Meta-análisis de correlaciones en R

Guía práctica

Juan David Leongómez[1]

17 febrero, 2022

[1] Laboratorio de Análisis del Comportamiento Humano (LACH), Facultad de Psicología, Universidad El Bosque, Bogotá, Colombia. Email: jleongomez@unbosque.edu.co. Web: jdleongomez.info.

Descripción

Este documento contiene todo el código explicaciones básicas, paso a paso, para hacer un meta-análisis en R, usando los paquetes `metafor` (Viechtbauer, 2010) y `robumeta` (Fisher & Tipton, 2015). Está principalmente basado en este video, creado por Daniel S. Quintana (2021), pero contiene citas a fuentes primarias, además de información que he agregado.

Esta guía asume una comprensión básica del meta-análisis, así como un manejo básico de R. Sin embargo, de ser necesario, como introducción al meta-análisis recomiendo ver el video introductorio sobre meta-análisis en *jamovi* (Leongómez, 2021) que publiqué anteriormente en mi canal de YouTube *Investigación Abierta*.

Cita éste trabajo como:

Leongómez, J. D. (2022). *Meta-análisis de correlaciones en R: Guía práctica.* Zenodo. https://doi.org/10.5281/zenodo.5640182

Índice

1. Base de datos de ejemplo

Para los ejemplos usados en ésta guía, usaré la base de datos `dat.molloy2014`, tomada de Molloy et al. (2013).

Esta base de datos viene incluida con el paquete `{metafor}` de R. Básicamente, Molloy et al. (2013) estudiaron si existe una asociación entre la diligencia (*conscientiousness*) y la adherencia a la medicación. En otras palabras, ¿las personas más diligentes son más propensas a cumplir con la medicación prescrita?

Primero, primero, debemos cargar los paquetes que usaremos, incluyendo `{metafor}` y `{robumeta}` para hacer metaánálisis, así como `{dplyr}` para manipular y organizar la base de datos.

```
library(robumeta)
library(metafor)
library(dplyr)
```

Una vez cargado el paquete `{metafor}`, ya puedo cargar la base de datos `dat.molloy2014`. En éste caso, para poder *llamarla* cuando sea necesario, la asignaré a un objeto llamado `dat`.

```
dat <- get(data(dat.molloy2014)) %>%
  mutate(study_id = 1:16)  %>% #agregar columna study_id
  select(study_id, authors:quality) #mover study_id como primera columna
```

La base da datos, que he asignado a un objeto llamado `dat`, tiene ahora la siguiente estructura (Tabla 1):

Tabla 1

Estructura de la base de datos

study_id	authors	year	ni	ri	controls	design	a_measure	c_measure	meanage	quality
1	Axelsson et al.	2009	109	0.187	none	cross-sectional	self-report	other	22.00	1
2	Axelsson et al.	2011	749	0.162	none	cross-sectional	self-report	NEO	53.59	1
3	Bruce et al.	2010	55	0.340	none	prospective	other	NEO	43.36	2
4	Christensen et al.	1999	107	0.320	none	cross-sectional	self-report	other	41.70	1
5	Christensen & Smith	1995	72	0.270	none	prospective	other	NEO	46.39	2
6	Cohen et al.	2004	65	0.000	none	prospective	other	NEO	41.20	2
7	Dobbels et al.	2005	174	0.175	none	cross-sectional	self-report	NEO	52.30	1
8	Ediger et al.	2007	326	0.050	multiple	prospective	self-report	NEO	41.00	3
9	Insel et al.	2006	58	0.260	none	prospective	other	other	77.00	2
10	Jerant et al.	2011	771	0.010	multiple	prospective	other	NEO	78.60	3
11	Moran et al.	1997	56	-0.090	multiple	prospective	other	NEO	57.20	2
12	O'Cleirigh et al.	2007	91	0.370	none	prospective	self-report	NEO	37.90	2
13	Penedo et al.	2003	116	0.000	none	cross-sectional	self-report	NEO	39.20	1
14	Quine et al.	2012	537	0.150	none	prospective	self-report	other	69.00	2
15	Stilley et al.	2004	158	0.240	none	prospective	other	NEO	46.20	3
16	Wiebe & Christensen	1997	65	0.040	none	prospective	other	NEO	56.00	1

Nota: Datos tomados de Molloy et al. (2013).

La columna `ri` contiene los coeficientes de correlación de Pearson (la columna `ni` contiene los tamaños de muestra de cada estudio). Dado que los coeficientes de Pearson no tienen una distribución normal, esto podría llevar a calcular varianzas incorrectas, especialmente cuando se trata de correlaciones con tamaños de muestra pequeños.

Adicionalmente, en este ejemplo tenemos una serie de moderadores:

- `controls`: número de variables controladas
- `design`: si se utilizó un diseño transversal o prospectivo
- `a_measure`: tipo de medida de adherencia (autoinforme u otro)
- `c_measure`: tipo de medida de diligencia (NEO u otra)
- `meanage`: edad promedio de la muestra
- `quality`: calidad metodológica

doi: 10.5281/zenodo.5640182

2. Transformación de r de Pearson a z de Fisher

Por esto, vamos a transformar los coeficientes r de Pearson a z de Fisher, que no tienen este problema. Para esto, vamos a usar la función `escalc` del paquete `metafor`.

```
dat <- escalc(measure = "ZCOR",
              ri = ri,
              ni = ni,
              data= dat,
              slab = paste(authors, year, sep = ", "))
```

Esto ha creado dos nuevas variables en nuestra tabla: `yi`, que es el tamaño de efecto, y `vi` que es la varianza.

Tabla 2

Estructura de la base de datos, con transformación de los r de Pearon a z de Fisher

study_id	authors	year	ni	ri	controls	design	a_measure	c_measure	meanage	quality	yi	vi
1	Axelsson et al.	2009	109	0.187	none	cross-sectional	self-report	other	22.00	1	0.1892266	0.0094340
2	Axelsson et al.	2011	749	0.162	none	cross-sectional	self-report	NEO	53.59	1	0.1634399	0.0013405
3	Bruce et al.	2010	55	0.340	none	prospective	other	NEO	43.36	2	0.3540925	0.0192308
4	Christensen et al.	1999	107	0.320	none	cross-sectional	self-report	other	41.70	1	0.3316471	0.0096154
5	Christensen & Smith	1995	72	0.270	none	prospective	other	NEO	46.39	2	0.2768638	0.0144928
6	Cohen et al.	2004	65	0.000	none	prospective	other	NEO	41.20	2	0.0000000	0.0161290
7	Dobbels et al.	2005	174	0.175	none	cross-sectional	self-report	NEO	52.30	1	0.1768200	0.0058480
8	Ediger et al.	2007	326	0.050	multiple	prospective	self-report	NEO	41.00	3	0.0500417	0.0030960
9	Insel et al.	2006	58	0.260	none	prospective	other	other	77.00	2	0.2661084	0.0181818
10	Jerant et al.	2011	771	0.010	multiple	prospective	other	NEO	78.60	3	0.0100003	0.0013021
11	Moran et al.	1997	56	-0.090	multiple	prospective	other	NEO	57.20	2	-0.0902442	0.0188679
12	O'Cleirigh et al.	2007	91	0.370	none	prospective	self-report	NEO	37.90	2	0.3884231	0.0113636
13	Penedo et al.	2003	116	0.000	none	cross-sectional	self-report	NEO	39.20	1	0.0000000	0.0088496
14	Quine et al.	2012	537	0.150	none	prospective	self-report	other	69.00	2	0.1511404	0.0018727
15	Stilley et al.	2004	158	0.240	none	prospective	other	NEO	46.20	3	0.2447741	0.0064516
16	Wiebe & Christensen	1997	65	0.040	none	prospective	other	NEO	56.00	1	0.0400214	0.0161290

Nota: Datos tomados de Molloy et al., (2013).

