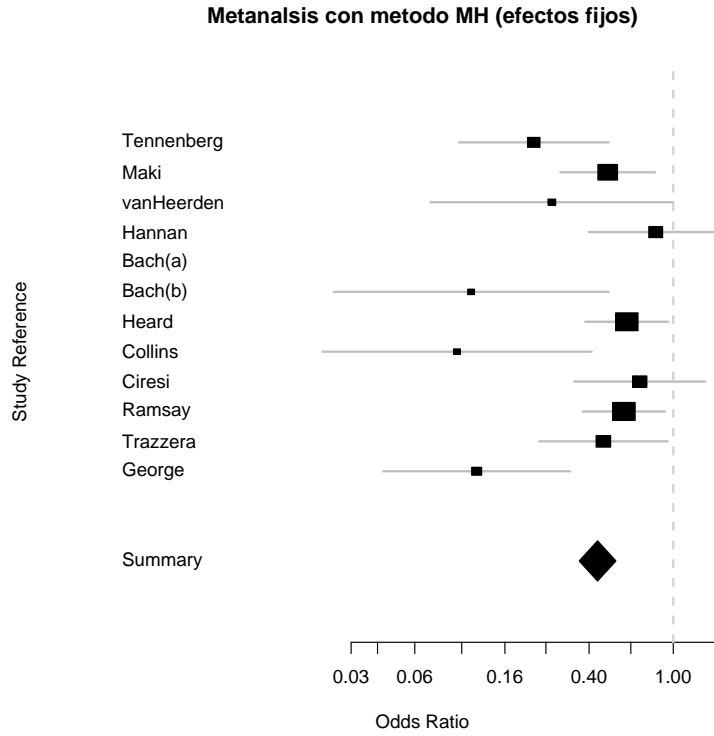


Figura 7: Modelos de Efectos fijos

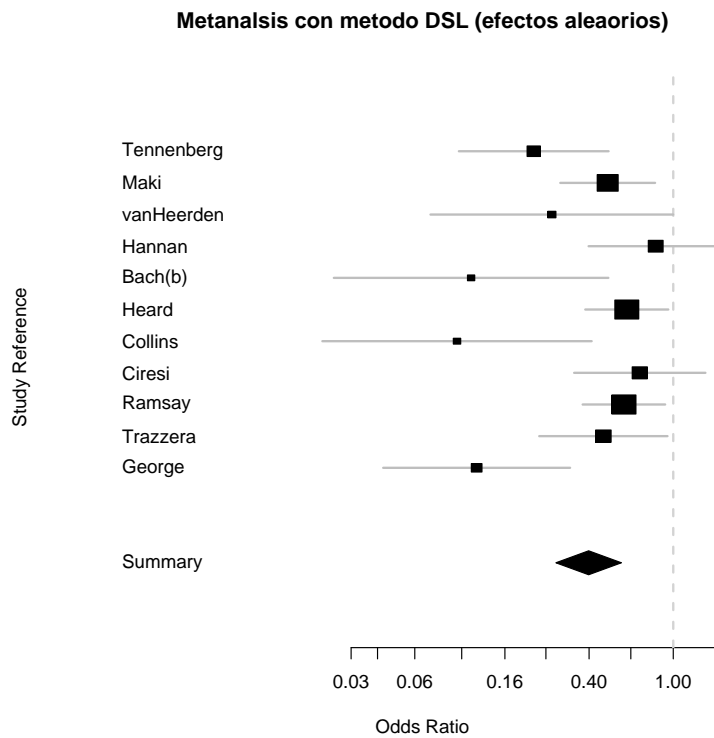


Figura 8: Modelos de Efectos aleatorios

13.1 librería rmeta()

