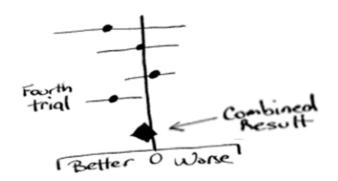




Revisão Sistemática e Meta-análise



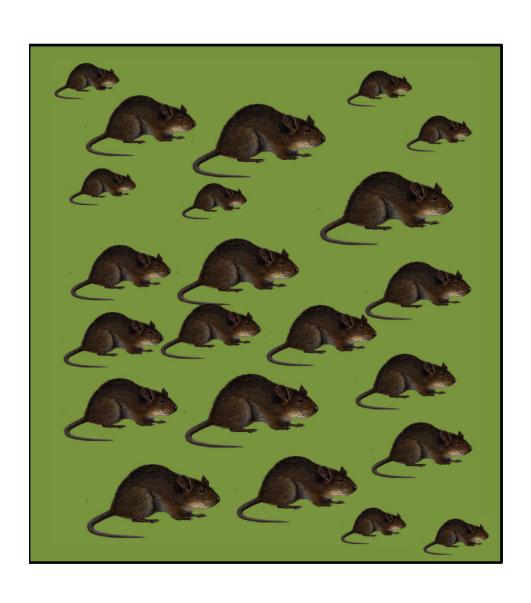
Marcelo M. Weber (mweber.marcelo@gmail.com)

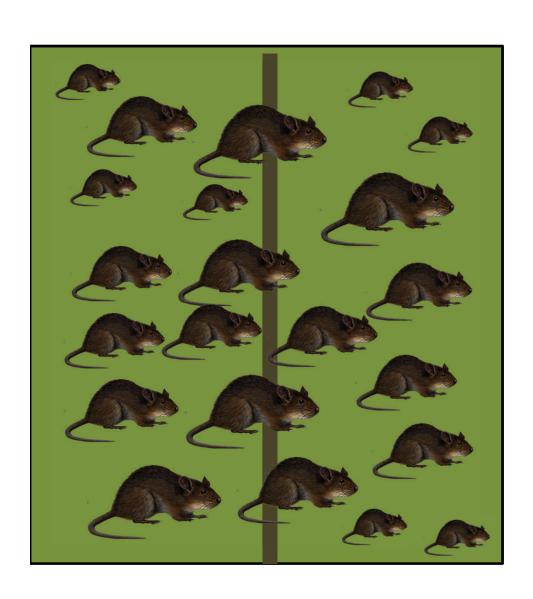
Nicholas A. C. Marino (nac.marino@gmail.com)