3. Hacer el meta-análisis

Para hacer el meta-análisis, usaremos la función `rma` del paquete `metafor`, para el que tenemos que especificar los tamaños de efecto (`yi`) y varianzas (`vi`) de los estudios a meta-analizar. En este caso, las columnas donde tenemos estos valores, tienen los mismos nombres (`yi`, `vi`). Asignaré los resultados del meta-análisis a un objeto llamado `res`.

```
res <- rma(yi = yi, vi = vi, data = dat)
```

Los resultados, son los siguientes:

```
res
```

```
##
## Random-Effects Model (k = 16; tau^2 estimator: REML)
##
## tau^2 (estimated amount of total heterogeneity): 0.0081 (SE = 0.0055)
## tau (square root of estimated tau^2 value):      0.0901
## I^2 (total heterogeneity / total variability):   61.73%
## H^2 (total variability / sampling variability):  2.61
##
## Test for Heterogeneity:
## Q(df = 15) = 38.1595, p-val = 0.0009
##
## Model Results:
##
## estimate      se    zval    pval   ci.lb   ci.ub
##   0.1499  0.0316  4.7501  <.0001  0.0881  0.2118  ***
```


doi: 10.5281/zenodo.5640182

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

Primero, nos confirma que ajustamos un modelo con efectos aleatorios (`Random-Effects Model`), a partir de 16 estudios (`k = 16`), y que para estimar τ^2 (tau cuadrado[1]) usamos el método de **máxima verosimilitud restringida**[2] (`tau^2 estimator: REML`), que se designa como *REML* por sus siglas en inglés .

Posteriormente, nos provee los valores de una serie de estimadores de heterogeneidad o varianza:

- τ^2: `tau^2 (estimated amount of total heterogeneity): 0.0081 (SE = 0.0055)`
- τ: `tau (square root of estimated tau^2 value): 0.0901`
- I^2: `I^2 (total heterogeneity / total variability): 61.73%`, y
- H^2: `H^2 (total variability / sampling variability):  2.61`

La tercera parte, reporta una prueba de heterogeneidad, usando el estadístico Q:

- `Test for Heterogeneity:`

  `Q(df = 15) = 38.1595, p-val = 0.0009`

De todos estos, los más comúnmente reportados son τ^2, τ, I^2 y Q. Cada una de estas medidas tiene ventajas y desventajas, por lo cual tiene sentido reportarlas todas.

I^2, por ejemplo, tiene la ventaja de ser sencillo de interpretar, pues hay criterios generales para heterogeneidad baja, moderada y alta (típicamente 25 %, 50 %, and 75 %, respectivamente). Sin embargo, es muy sensible a los tamaños de muestra de los estudios meta-analizados (por ejemplo, si en tu meta-análisis hay estudios con tamaños de muestra muy grandes, esto va a sesgar tu I^2).

Q, aunque no es sensible al tamaño de muestra, es sensible al número de estudios meta-analizados. Tiene la ventaja de ser un test de hipótesis, y como tal, puede ser interpretado a partir de su valor p.

τ^2 no tiene estos problemas, pero es más difícil de interpretar.

En nuestro caso, el estadístico Q sugiere que hay una heterogeneidad significativa en los estudios meta-analizados ($p = 0.0009$). I^2, sugiere una heterogeneidad moderada, lo que quiere decir que más de la mitad (61.73 %) de la varianza se estima que se deriva de diferencias en los tamaños de efecto.

Por último, tenemos los resultados del modelo de meta-análisis (`Model results`). Nos provee un estimado de la asociación positiva entre diligencia y adherencia a la medicación (0.1499 ± 0.0316), lo que equivale a un valor z de 4.7501 , y sugiere que esa asociación es significativa ($p < .0001$). Así mismo, nos provee los límites inferior (0.0881) y superior (0.2118) de los intervalos de confianza.

3.1. Más información sobre heterogeneidad

Además de reportar los estadísticos τ^2, τ, I^2 y Q, podemos fácilmente calcular los intervalos de confianza para τ^2, τ, e I^2 con la función `confint`, que también pueden ser reportado junto a estos estadísticos.

```r
confint(res)
```

```
##
##        estimate   ci.lb   ci.ub
## tau^2    0.0081  0.0017  0.0378
## tau      0.0901  0.0412  0.1944
## I^2(%)  61.7324 25.2799 88.2545
## H^2      2.6132  1.3383  8.5139
```

[1] τ^2 es una estimación de la varianza de los tamaños de los efectos reales entre los estudios meta-analizados. Se usa, principalmente, para asignar pesos a cada estudio. Para más información, ver Borenstein et al. (2009).

[2] Hay varios métodos disponibles como estimador, además de **máxima verosimilitud restringida** (REML). Sin embargo, si no estás seguro, REML es una buena opción. Cada método tiene ventajas y desventajas que, si tienes interés en mirar, están descritas en la documentación de la función `rma`.

doi: 10.5281/zenodo.5640182

Para el τ^2, el hecho de que los intervalos de confianza no crucen el 0 (en nuestro caso 0.0017 — 0.0378), sugiere que de hecho también que hay heterogeneidad entre los estudios que meta-analizamos.

3.2. Diagnóstico de influencia

Otro aspecto importante de un meta-análisis, es determinar si alguno(s) de los estudios meta-analizados es(son) particularmente influyente(s) en nuestro resultado[3]. Para esto, podemos usar la función `influence`, cuyo resultado en este caso asignaré a un objeto llamado `inf`.

```
inf <- influence(res)
```

Ya que lo asigné a un objeto (`inf`), para ver el resultado, tengo que correrlo para ver su resultado.

```
inf
```

```
##
##                             rstudent  dffits cook.d  cov.r tau2.del  QE.del
## Axelsson et al., 2009         0.2918  0.0485 0.0025 1.1331   0.0091 37.7109
## Axelsson et al., 2011         0.1196 -0.0031 0.0000 1.2595   0.0100 36.7672
## Bruce et al., 2010            1.2740  0.2595 0.0660 0.9942   0.0075 35.3930
## Christensen et al., 1999      1.4711  0.3946 0.1439 0.9544   0.0068 33.5886
## Christensen & Smith, 1995     0.8622  0.1838 0.0339 1.0505   0.0082 36.5396
## Cohen et al., 2004           -0.9795 -0.2121 0.0455 1.0639   0.0084 37.1703
## Dobbels et al., 2005          0.2177  0.0296 0.0010 1.1740   0.0094 37.6797
## Ediger et al., 2007          -0.9774 -0.3120 0.1001 1.1215   0.0084 36.1484
## Insel et al., 2006            0.7264  0.1392 0.0195 1.0561   0.0083 37.0495
## Jerant et al., 2011          -1.8667 -0.5861 0.2198 0.8502   0.0047 25.0661
## Moran et al., 1997           -1.4985 -0.2771 0.0756 1.0073   0.0077 35.6617
## O'Cleirigh et al., 2007       1.8776  0.4918 0.2148 0.8819   0.0059 31.9021
## Penedo et al., 2003          -1.1892 -0.2939 0.0859 1.0550   0.0080 36.3291
## Quine et al., 2012           -0.0020 -0.0423 0.0021 1.2524   0.0100 37.7339
## Stilley et al., 2004          0.8066  0.2126 0.0459 1.0907   0.0083 35.8385
## Wiebe & Christensen, 1997    -0.7160 -0.1656 0.0280 1.0853   0.0087 37.7017
##                                 hat  weight    dfbs inf
## Axelsson et al., 2009        0.0568  5.6776  0.0481
## Axelsson et al., 2011        0.1054 10.5396 -0.0032
## Bruce et al., 2010           0.0364  3.6432  0.2623
## Christensen et al., 1999     0.0562  5.6195  0.3994
## Christensen & Smith, 1995    0.0441  4.4069  0.1837
## Cohen et al., 2004           0.0411  4.1094 -0.2112
## Dobbels et al., 2005         0.0714  7.1362  0.0296
## Ediger et al., 2007          0.0889  8.8886 -0.3128
## Insel et al., 2006           0.0379  3.7886  0.1387
## Jerant et al., 2011          0.1058 10.5826 -0.5430
## Moran et al., 1997           0.0369  3.6922 -0.2791
## O'Cleirigh et al., 2007      0.0511  5.1150  0.5059
## Penedo et al., 2003          0.0587  5.8732 -0.2941
## Quine et al., 2012           0.0998  9.9778 -0.0434
## Stilley et al., 2004         0.0684  6.8403  0.2125
## Wiebe & Christensen, 1997    0.0411  4.1094 -0.1642
```

Esto me muestra gran cantidad de información de cada estudio (en este caso, una tabla sin formato, que es muy ancha para poder imprimirse en esta página, por lo cual está reportada en dos partes). Sin embargo, lo más importante ahora es mirar la última columna, también llamada `inf`. Si ahí aparecieran asteriscos (que no es nuestro caso), sugeriría que ese estudio es particularmente influyente.