Otro ejemplo sobre un estudio sobre uso de goma de nicotina para deja de fumar

```
smokingOR<-meta.MH(smoking[["tt"]],smoking[["tc"]],smoking[["qt"]],
smoking[["qc"]],names=rownames(smoking))
summary(smokingOR)
Fixed effects ( Mantel-Haenszel ) meta-analysis
Call: meta.MH(ntrt = smoking[["tt"]],
nctrl = smoking[["tc"]], ptrt = smoking[["qt"]],
pctrl = smoking[["qc"]], names = rownames(smoking))
```

```
-----
              OR (lower 95% upper)
Blondal89    1.85    0.99    3.46
Campbell91   0.98    0.50    1.92
Fagerstrom82 1.76    0.80    3.89
Fee82        1.53    0.77    3.05
Garcia89     2.95    1.01    8.62
Garvey00     2.49    1.43    4.34
Gross95      2.62    1.03    6.71
Hall85       2.03    0.78    5.29
Hall87       2.82    1.33    5.99
Hall96       0.87    0.46    1.64
Hjalmarson84 2.17    1.10    4.28
Huber88      6.00    2.57   14.01
Jarvis82     3.33    1.37    8.08
Jensen91     1.43    0.84    2.44
Killen84     1.33    0.43    4.15
Killen90     1.23    0.93    1.64
Malcolm80    3.52    0.85   14.54
McGovern92   1.17    0.70    1.94
Nakamura90   3.82    1.15   12.71
Niaura94     1.34    0.35    5.19
Pirie92      1.84    1.20    2.82
Puska79      1.46    0.78    2.75
Schneider85  1.71    0.52    5.62
Tonnesen88   2.12    0.93    4.86
Villa99      1.76    0.55    5.64
Zelman92     1.46    0.68    3.14
-----
```

13.1 librería rmeta()

```
Mantel-Haenszel OR =1.67 95% CI ( 1.47,1.9 )
Test for heterogeneity: X^2( 25 ) = 34.9 ( p-value 0.09 )
>
```

```
smokingDSL<-meta.DSL(smoking[["tt"]],smoking[["tc"]],
smoking[["qt"]],smoking[["qc"]],names=rownames(smoking))
summary(smokingDSL)
Random effects ( DerSimonian-Laird ) meta-analysis
Call: meta.DSL(ntrt = smoking[["tt"]],
  nctrl = smoking[["tc"]], ptrt = smoking[["qt"]],
  pctrl = smoking[["qc"]], names = rownames(smoking))
```

```
-----
              OR (lower 95% upper)
Blondal89    1.85    0.99    3.46
Campbell91   0.98    0.50    1.92
Fagerstrom82 1.76    0.80    3.89
Fee82        1.53    0.77    3.05
Garcia89     2.95    1.01    8.62
Garvey00     2.49    1.43    4.34
Gross95      2.62    1.03    6.71
Hall85       2.03    0.78    5.29
Hall87       2.82    1.33    5.99
Hall96       0.87    0.46    1.64
Hjalmarson84 2.17    1.10    4.28
Huber88      6.00    2.57   14.01
Jarvis82     3.33    1.37    8.08
Jensen91     1.43    0.84    2.44
Killen84     1.33    0.43    4.15
Killen90     1.23    0.93    1.64
Malcolm80    3.52    0.85   14.54
McGovern92   1.17    0.70    1.94
Nakamura90   3.82    1.15   12.71
Niaura94     1.34    0.35    5.19
Pirie92      1.84    1.20    2.82
Puska79      1.46    0.78    2.75
Schneider85  1.71    0.52    5.62
Tonnesen88   2.12    0.93    4.86
```