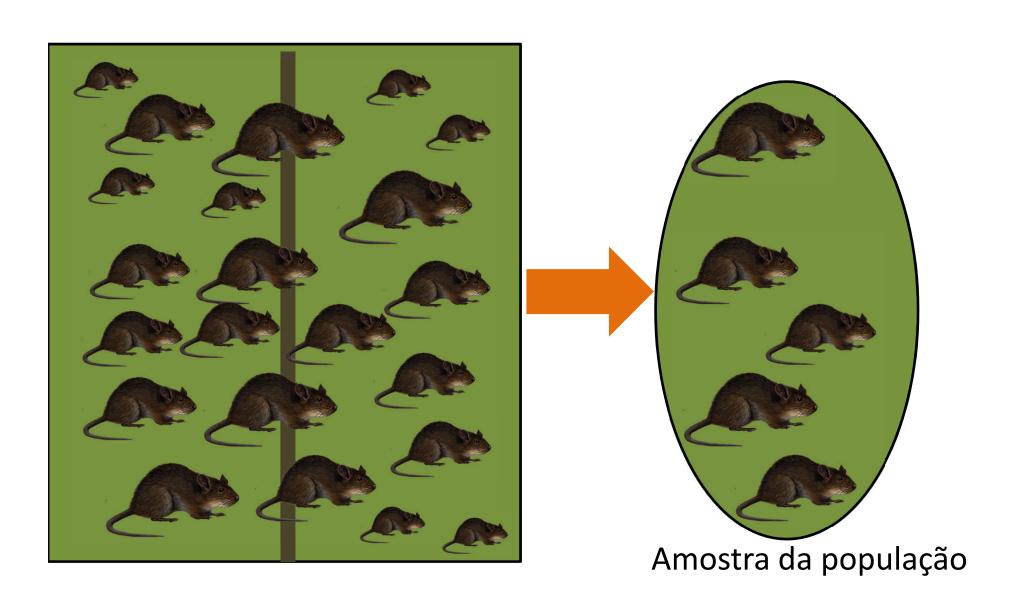
github.com/nacmarino/maR

Programa

- 1. Viés de publicação;
 - 1.1. Definição e origem;
 - 1.2. Métodos de avaliação gráfica e estatística;
- 2. Tamanhos de efeito não-independentes
 - 2.1. Origem da não-independência;
 - 2.2. Tipo de não-independência;
 - 2.3. Modelo hierárquico multivariado.











Viés de publicação ocorre sempre que a divulgação da pesquisa é tal que os TEs incluídos na meta-análise geram conclusões diferentes daquelas obtidas se todos os TEs para todos os testes conduzidos fossem incluídos na análise.



Problema de faltar estudos

- Mundo ideal: localizar todos os estudos que satisfaçam nossos critérios de busca.
- Mundo real: raramente possível.

Problema de faltar estudos

- Mundo ideal: localizar todos os estudos que satisfaçam nossos critérios de busca.
- Mundo real: raramente possível.

Se os estudos faltantes são um *subconjunto aleatório* de todos os estudos → menos informação, IC amplos e testes menos poderosos → sem impacto sistemático no TE.

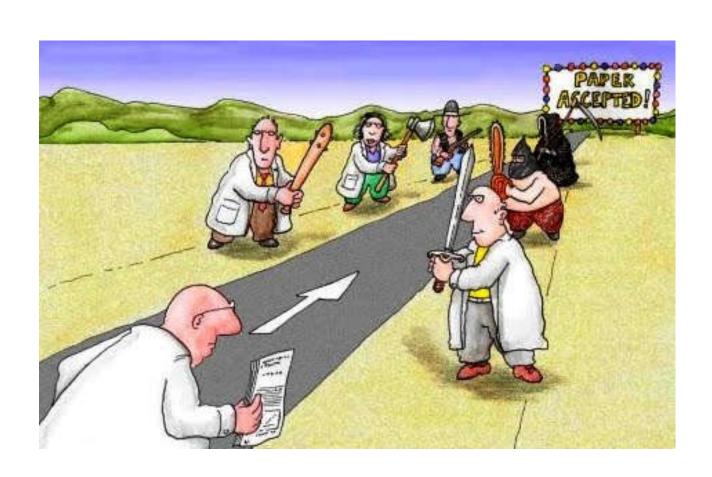
Problema de faltar estudos

- Mundo ideal: localizar todos os estudos que satisfaçam nossos critérios de busca.
- Mundo real: raramente possível.

Se os estudos faltantes são um *subconjunto aleatório* de todos os estudos → menos informação, IC amplos e testes menos poderosos → sem impacto sistemático no TE.

Se os estudos faltantes são *sistematicamente diferentes* de todos os estudos → alto impacto no TE → superestimativa.