Por último, podemos también ver ésta información que tenemos guardada en el objeto `inf`, de manera gráfica,

[3] Por ejemplo, si estuviésemos meta-analizando 20 estudios, de los cuales 19 tienen un n de 100, pero el otro tiene un n de 10.000, éste último tendrá una influencia enorme en nuestro resultado. Sería preocupante que tu meta-análisis sea dependiente de un único estudio.

doi: 10.5281/zenodo.5640182

usando la función `plot`.

```r
plot(inf)
```

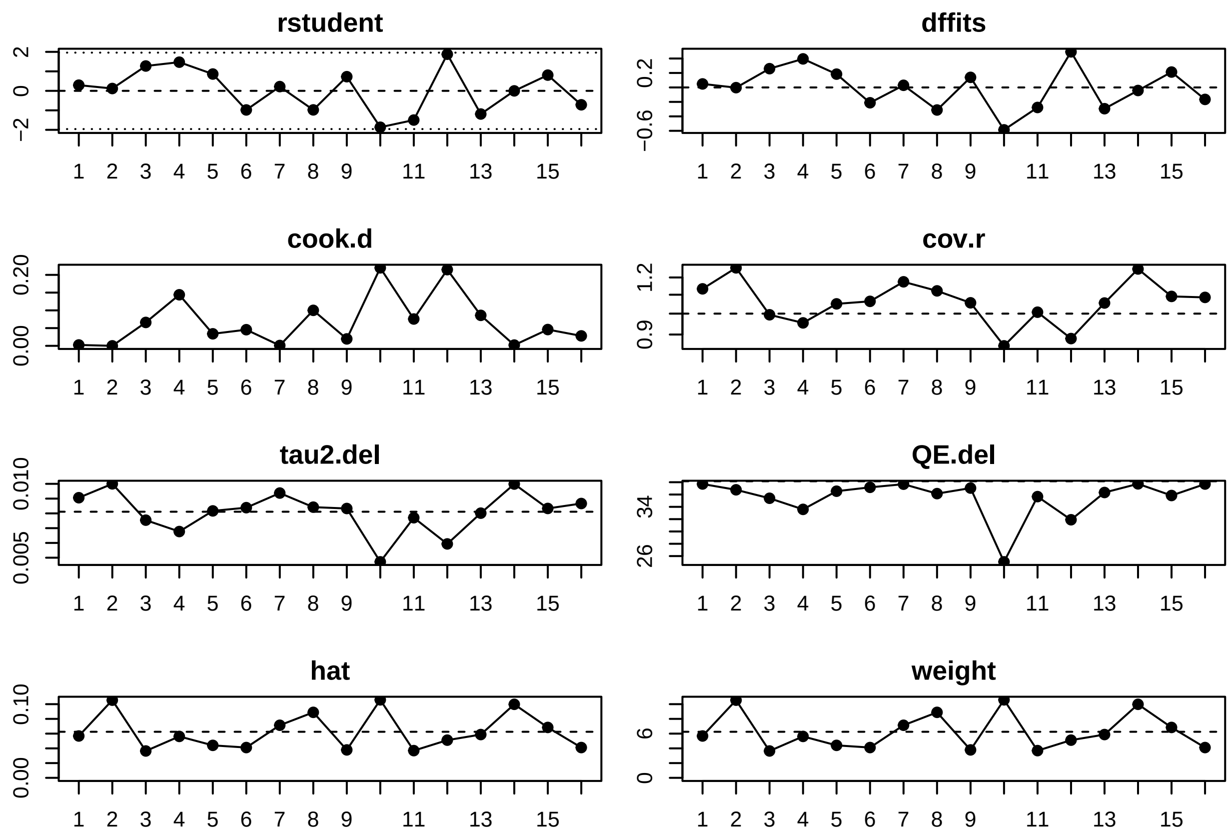

Figura 1. Diagnóstico de influencia. Estudios particularmente influyentes serpian representados con un punto rojo.

En este caso, no hay ningún estudio que se considere demasiado influyente, por lo que podemos estar tranquilos con nuestro meta-análisis.

3.3. *Forest plot* (diagrama de bosque)

Para hacer un diagrama de bosque (*forest plot*) con `metafor` resumiendo nuestro meta-análisis, solo tenemos que usar la función `forest`, usando como argumento el objeto al que asignamos los resultados de nuestro meta-análisis (`res`).

Como se puede ver en las Figuras 2, 3 y 4 (que son 3 versiones del mismo f*orest plot*), no es una sorpresa que el análisis nos sugiera bastante heterogeneidad; las correlaciones encontradas entre los diferentes estudios varían mucho (están entre -0.09 y 0.37), y aunque son positivas en la mayoría de los casos (en algunos claramente positivas), en algunos son prácticamente 0 o incluso negativas.

```r
forest(res)
```


doi: 10.5281/zenodo.5640182

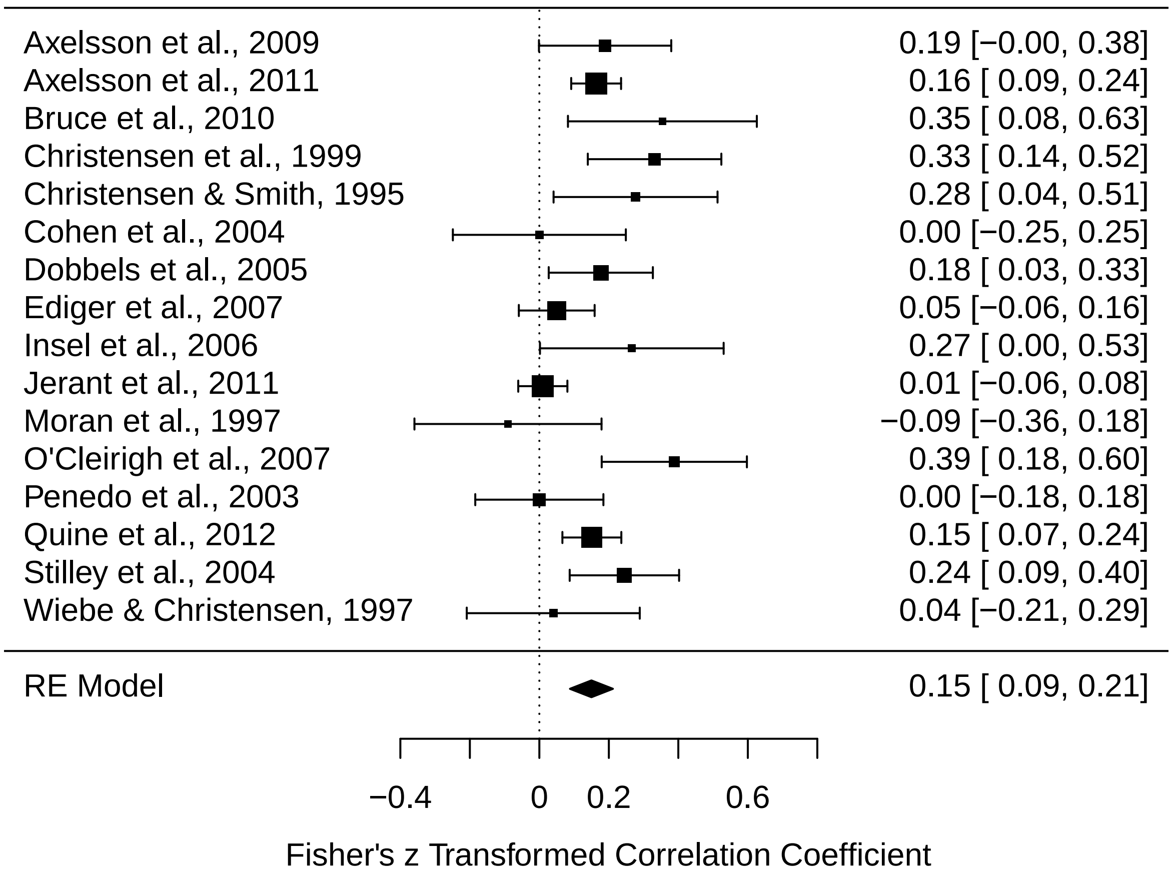

Figura 2. Forest plot básico de metafor. Para cada estudio meta-analizado, tenemos el efecto (correlación, en este caso en valores z de Fisher), así como sus intervalos de confianza entre paréntesis cuadrados. Esta misma información está representada gráficamente, con los cuadrados representando el efecto de cada estudio así como sus intervalos de confianza, y el tamaño de muestra (representado por el tamaño del cuadrado). Bajo estos resultados, tenemos nuestro meta-análisis, con el mismo formato en texto, pero representando el efecto y sus intervalos de confianza con un diamante.