13.1 librería rmeta()

```
Villa99      1.76    0.55      5.64
Zelman92    1.46    0.68      3.14
-----
SummaryOR= 1.75  95% CI ( 1.48,2.07 )
Test for heterogeneity: X^2( 25 ) = 34.87 ( p-value 0.0905 )
Estimated random effects variance: 0.05
>

BCG_OR<-meta.MH(BCG[[3]],BCG[[5]],BCG[[2]],BCG[[4]],names=BCG$Study)
summary(BCG_OR)
Fixed effects ( Mantel-Haenszel ) meta-analysis
Call: meta.MH(ntrt = BCG[[3]],
nctrl = BCG[[5]], ptrt = BCG[[2]], pctrl = BCG[[4]],
names = BCG$Study)
-----
      OR (lower 95% upper)
1  0.39    0.12    1.26
2  0.19    0.08    0.46
3  0.25    0.07    0.91
4  0.23    0.18    0.31
5  0.80    0.51    1.26
6  0.38    0.32    0.47
7  0.20    0.08    0.50
8  1.01    0.89    1.15
9  0.62    0.39    1.00
10 0.25    0.14    0.42
11 0.71    0.57    0.89
12 1.56    0.37    6.55
13 0.98    0.58    1.66
-----
Mantel-Haenszel OR =0.62 95% CI ( 0.57,0.68 )
Test for heterogeneity: X^2( 12 ) = 163.94 ( p-value 0 )
>
}

BCG_DSL<-meta.DSL(BCG[[3]],BCG[[5]],BCG[[2]],BCG[[4]],names=BCG$Study)
summary(BCG_DSL)
Random effects ( DerSimonian-Laird ) meta-analysis
Call: meta.DSL(ntrt = BCG[[3]], nctrl = BCG[[5]], ptrt = BCG[[2]],
```

13.2 librería meta()

```
pctrl = BCG[[4]], names = BCG$Study)
-----
      OR (lower 95% upper)
1  0.39   0.12   1.26
2  0.19   0.08   0.46
3  0.25   0.07   0.91
4  0.23   0.18   0.31
5  0.80   0.51   1.26
6  0.38   0.32   0.47
7  0.20   0.08   0.50
8  1.01   0.89   1.15
9  0.62   0.39   1.00
10 0.25   0.14   0.42
11 0.71   0.57   0.89
12 1.56   0.37   6.55
13 0.98   0.58   1.66
-----
SummaryOR= 0.47 95% CI ( 0.32,0.69 )
Test for heterogeneity: X^2( 12 ) = 163.16 ( p-value 0 )
Estimated random effects variance: 0.37
>
```

13.2. librería meta()

Schwarzer [16] La librería *meta* permite hacer metanálisis de datos binarios, continuos, análisis acumulado y gráfico de influencia

```
metainf(meta1)
Influential analysis (Fixed effect model)
```

	RR	95%-CI	p.value	I ²
Omitting 41	0.4510	[0.2413; 0.8431]	0.0126	49.2%
Omitting 47	0.3603	[0.1889; 0.6873]	0.0019	0.0%
Omitting 51	0.5303	[0.2745; 1.0244]	0.059	8.3%
Omitting 59	0.4393	[0.1650; 1.1697]	0.0997	50.0%
Pooled estimate	0.4407	[0.2416; 0.8039]	0.0075	25.1%

13.2 librería meta()

```
Method: Inverse variance method
>
}

meta1.copas <- metabin(event.e, n.e, event.c, n.c, data=Fleiss93, sm="OR")
summary(meta1.copas) summary(meta1.copas)
Number of trials combined: 7
              OR           95%-CI           z  p.value
Fixed effects model  0.8969  [0.8405; 0.9570] -3.2876  0.001
Random effects model 0.8763  [0.7743; 0.9917] -2.0918  0.0365

Quantifying heterogeneity:
tau^2 = 0.0096; H = 1.29 [1; 1.99]; I^2 = 39.7% [0%; 74.6%]

Test of heterogeneity:
      Q  d.f.  p.value
9.95   6    0.1269

Method: Mantel-Haenszel method

Un análisis acumulado

summary(d)
Cumulative meta-analysis
Call: cummeta.summaries(effects = b$logos, stderrs = b$selogs, names = b$names,
      method = "random", logscale = TRUE)
-----
              Effect (lower 95% upper)
Tennenberg  0.22   0.10   0.49
Maki        0.35   0.16   0.76
vanHeerden  0.35   0.20   0.61
Hannan      0.43   0.24   0.75
Bach(b)     0.36   0.20   0.66
Heard       0.42   0.27   0.67
Collins     0.36   0.22   0.60
Ciresi      0.41   0.26   0.63
Ramsay      0.44   0.31   0.64
Trazzera    0.45   0.33   0.62
```