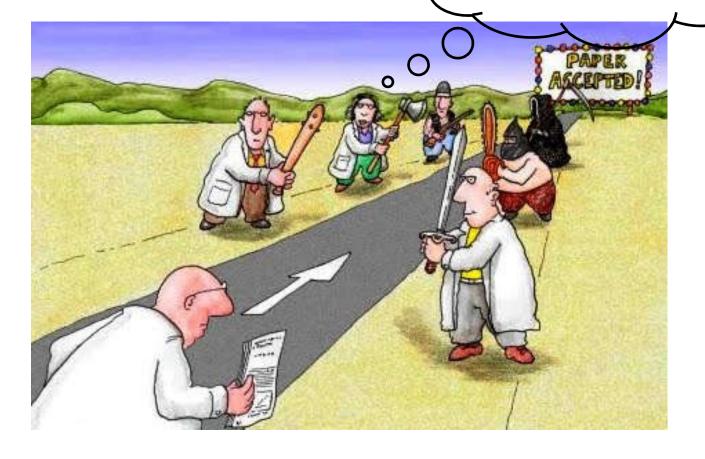
Fontes de viés







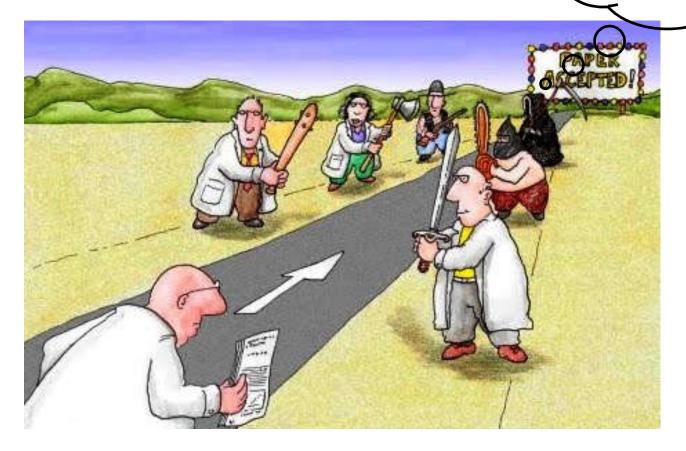
Réplicas? Poderia ter mais.



Fontes de viés

Non-significant results?

Meh!!

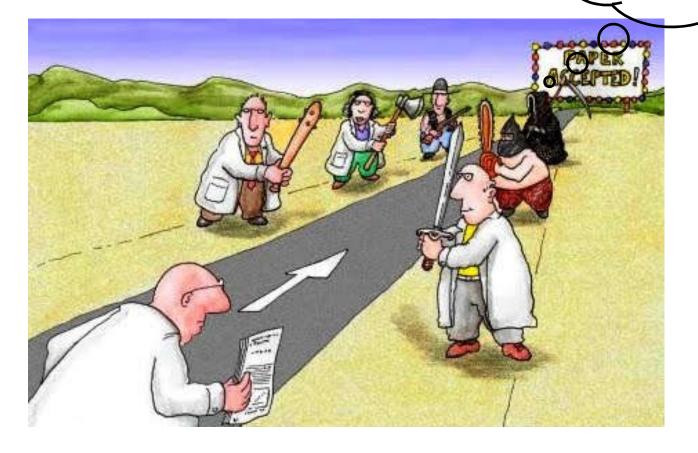


Fontes de viés

1. Resultados estatisticamente significantes são mais prováveis de serem publicados.

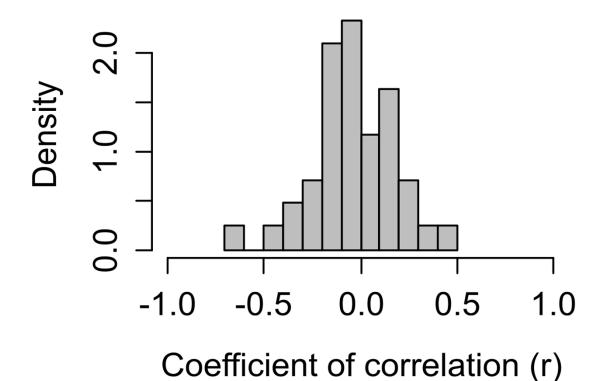
Non-significant results?

Meh!!