Para una versión más completa y anotada, también usando el *plot* básico de `metafor`, pero representando coeficientes de correlación de Pearson (r) en vez de valores z, así como agregando una columna con los pesos dados a cada estudio, y detalles del modelo final, podemos agregar algunas opciones (explicadas acá):

```
# forest plot con anotaciones adicionales
forest(res, cex = 0.75, xlim = c(-1.6, 1.6),
       showweights = TRUE,
       atransf = transf.ztor,
```


doi: 10.5281/zenodo.5640182

```r
       xlab = "Correlación (r)",
       digits = c(2,3L),
       at = transf.rtoz(c(-0.4, -0.2, 0, 0.2, 0.4, 0.6)),
       mlab = bquote(paste("Modelo EA: Q(", .(res$k - res$p), ") = ",
     .(formatC(res$QE, digits=2, format="f")),
     ", p ", .(scales::pvalue(res$pval)), "; ", I^2, " = ",
     .(formatC(res$I2, digits=1, format="f")), "%")))
# agregar encabezados a las columnas (valores de X y Y deben ser ajustados)
op <- par(cex = 0.8, font=2)
text(x = -1.6, y = 18, labels = "Autor(es), Año", pos = 4)
text(x = 0, y = 18, labels = "Efecto e IC", pos = 4)
text(x = 1, y = 18, labels = "Peso", pos = 2)
text(x = 1.6, y = 18, labels = "Corr. [95% IC]", pos = 2)
```

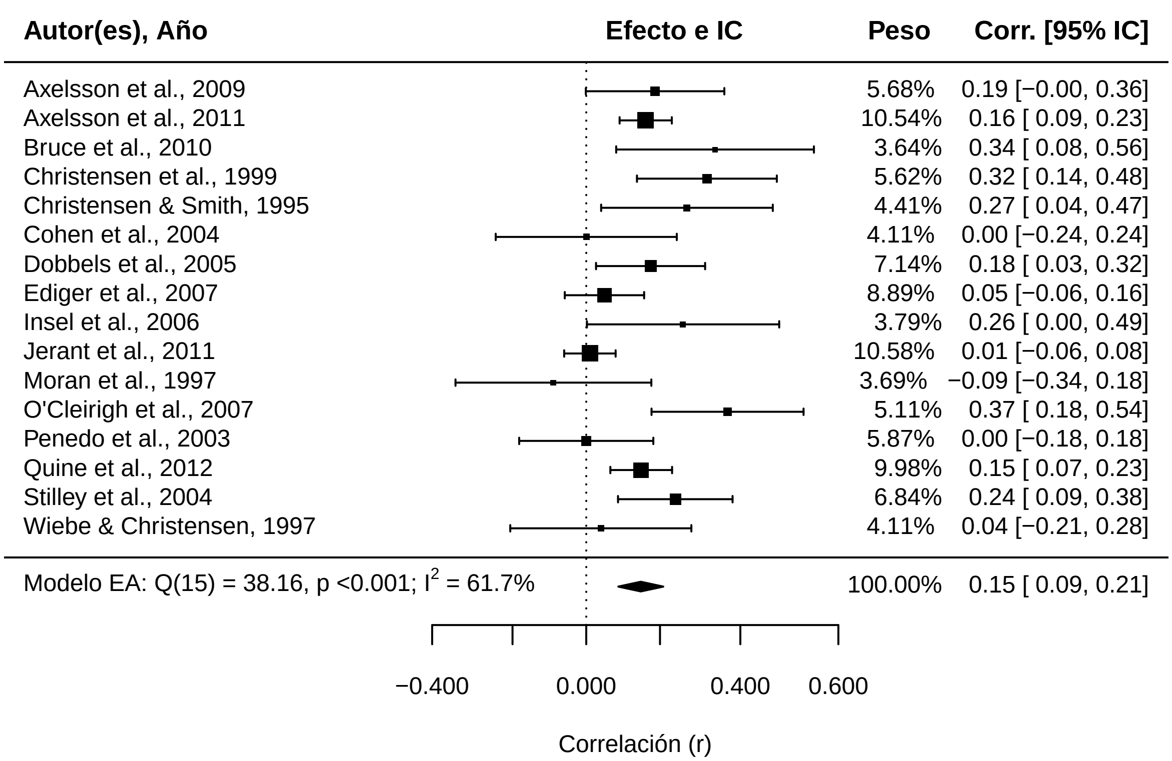

Figura 3. Forest plot anotado, creado con metafor. En esta versión agregué algunos encabezados en español, así como estadísticos generales del modelo de meta-análisis. Modelo EA se refiere al modelo meta-analizado, de efectos aleatorios.

O, para una incluso más sofisticada, se puede usar la función `viz_forest` del paquete `metaviz`, en este caso usando la variante `rain` (no es necesaria).

```r
library(metaviz)
viz_forest(res,
```


doi: 10.5281/zenodo.5640182

```
            study_labels = paste(dat$authors, dat$year, sep = ", "),
            xlab = "Correlación",
            variant = "rain",
            annotate_CI = TRUE,
            summary_label = "Resumen",
            text_size = 2.6)
```

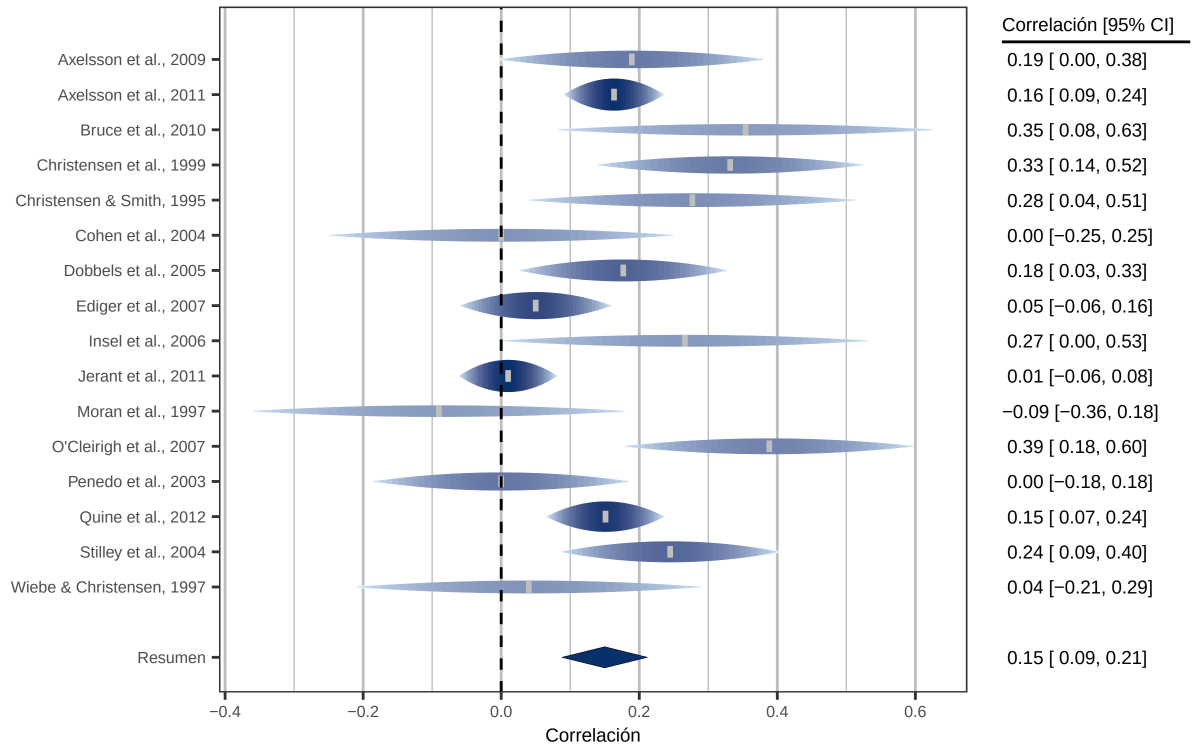

Figura 4. Forest plot creado con metaviz.

3.4. *Funnel plot* (diagrama de embudo) y sesgo de estudios pequeños

En este punto, es en donde más errores se cometen. Las pruebas más comunes para evaluar sesgos de publicación, son la evaluación de la asimatría en el *funnel plot* (diagrama de embudo), y la regresión (o test) de Egger (Egger et al., 1997).