13.2 librería meta()

George 0.40 0.28 0.57

summary(meta1)

Number of trials combined: 4

	RR	95%-CI	z	p.value
Fixed effects model	0.4407	[0.2416; 0.8039]	-2.6716	0.0075
Random effects model	0.4434	[0.2038; 0.9648]	-2.0503	0.0403

Quantifying heterogeneity:

tau² = 0.1685; H = 1.16 [1; 1.86]; I² = 25.1% [0%; 71.1%]

Test of heterogeneity:

Q	d.f.	p.value
4	3	0.2611

Method: Inverse variance method

meta1

	RR	95%-CI	%W(fixed)	%W(random)
41	0.3333	[0.0380; 2.9253]	7.66	11.27
47	1.6250	[0.3140; 8.4103]	13.37	18.04
51	0.1745	[0.0400; 0.7613]	16.65	21.45
59	0.4415	[0.2062; 0.9455]	62.32	49.25

Number of trials combined: 4

	RR	95%-CI	z	p.value
Fixed effects model	0.4407	[0.2416; 0.8039]	-2.6716	0.0075
Random effects model	0.4434	[0.2038; 0.9648]	-2.0503	0.0403

Quantifying heterogeneity:

tau² = 0.1685; H = 1.16 [1; 1.86]; I² = 25.1% [0%; 71.1%]

Test of heterogeneity:

Q	d.f.	p.value
4	3	0.2611

13.3 librería copas()

Method: Inverse variance method

13.3. librería *copas()*

Schwarzer [15]

```
> meta1.copas <- metabin(event.e, n.e, event.c, n.c, data=Fleiss93, sm="OR")
> summary(meta1.copas)
Number of trials combined: 7
```

	OR	95%-CI	z	p.value
Fixed effects model	0.8969	[0.8405; 0.9570]	-3.2876	0.001
Random effects model	0.8763	[0.7743; 0.9917]	-2.0918	0.0365

Quantifying heterogeneity:

$\tau^2 = 0.0096$; $H = 1.29$ [1; 1.99]; $I^2 = 39.7\%$ [0%; 74.6%]

Test of heterogeneity:

Q	d.f.	p.value
9.95	6	0.1269

Method: Mantel-Haenszel method

```
cop1 <- copas(meta1.copas)
summary(cop1)
```

Summary of Copas selection model analysis:

	publprob	OR	95%-CI	pval.treat	pval.rsb	N.unpubl
	1.00	0.8804	[0.7661; 1.0118]	0.0728	0.1591	0
	0.94	0.8870	[0.7739; 1.0166]	0.0847	0.1814	0
	0.85	0.8959	[0.7864; 1.0207]	0.0985	0.2211	1
	0.77	0.9049	[0.7970; 1.0275]	0.1232	0.2701	2
	0.70	0.9140	[0.8047; 1.0382]	0.1666	0.3261	2
	0.64	0.9232	[0.8097; 1.0526]	0.2325	0.3871	3
	0.59	0.9325	[0.8121; 1.0708]	0.3218	0.4508	4
	0.54	0.9419	[0.8109; 1.0941]	0.4338	0.5159	4
	0.50	0.9515	[0.8038; 1.1264]	0.5635	0.5809	5
	0.46	0.9615	[0.7850; 1.1777]	0.7044	0.6492	6
	0.40	0.9706	[0.7717; 1.2208]	0.7988	0.74	8
Copas model (adj)	0.8804	[0.7661; 1.0118]	0.0728	0.1591	0	