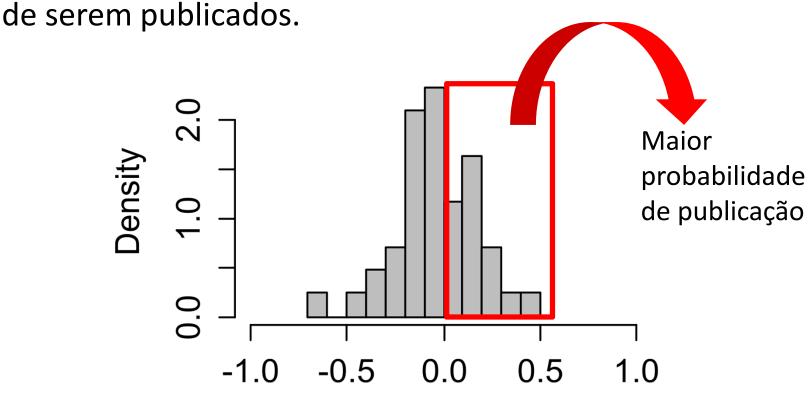
Fontes de viés

1. Resultados estatisticamente significantes são mais prováveis de serem publicados.



Fontes de viés

1. Resultados estatisticamente significantes são mais prováveis



Coefficient of correlation (r)

Fontes de viés

- 1. Resultados estatisticamente significantes são mais prováveis de serem publicados.
- Estudos com resultados não-significativos são menos prováveis de serem publicados e, quando são, o processo de publicação é mais longo;
- Alguns pesquisadores reportam seletivamente os resultados e podem, em alguns casos, mudar a hipótese pensada *a priori*.

Fontes de viés

- 2. Estudos publicados são mais prováveis de serem incluídos em uma meta-análise.
- ➤ Se alguém realizando uma revisão sistemática foi capaz de localizar estudos publicados na literatura cinza, então o fato de que estudos com maiores TE's são mais prováveis de serem publicados não seria um problema;
- > Raramente esse é o caso. Maioria dos estudos não insere literatura cinza.

Fontes de viés

2. Estudos publicados são mais prováveis de serem incluídos em uma meta-análise.

Literatura cinza: qualquer literatura produzida que não é controlada por editoras comerciais, tais como relatórios, resumos de congressos, teses, dissertações, monografias e livros.

Fontes de viés

2. Estudos publicados são mais prováveis de serem incluídos em uma meta-análise.

Literatura cinza: qualquer literatura produzida que não é controlada por editoras comerciais, tais como relatórios, resumos de congressos, teses, dissertações, monografias e livros.

Alguns sugerem que é legítimo excluir estudos que não foram publicados em revistas com revisão por pares porque esses estudos tendem a possuir baixa qualidade.

Fontes de viés

- 2. Estudos publicados são mais prováveis de serem incluídos em uma meta-análise.
- ➤ Não é óbvio que o processo de revisão por pares garante alta qualidade, nem que é o único mecanismo para selecionar os estudos.
- > Simples razão: nem todos os pesquisadores pensam em publicar em revistas especializadas:
- (i) pesquisadores vinculados a órgãos governamentais focam em produzir relatórios;
- (ii) teses e dissertações de alta qualidade podem ser improváveis de serem publicadas se o indivíduos não almeja carreira acadêmica.
- Revisão por pares pode ser enviesada, não-confiável ou de baixa qualidade.

Fontes de viés

- 3. Outros vieses
- Língua;
- > Acessibilidade;
- > Ano de publicação;
- ➤ Viés de citação;
- > Familiaridade.

Questões para refletir:

Biólogos são conhecidos por testar hipóteses nulas que eles esperam rejeitar ao invés de delinear hipóteses nulas que são refutáveis menos prontamente. Ex.: Testar se machos e fêmeas de uma espécie possuem o mesmo tamanho quando é obvio que eles diferem, ao invés de testar uma hipótese nula baseada em conhecimento prévio de que os machos são 20% maiores do que as fêmeas.