El principal error que la mayoría de los investigadores (meta-analistas) cometen, es que simplemente basándose en éstos métodos, concluyen que un meta-análisis tiene (o no) riesgo de sufrir de un sesgo de publicación. Sin embargo, estos métodos, no son pruebas exclusivas de sesgo de publicación, sino de sesgo de estudios de tamaño muestral pequeño (ver e.g. Schwarzer et al., 2015), que pueden incluir sesgo de publicación, pero no se centran exclusivamente en éste.

A pesar de esto, tanto la regresión de Egger como el *funnel plot*, son intersantes dado que el sesgo de estudios pequeños es importante.

3.4.1. *Funnel plot*

Para crear un *funnel plot* con `metafor`, de nuestro meta-análisis, solo tenemos que usar la función `funnel`, usando como argumento el objeto al que asignamos los resultados de nuestro meta-análisis (`res`).

```
funnel(res)
```


doi: 10.5281/zenodo.5640182

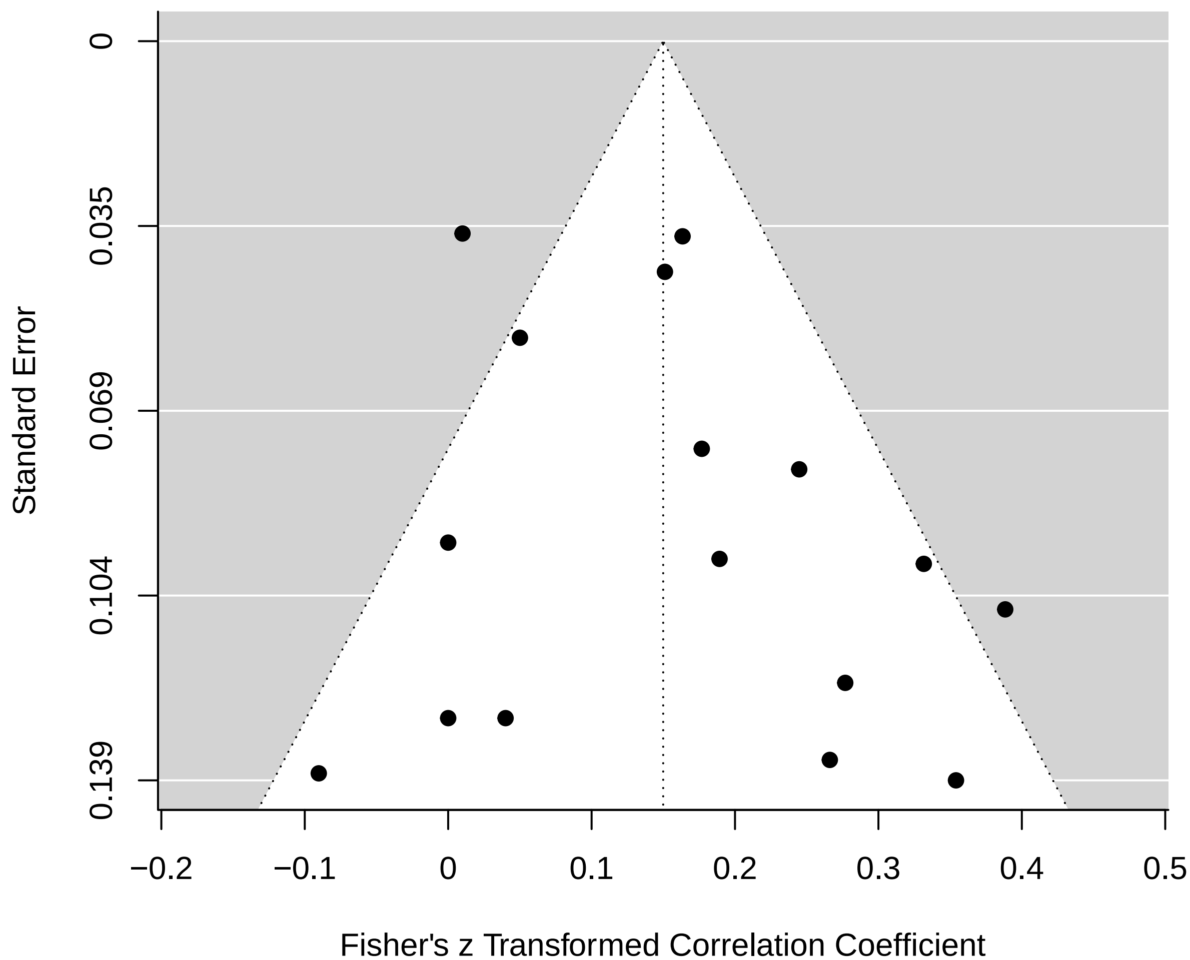

Figura 5. Funnel plot básico de metafor. Para cada estudio meta-analizado, tenemos el efecto (correlación, en este caso en valores z de Fisher) en el eje X, así como su error estándar en el eje Y. La línea punteada vertical representa el efecto meta-analizado que hemos encontrado, así que podemos ver los estudios que encontraron un efecto mayor (derecha de la línea punteada) o menor (izquierda) de éste. A primera vista no parece haber mucha asimetría, pero es importante tener en cuenta que es un análisis muy subjetivo.

De nuevo, se puede usar el paquete `metaviz`, usando la función `viz_funnel`. Hay muchas opciones, pero como ejemplo, usaré la versión por defecto, agregando solo la línea de la regressión de Egger (`egger = TRUE`; ver sección 3.4.2, a continuación), y transformando los tamaños de efecto de regreso a r de Pearson (`x_trans_function = tanh`).

```
viz_funnel(res,
           egger = TRUE,
           x_trans_function = tanh)
```


doi: 10.5281/zenodo.5640182

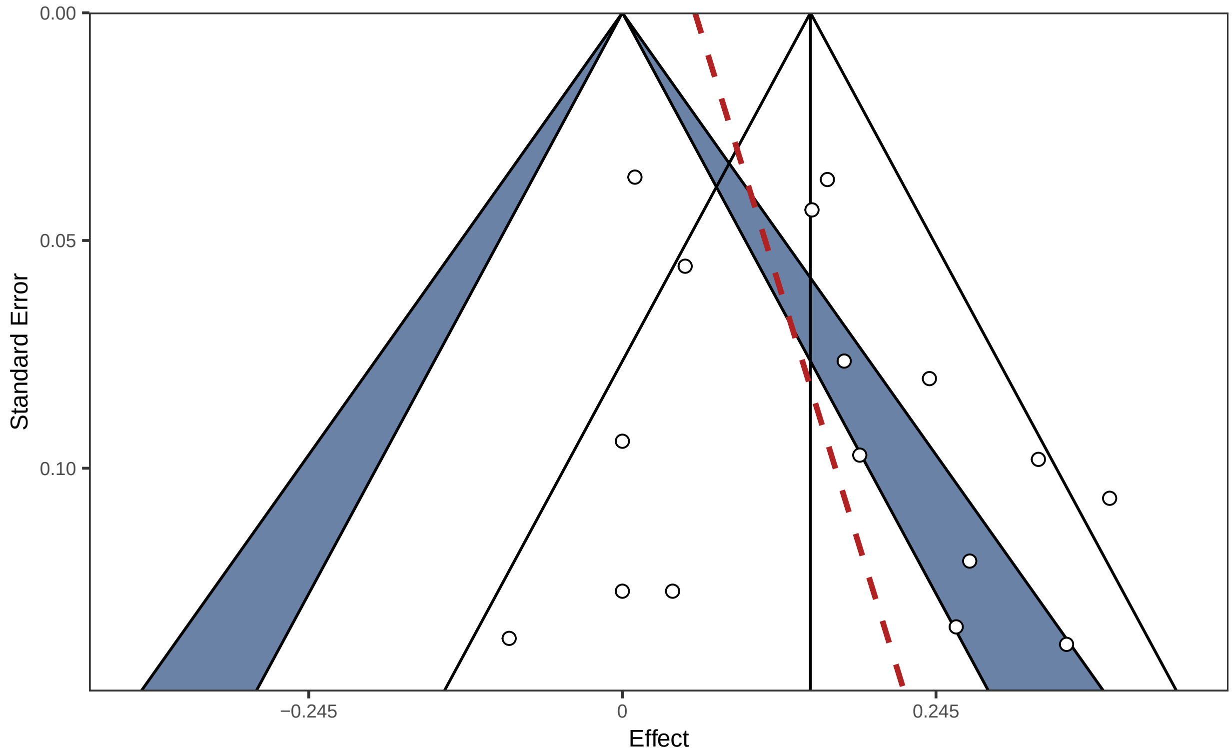

Figura 6. Funnel plot creado con metaviz. En azul, se representa el área donde estudios, segpun su error (y su tamaño de muestra), tendrían un efecto significativo al 5 % (i.e. $p > 0.05$), y fuera de ésta, donde tendrían un efecto significativo al 1 % (i.e. $p > 0.01$). La línea negra vertical representa el efecto meta-analizado, y el triángulo a partir de su inicio, el área donde se ubican los estudios que no se diferencian significativamente del resultado del meta-análisis. La línea roja punteada, representa la regresión de Egger.