13.3 librería copas()

Random effects model 0.8763 [0.7743; 0.9917] 0.0365

Legend:

publprob - Probability of publishing the study with the largest standard error

pval.treat - P-value for hypothesis that the treatment effect is equal in both groups

pval.rsb - P-value for hypothesis that no further selection remains unexplained

N.unpubl - Approximate number of studies the model suggests remain unpublished

=====

===== COPAS ANALYSIS =====

=====

1) Summary statistics and test for heterogeneity

=====

Number of trials combined: 7

	OR	95%-CI	z	p.value
Fixed effects model	0.8969	[0.8405; 0.9570]	-3.2876	0.001
Random effects model	0.8763	[0.7743; 0.9917]	-2.0918	0.0365

Quantifying heterogeneity:

$\tau^2 = 0.0096$; $H = 1.29$ [1; 1.99]; $I^2 = 39.7\%$ [0%; 74.6%]

Test of heterogeneity:

Q	d.f.	p.value
9.95	6	0.1269

Method: Mantel-Haenszel method

2) Ranges of gamma0, gamma1 used for calculating contour plot

=====

gamma0 ranges from -0.5 to 2.1

gamma1 ranges from 0 to 0.08

Range of probability publishing trial with largest SE: (0.309, 0.992)

Range of probability publishing trial with smallest SE: (0.309, 1)

> summary(cop2)

Summary of Copas selection model analysis:

13.3 librería copas()

publprob	OR	95%-CI	pval.treat	pval.rsb	N.unpubl
1.00	0.8804	[0.7661; 1.0118]	0.0728	0.1591	0
0.94	0.8870	[0.7738; 1.0167]	0.0851	0.1821	0
0.77	0.9049	[0.7980; 1.0263]	0.1198	0.2796	1
0.69	0.9140	[0.8059; 1.0366]	0.1614	0.3429	2
0.58	0.9325	[0.8139; 1.0682]	0.3134	0.4849	4
0.50	0.9509	[0.8054; 1.1227]	0.5525	0.6266	5
Copas model (adj) 0.8804 [0.7661; 1.0118] 0.0728 0.1591 0					
Random effects model 0.8763 [0.7743; 0.9917] 0.0365					

Legend:

publprob - Probability of publishing the study with the largest standard error
pval.treat - P-value for hypothesis that the treatment effect is equal in both groups
pval.rsb - P-value for hypothesis that no further selection remains unexplained
N.unpubl - Approximate number of studies the model suggests remain unpublished