Questões para refletir:

- Biólogos são conhecidos por testar hipóteses nulas que eles esperam rejeitar ao invés de delinear hipóteses nulas que são refutáveis menos prontamente. Ex.: Testar se machos e fêmeas de uma espécie possuem o mesmo tamanho quando é obvio que eles diferem, ao invés de testar uma hipótese nula baseada em conhecimento prévio de que os machos são 20% maiores do que as fêmeas.
- Statistical fishing e selective reporting:

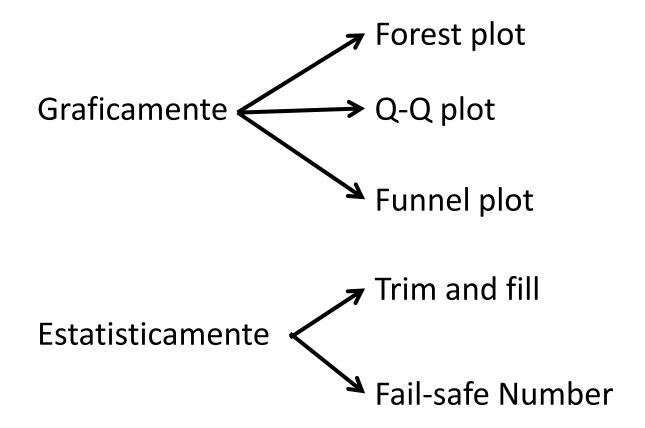


Fontes de viés

- 3. Outros vieses
- ➤ Língua;
- > Acessibilidade;
- > Ano de publicação;
- ➤ Viés de citação;
- > Familiaridade.

É possível que os estudos na metaanálise possam sobrestimar o TE verdadeiro porque eles são baseados em uma amostra enviesada da população alvo de estudos. Como lidar com essa situação?

Métodos para avaliar



Forest plot

Ranquear os tamanhos de efeito de acordo com o N

```
library(metafor)

data <- read.table("data.txt", header = T, sep = '\t')

#Calcular os tamanhos de efeito e as variâncias
model_data <- escalc(measure = "ZCOR", ri = data$correlation, ni =
data$N, data=data, method="REML")</pre>
```

Forest plot

Ranquear os tamanhos de efeito de acordo com o N

#Visualiza os tamanhos de efeito e as variâncias head(model_data)

	Study	N	correlation	yi	vi
1	Study_1	50	-0.409	-0.4344	0.0213
2	Study_2	38	-0.236	-0.2405	0.0286
3	Study_3	12	-0.346	-0.3609	0.1111
4	Study_4	14	-0.291	-0.2997	0.0909
5	Study_5	10	-0.431	-0.4611	0.1429
6	Study_6	89	-0.316	-0.3272	0.0116

```
#Gera o ajuste ao modelo selecionado (random effects)
model data rma <- rma(yi, vi, data = model data, method="REML")
#Visualiza os resultados
model data rma
  Random-Effects Model (k = 11; tau<sup>2</sup> estimator: REML)
  tau^2 (estimated amount of total heterogeneity): 0.0000 (SE = 0.0166)
  tau (square root of estimated tau^2 value):
                                                  0.0020
  I^2 (total heterogeneity / total variability):
                                                  0.01%
  H^2 (total variability / sampling variability):
                                                  1.00
  Test for Heterogeneity:
  Q(df = 10) = 8.0928, p-val = 0.6198
  Model Results:
                                pval ci.lb ci.ub
  estimate se zval
   -0.3079 0.0657 -4.6900 <.0001 -0.4366 -0.1793
                                                            常常常
  Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Forest plot

Ranquear os tamanhos de efeito de acordo com o N