Alternativamente, el paquete `metaviz` tiene la función `viz_sunset`, que permite además mostrar el poder estadístico (o potencia) de los estudios meta-analizados para detectar un efecto de interés mediante una prueba de Wald de dos colas. A continuación, muestro dos versiones de esta función. En ambos casos, agregué el efecto *real* encontrado con el meta-análisis (`contours = TRUE`), y transformé los tamaños de efecto de regreso a r de Pearson (`x_trans_function = tanh`).

```
# Escala de poder discreta
viz_sunset(res,
           contours = TRUE,
           x_trans_function = tanh)

# Escala de poder contínua
viz_sunset(res,
           contours = TRUE,
           x_trans_function = tanh,
           power_contours = "continuous")
```


doi: 10.5281/zenodo.5640182

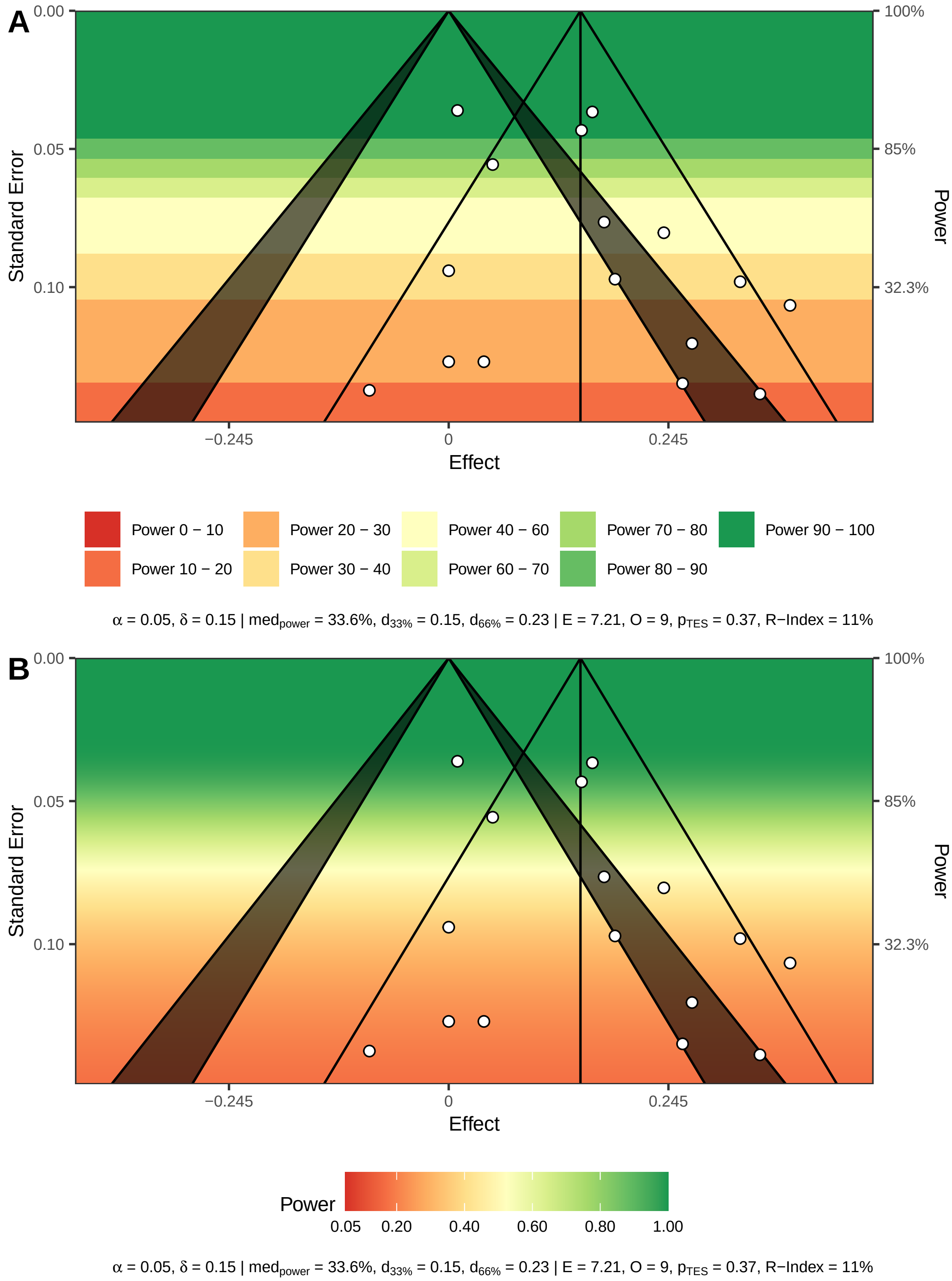

Figura 7. Dos versiones de *funnel plot* creados con metaviz, usando la función viz-sunset, que estima el poder de cada estudio para detectar un efecto de interés. **A.** Poder representado por bandas dicretas de color. **B.** Poder representado de manera contpinua en una escala de color. En ambos casos, y tal como en la Fig. 6, el efecto real está representado como una línea vertical, y el triángulo a partir de su inicio representa el área donde se ubican los estudios que no se diferencian significativamente del resultado del meta-análisis.

doi: 10.5281/zenodo.5640182

3.4.2. Regresión de Egger

Para hacer una prueba formal de sesgo de estudios pequeños, podemos hacer una prueba o regresión de Egger (Egger et al., 1997). En `metafor`, esto se hace con la función `regtest`, de nuevo usando como argumento el objeto al que asignamos el resultado de nuestro meta-análisis (`res`).

```
regtest(res)
```

Como se puede ver, la prueba de Egger no muestra un resultado significativo (`z = 1.0216, p = 0.3070`).

Con base en esto, y la inspección visual subjetiva del *funnel plot*, muchos investigadores concluyen que no hay sesgo de publicación. Sin embargo, como mencioné antes, estas pruebas no se centran en el sesgo de publicación sino en el sesgo de estudios pequeños. En otras palabras, con base en esto, lo único que podemos concluir correctamente, es que no hay sesgo de estudios pequeños (más adelante, en la sección XXXXX, explicaré cómo evaluar si hay sesgo de publicación).

4. Meta-análisis de correlación con moderador

4.1. Ejemplo 1: Moderación de la edad promedio de los participantes

Primero, y como ejemplo, vamos a ver si la edad (en nuestros datos, `meanage`) modera el resultado. Esto es importante, pues hay una enorme variación entre las edades medias de los participantes de los diferentes estudios[4], lo que podría moderar (afectar) la asociación entre diligencia (*conscientiousness*) y adherencia a la medicación prescrita.

Para esto, de nuevo podemos usar la función `rma` de paquete `metafor` y de la misma manera que en la sección 3, pero agregando nuestra variable moderadora (`meanage`) al argumento `mods`. En este caso voy a asignar a un objeto llamado `res.modage`, para diferenciarlo del objeto `res` al que asigné el meta-análisis básico, sin moderadores.

```
res.modage <- rma(yi = yi, vi = vi, mods = ~meanage, data = dat)
```

Los resultados, son los siguientes:

```
res.modage
```

```
##
## Mixed-Effects Model (k = 16; tau^2 estimator: REML)
##
## tau^2 (estimated amount of residual heterogeneity):     0.0072 (SE = 0.0054)
## tau (square root of estimated tau^2 value):             0.0846
## I^2 (residual heterogeneity / unaccounted variability): 56.50%
## H^2 (unaccounted variability / sampling variability):   2.30
## R^2 (amount of heterogeneity accounted for):            11.76%
##
## Test for Residual Heterogeneity:
## QE(df = 14) = 30.9050, p-val = 0.0057
##
## Test of Moderators (coefficient 2):
## QM(df = 1) = 1.4286, p-val = 0.2320
##
## Model Results:
##
##          estimate      se     zval    pval    ci.lb   ci.ub
## intrcpt    0.2741  0.1090   2.5147  0.0119   0.0605  0.4877  *
## meanage   -0.0024  0.0020  -1.1952  0.2320  -0.0063  0.0015
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

[4] De hecho, mientras que en el estudio de Axelsson et al. (2009) la edad promedio fue de 22, en el estudio de Jerant et al. (2011) la edad promedio fue de 78.6.

doi: 10.5281/zenodo.5640182

Los resultados, que tienen la misma organización que los del análisis sin moderadores (sección 3) resultado nos muestra que, a pesar de la gran diferencia de edad entre estudios, la edad no tiene un efecto significativo, como se puede ver en la columna `pval` para el efecto de `meanage` (0.232).

4.1.1. *Forest plot* y *funnel plot*

Por supuesto, de estos resultados también puedo crear *forest plots* y *funnel plots*, siguiendo los ejemplos y código de la sección 3.

Para el *forest plot*, hago a continuación un ejemplo usando la opción básica (por supuesto, se pueden crear versiones anotadas y mejoradas, pero la función `viz_forest` del paquete `metaviz` tendrá problemas para crear un *forest plot* de un meta-análisis con moderadores).