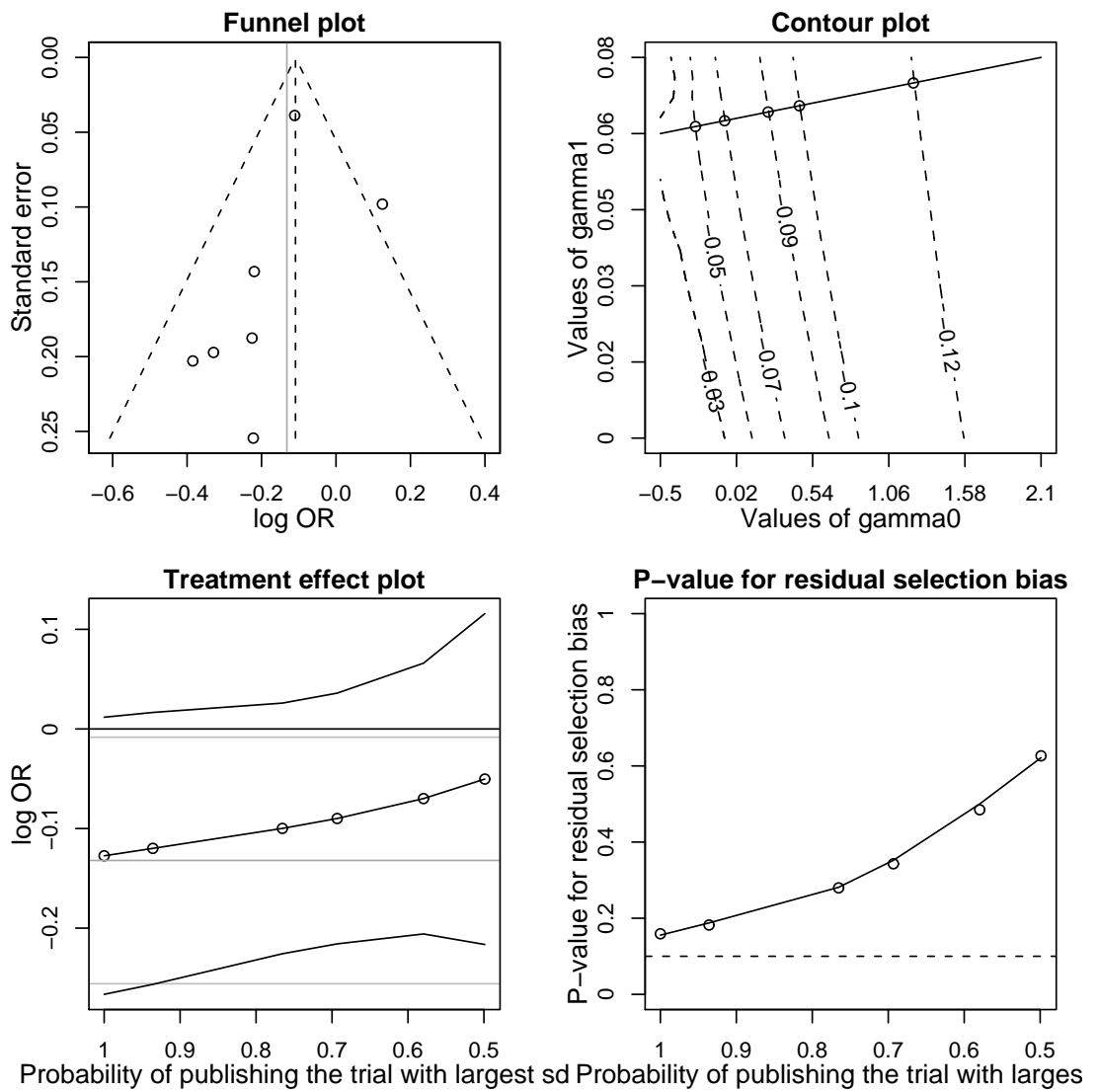


Figura 9: Sesgo de publicación

Índice alfabético

- Error Estandar*, 10
- intervalos de confianza*, 10
- celdas vacías, 10
- cocientes de probabilidad, 12
- Coeficiente de precisión, 25
- curva de respuesta característica, 12
- Curva Roc, 23

- DerSimonian-Laird, 14
- desvío estándar, 8
- distribución a posteriori, 7
- distribución a priori, 7

- Efecto umbral, 23
- Enfoque Bayesiano, 7
- error de muestreo, 14
- estadístico *d* de Cohen, 11
- estadístico *g* de Hedges, 11

- heterogeneidad, 13

- inconsistencia, 14
- inferencias, 7

- método de Peto, 10
- Metanálisis, 4
- Modelo Efectos Aleatorios, 15
- Modelo Efectos Fijos, 15
- Modelo probit de selección, 25

- Odds Ratios Diagnósticos, 15

- perpéciva frecuentista, 7
- pruebas de homogeneidad para sensibilidad
y especificidad, 23