```
#Transforma os resultados de volta a escala inicial
```

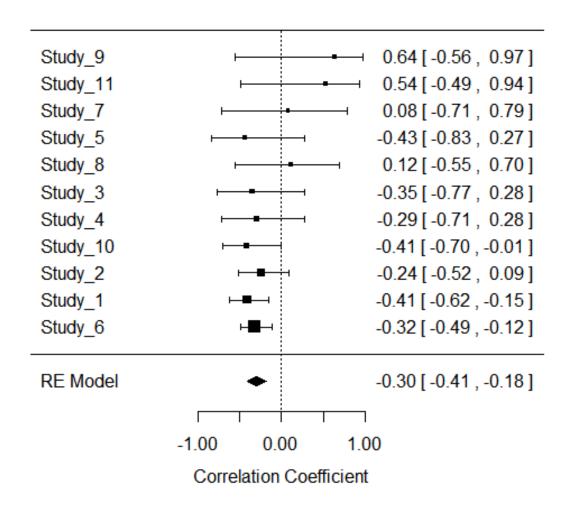
```
predict(model_data_rma, transf = transf.ztor)
```

#Plota um forest plot ranqueado baseado no N

```
forest(model_data_rma, slab = paste(data$Study), order = order(data$N), transf = transf.ztor, cex = 1)
```

Forest plot

Ranquear os tamanhos de efeito de acordo com o N



Forest plot

Examinar o forest plot de uma meta-análise cumulativa

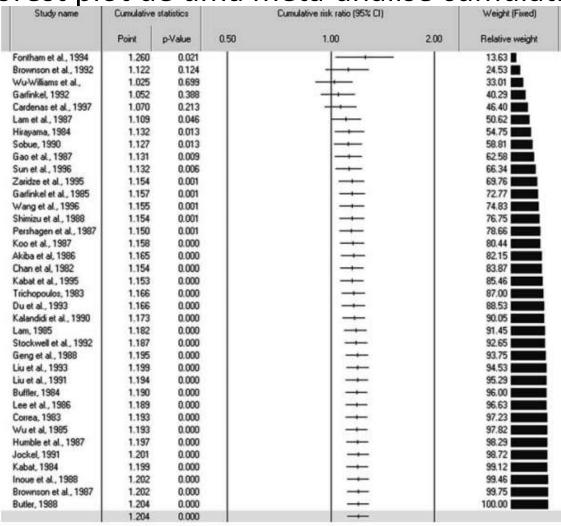
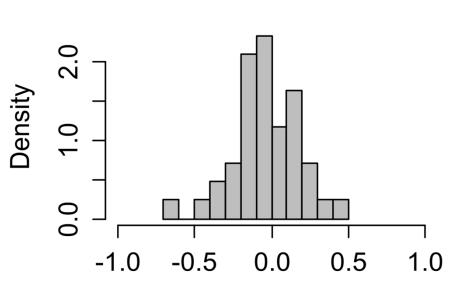


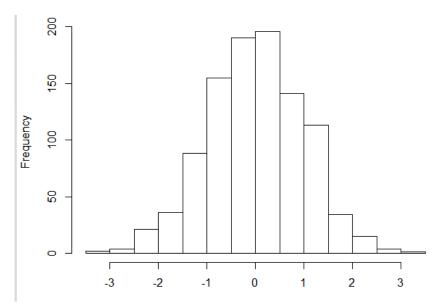
Figure 30.4 Passive smoking and lung cancer - cumulative forest plot.

Q-Q plot (quantil-quantil plot)

 Os quantis da distribuição dos dados observados são plotados contra os quantis teóricos de uma distribuição normal padrão;



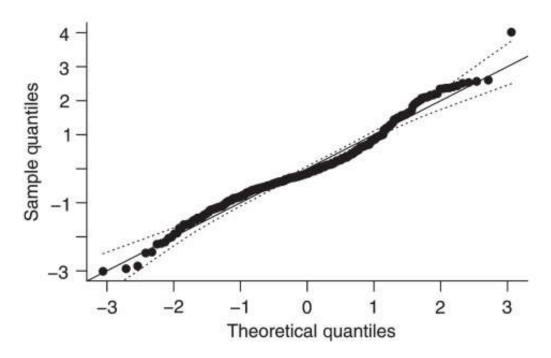
Coefficient of correlation (r)



Distribuição normal padrão

Q-Q plot (quantil-quantil plot)

- Os quantis da distribuição dos dados observados são plotados contra os quantis teóricos de uma distribuição normal padrão;
- Se os dados observados tem uma distribuição normal, os pontos cairão próximo à linha y=x



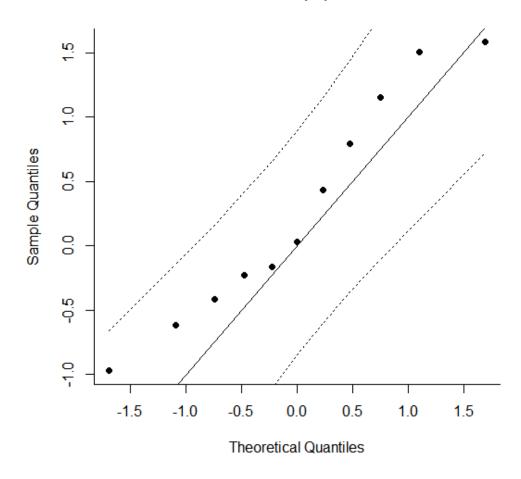
Q-Q plot (quantil-quantil plot)