```
# forest plot con anotaciones adicionales
forest(res.modage, xlim = c(-1.6, 1.6))
text(-1.6, 18, "Autor(es), Año", pos = 4)
text(1.6, 18, "Correlación [95% CI]", pos = 2)
```

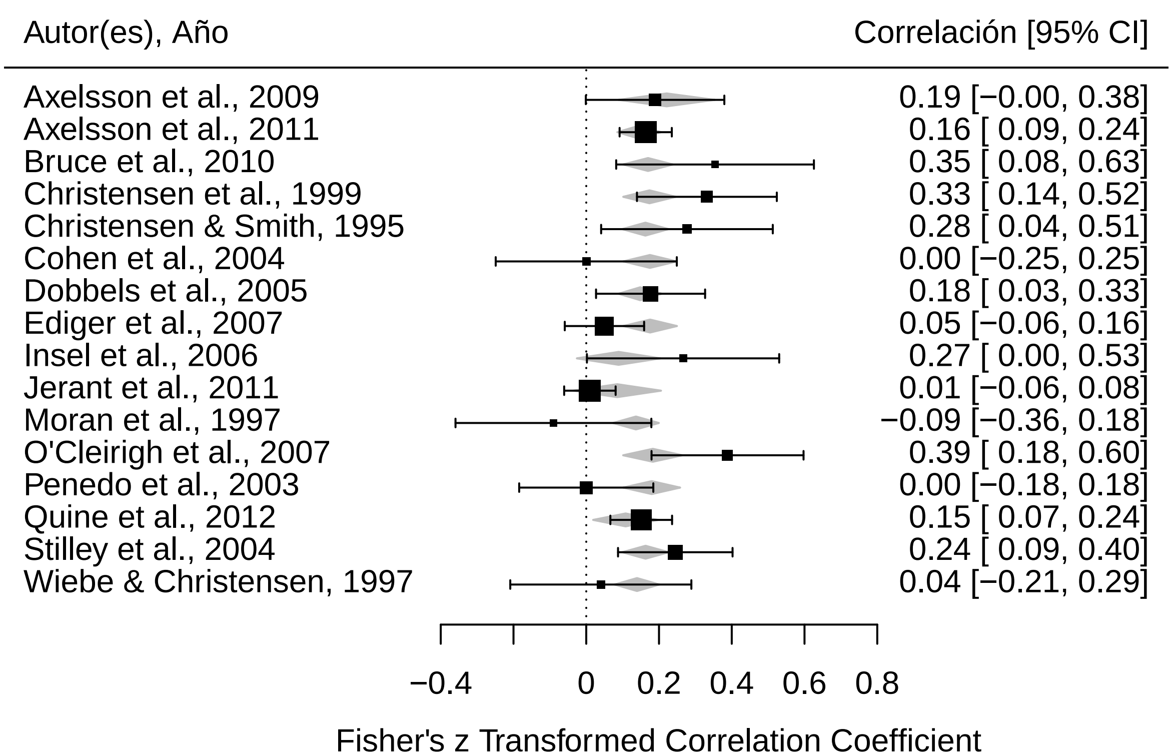

Figura 8. Forest plot básico de metafor, para un meta-análisis incluyendo la edad promedio de los participantes como moderador. En la ilustración gráfica, además de los efectos originales, se puede ver el efecto de cada estudio estimado cuando se incluye el moderador como polígonos (diamantes) de color gris. Sin embargo, ya no obtenemos una fila representando el efecto promediado del meta-análisis, ya que no tenemos un solo efecto.

De manera similar, podemos obtener un *funnel plot* de nuestro meta-análisis con moderador, pero éste nos mostrará,

doi: 10.5281/zenodo.5640182

en vez de los coeficientes de correlación (transformados a z de Fisher), los valores residuales de cada estudio (es decir, qué tanto se alejan del resultado de nuestro meta-análisis):

```
funnel(res.modage)
```

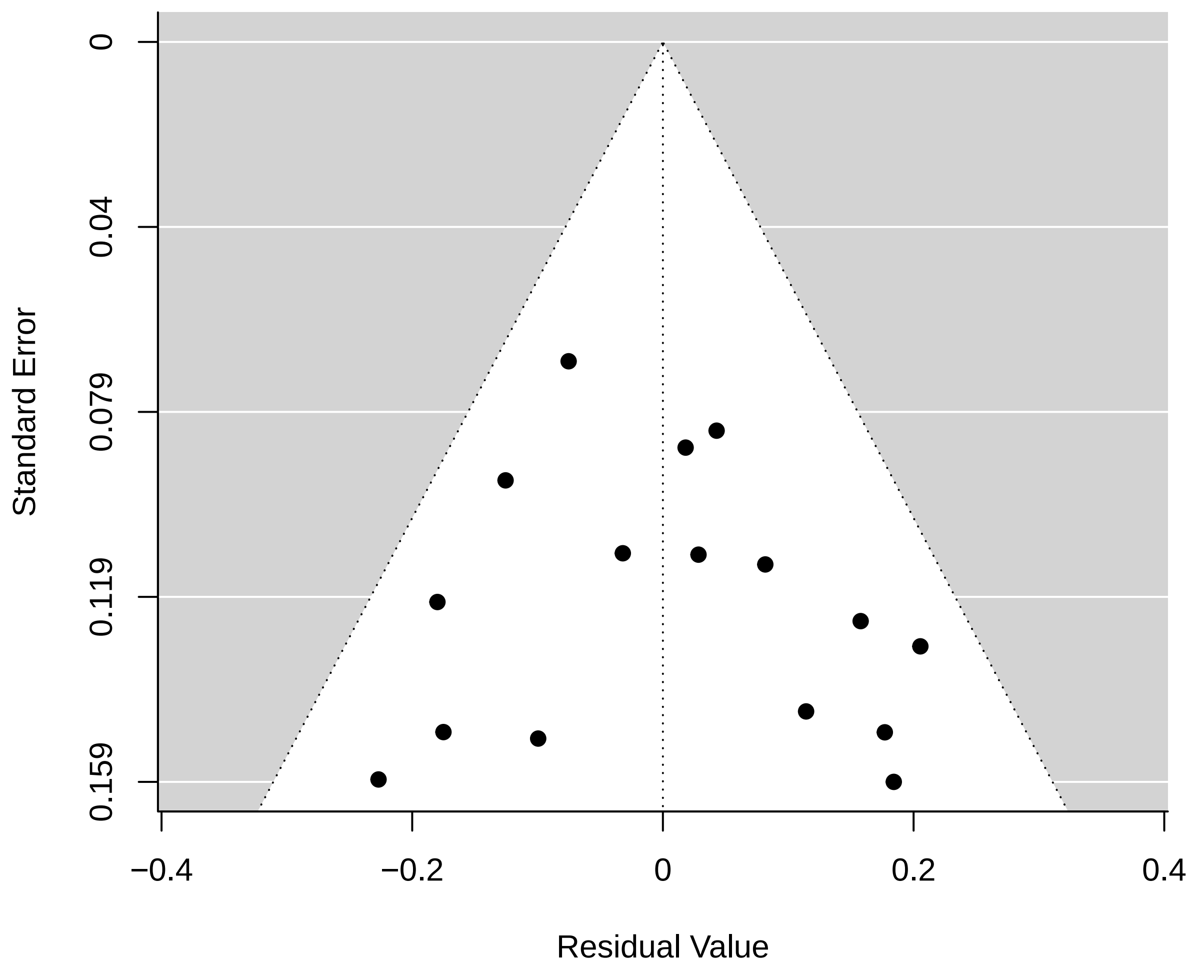

Figura 9. Funnel plot básico de metafor, para un meta-análisis incluyendo la edad promedio de los participantes como moderador. La línea punteada vertical representa el efecto meta-analizado que hemos encontrado, así que podemos ver los estudios que encontraron un efecto mayor (derecha de la línea punteada) o menor (izquierda) de éste.

4.2. Ejemplo 2: Moderación de la calidad de los estudios meta-analizados

La base de datos con tiene una medida de calidad metodológica de los estudios (variable `quality`). Dicha calidad, también podría moderar la asociación entre diligencia (*conscientiousness*) y adherencia a la medicación prescrita. Siguiendo los mismos pasos, puedo hacer éste análisis, pero voy a asignar este meta-análisis a un objeto llamado `res.modq` para diferenciarlo de los demás.

doi: 10.5281/zenodo.5640182