- Revisiones COCHRANE, 27

- Test de razón de verosimilitud, 23

- tiempo transcurrido hasta el evento, 12

- umbrales, 23

- Verosimilitud, 7

Referencias

- [1] Cochrane Collaboration. Cochrane database of systematic reviews. In *The Cochrane Library*, volume Issue 1. Oxford: Update Software, 2003.
- [2] Cochrane Collaboration. *Cochrane Collaboration Handbook*. Cochrane Collaboration. URL <http://www.cochrane.org>.
- [3] JQ Copas, JB Shi. A sensitivity analysis for publication bias in systematic reviews. *Statistical Methods in Medical Research*, 10:251265, 2001.
- [4] J Deeks, J Higgins. *Statistical algorithms in Review Manager 5*. Statistical Methods Group of The Cochrane Collaboration, November 2007. URL http://www.cochrane.org/resources/handbook/Methods_in_RevMan5.pdf.
- [5] N DerSimonian, R Laird. Meta-analysis in clinical trials. *Controlled Clinical Trials*, 7:177188, 1986.
- [6] Schneider M Minder C. Egger M, Smith GD. Bias in meta-analysis detect by a simple, graphical test. *BMJ*, 315:629–34, 1997.
- [7] T Everitt, B & Hothorn. *A Handbook of Statistical Analyses Using R*. Chapman&Hall/CRC, 2006.
- [8] RF Galbraith. Graphical display of estimates having differing standard errors. *Technometrics*, 30:271281, 1988.
- [9] Stern Hal S. Rubin Donald B. Gelman Andrew, Carlin John B. *Bayesian Data Analysis*. 2002.
- [10] GV Glass. Primary, secondary, and meta-analysis of research. *Educational Research*, 5:3–8, 1976.
- [11] O’Rourke K. L’Abbe KA, Detsky AS. Meta-analysis in clinical research. *Annals Internal Medicine*, 107:224–33, 1987.
- [12] DB Light, RJ & Pillemer. Summing up. the science of reviewing research. Cambridge: Havard University Press, 1984.
- [13] Tommas Lumley. *Rmeta Package*, febrero 2008.
- [14] MD Mitchell. Validation of the summary roc for diagnostic test meta-analysis: A monte carlo simulation. *Academic Radiology*, 10:25–31, 2003.
- [15] G Schwarzer. *Copas Package*, febrero 2008.

REFERENCIAS

- [16] G Schwarzer. *meta Package*, febrero 2008.
- [17] A Ebrahim S. Smeeth, L Haines. Numbers needed to treat derived from meta-analyses - sometimes informative, usually misleading. *British Medical Journal*, 318:1548–51, 1999.
- [18] Donald Stangl, D Berry. *Meta-Analysis in Medicine and Health Policy*. Number 4 in Biostatistics: A Series of References and Textbooks. Marcel Dekker Inc, 2000.
- [19] Madrid Unidad de Bioestadística Clínica del Hospital Ramón y Cajal. *Métodos estadísticos*, 2003. URL <http://www.hrc.es/investigacion/metadiscen.htm>.
- [20] SD Walter. Properties of the summary receiver operating characteristic (sroc) curve for diagnostic test data. *Stat Med*, 21:1237–1256, 2002.
- [21] Abraira V. Zamora J, Muriel A. *Métodos Estadísticos*. Unidad de Bioestadística Clínica del Hospital Ramón y Cajal, Madrid,, 2007.

Índice de figuras

1.	Funnel plot para datos(catheter) MEF (MH)	17
2.	Gráfico acumulativo	18
3.	Gráfico de Galbraith (radial) para datos(Olkin95)	20
4.	Gráfico de embudo (Funnel plot) para (MEA) con datos=smoking	22
5.	Sesgo de Publicación(MOdelo librería copas()de R)	23
6.	Ejemplo de asociación entre tratamiento y covariables	26
7.	Modelos de Efectos fijos	30
8.	Modelos de Efectos aleatorios	30
9.	Sesgo de publicación	40

Índice de cuadros

2.	Datos Binarios (Estudios Individuales)	8
3.	Variable de respuesta a escala continua (Estudios Individuales)	9
4.	Diferencia de respuestas (Estudios Individuales)	9
5.	Prueba diagnósticas (Estudio Individual)	9