```
#Plota um quantile-quantile plot
qqnorm(model_data_rma, type = "rstandard", pch = 19)
```

Q-Q plot (quantil-quantil plot)

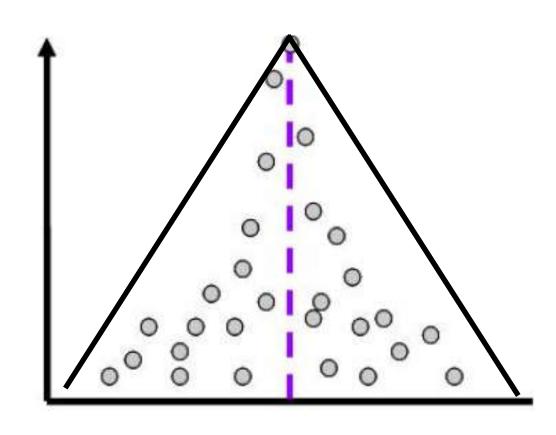
#Plota um quantile-quantile plot
qqnorm(model_data_rma, type = "rstandard", pch = 19)

Normal Q-Q Plot



Funnel plot

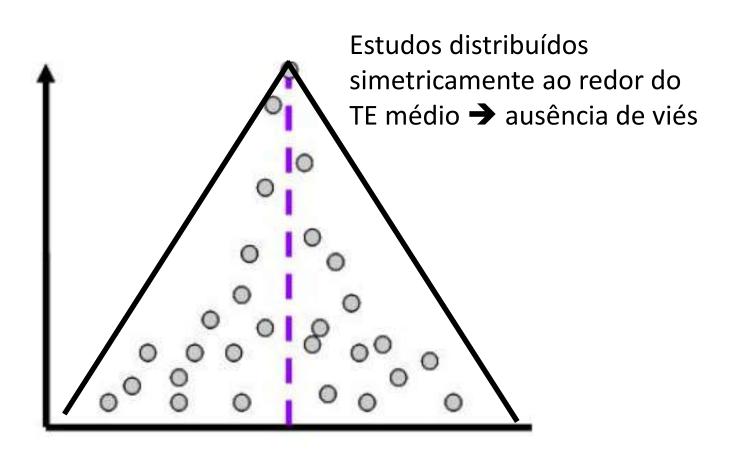
Tamanho do estudo ou variância ou erro padrão



Tamanho de efeito

Funnel plot

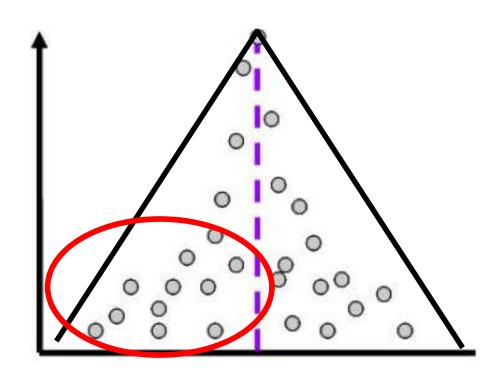
Tamanho do estudo ou variância ou erro padrão



Tamanho de efeito

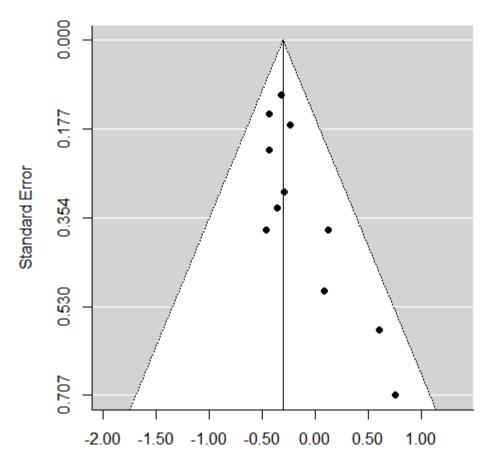
Funnel plot

Lembrando que o viés esperado é sempre em relação aos estudos com pequenos tamanho amostral e nãosignificativos. A tendência de estudos pequenos publicados é que eles tenham p<0.05 e que exibam um tamanho maior do que o esperado.



Funnel plot

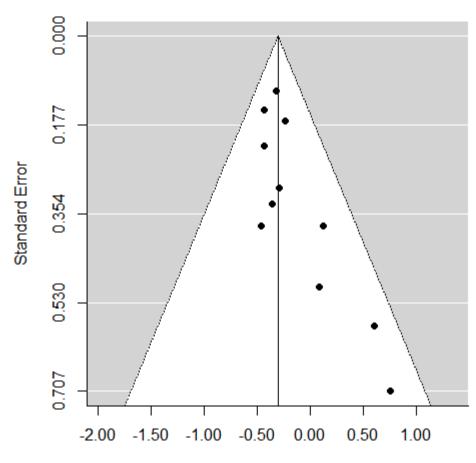
#Plota um funnel plot