```
res.modq <- rma(yi = yi, vi = vi, mods = ~quality, data = dat)
res.modq
```

```
##
## Mixed-Effects Model (k = 16; tau^2 estimator: REML)
##
## tau^2 (estimated amount of residual heterogeneity):     0.0078 (SE = 0.0057)
## tau (square root of estimated tau^2 value):             0.0884
## I^2 (residual heterogeneity / unaccounted variability): 57.79%
## H^2 (unaccounted variability / sampling variability):   2.37
## R^2 (amount of heterogeneity accounted for):            3.73%
##
## Test for Residual Heterogeneity:
## QE(df = 14) = 30.4205, p-val = 0.0067
##
## Test of Moderators (coefficient 2):
## QM(df = 1) = 0.6393, p-val = 0.4240
##
## Model Results:
##
##          estimate      se     zval    pval    ci.lb   ci.ub
## intrcpt    0.2082  0.0796   2.6149  0.0089   0.0521  0.3643  **
## quality   -0.0312  0.0391  -0.7995  0.4240  -0.1078  0.0453
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

De nuevo, encontramos que éste moderador (`quality`), al igual que la edad promedio (`meanage`), no tiene un efecto significativo, como se puede ver en la columna `pval` para el efecto de `quality` (0.424).

Por supuesto, *forest plots* y *funnel plots* pueden ser creados, tal y como describí en la sección 4.1.1.

4.3. Ejemplo 3: Moderación de las controles usados en cada estudio meta-analizado

Como último ejemplo, voy a mirar si el hecho de que los estudios tengan variables que fueron controladas, modera la asociación entre diligencia (*conscientiousness*) y adherencia a la medicación prescrita. Siguiendo los mismos pasos, voy hacer éste análisis, pero voy a asignar este meta-análisis a un objeto llamado `res.mes`. Son embargo, dado que la variable que contiene esta información (`controls`) es un factor, pero no está definido como tal, debo hacerlo en la usando la función `factor` al ingresar el argumento `mods` (i.e. `mods = ~factor(controls)`).

```
res.mes <- rma(yi = yi, vi = vi, mods = ~factor(controls), data = dat)
res.mes
```

```
##
## Mixed-Effects Model (k = 16; tau^2 estimator: REML)
##
## tau^2 (estimated amount of residual heterogeneity):     0.0000 (SE = 0.0015)
## tau (square root of estimated tau^2 value):             0.0002
## I^2 (residual heterogeneity / unaccounted variability): 0.00%
## H^2 (unaccounted variability / sampling variability):   1.00
## R^2 (amount of heterogeneity accounted for):            100.00%
##
## Test for Residual Heterogeneity:
## QE(df = 14) = 18.0370, p-val = 0.2051
##
## Test of Moderators (coefficient 2):
## QM(df = 1) = 20.1221, p-val < .0001
##
```

 doi: 10.5281/zenodo.5640182

```
## Model Results:
##
##                       estimate      se    zval    pval    ci.lb   ci.ub
## intrcpt                 0.0167  0.0296  0.5635  0.5731  -0.0413  0.0746
## factor(controls)none    0.1621  0.0361  4.4858  <.0001   0.0913  0.2329  ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

En éste caso, a diferencia de los ejemplos de moderación anteriores, la variable moderadora (`controls`) sí tiene un efecto significativo, como se puede ver en la columna `pval` para el efecto de `factor(controls)none` (<0.001), y en los asteriscos que aparecen al final de esa fila (`***`).

Por supuesto, *forest plots* y *funnel plots* pueden ser creados, tal y como describí en la sección 4.1.1.

5. Sesgo de publicación (*publication bias*)

Para determinar el sesgo de publicación, se puede usar la función `weightfunct` del paquete `{weightr}`, que nos permite "estimar tanto el modelo de función de peso para el sesgo de publicación que se publicó originalmente en Vevea y Hedges (1995) como la versión modificada presentada en Vevea y Woods (2005)," como se describe en la documentación de la función `weightfunct`.

```
library(weightr)
```

En este caso, usaré esta función, asignando el resultado a un objeto que llamaré `wf`.

```
wf <- weightfunct(effect = dat$yi, v = dat$vi, table = TRUE)
wf
```

```
##
## Unadjusted Model (k = 16):
##
## tau^2 (estimated amount of total heterogeneity): 0.0070 (SE = 0.0051)
## tau (square root of estimated tau^2 value):  0.0834
##
## Test for Heterogeneity:
## Q(df = 15) = 38.1595, p-val = 0.001436053
##
## Model Results:
##
##           estimate std.error z-stat      p-val   ci.lb  ci.ub
## Intercept   0.1486   0.03073  4.835 1.3288e-06 0.08837 0.2088
##
## Adjusted Model (k = 16):
##
## tau^2 (estimated amount of total heterogeneity): 0.0056 (SE = 0.0045)
## tau (square root of estimated tau^2 value):  0.0750
##
## Test for Heterogeneity:
## Q(df = 15) = 38.1595, p-val = 0.001436053
##
## Model Results:
##
##               estimate std.error z-stat    p-val    ci.lb  ci.ub
## Intercept      0.09153   0.04464  2.050 0.040341  0.00403 0.1790
## 0.025 < p < 1  0.24121   0.20122  1.199 0.230626 -0.15317 0.6356
##
## Likelihood Ratio Test:
## X^2(df = 1) = 2.98493, p-val = 0.084043
```


doi: 10.5281/zenodo.5640182

```
##
## Number of Effect Sizes per Interval:
##
##                        Frequency
## p-values <0.025                9
## 0.025 < p-values < 1           7
```

Referencias

Borenstein, M., Hedges, L. V., Higgns, J. P. T., & Rothstein, H. R. (2009). Identifying and Quantifying Heterogeneity. In *Introduction to Meta-Analysis* (pp. 107–125). Wiley. https://doi.org/10.1002/9780470743386.ch16

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, *315*(7109), 629–634. https://doi.org/10.1136/bmj.315.7109.629

Fisher, Z., & Tipton, E. (2015). Robumeta: An R-package for robust variance estimation in meta-analysis. *arXiv:1503.02220 [Stat]*. https://arxiv.org/abs/1503.02220

Leongómez, J. D. (2021). *Hacer meta-análisis en jamovi es muy fácil*. [Archivo de Vídeo]. YouTube. https://youtu.be/ntBbkOn9D_o.

Molloy, G. J., O'Carroll, R. E., & Ferguson, E. (2013). Conscientiousness and Medication Adherence: A Meta-analysis. *Annals of Behavioral Medicine*, *47*(1), 92–101. https://doi.org/10.1007/s12160-013-9524-4

Quintana, D. S. (2021). *How to perform a meta-analysis in R*. [Archivo de Vídeo]. YouTube. https://youtu.be/lH4VZMTEZSc.

Schwarzer, G., Carpenter, J. R., & Rücker, G. (2015). Small-Study Effects in Meta-Analysis. In G. Schwarzer, J. R. Carpenter, & G. Rücker (Eds.), *Meta-Analysis with R* (pp. 107–141). Springer International Publishing. https://doi.org/10.1007/978-3-319-21416-0_5

Vevea, J. L., & Hedges, L. V. (1995). A general linear model for estimating effect size in the presence of publication bias. *Psychometrika*, *60*(3), 419–435. https://doi.org/10.1007/BF02294384

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

Viechtbauer, W. (2010). Conducting Meta-Analyses inRwith themetaforPackage. *Journal of Statistical Software*, *36*(3). https://doi.org/10.18637/jss.v036.i03

doi: 10.5281/zenodo.5640182