funnel(x = model_data_rma, yaxis = "sei")



Fisher's z Transformed Correlation Coefficient

Funnel plot

#Plota um funnel plot
funnel(x = model_data_rma, yaxis = "sei")



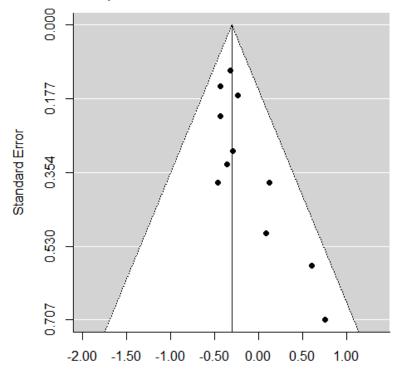
Fisher's z Transformed Correlation Coefficient

- Visualizar os *gaps*;
- Ajudam na interpretação;
- Interpretação subjetiva;
- Não são efetivos quando n<30 (assimetria pode surgir por acaso).

Funnel plot

#Calculo da significância da assimetria ranktest(model_data_rma)

Rank Correlation Test for Funnel Plot Asymmetry Kendall's tau = 0.5505, p = 0.0191



Fisher's z Transformed Correlation Coefficient

Funnel plot – Trim and Fill (apara e preenche)

- Procedimento iterativo para ajustar assimetria causada pelos estudos pequenos, produzindo estimativa não-enviesada;

1. Trim

- Remove os estudos com N muito pequeno e recalcula o TE a cada iteração até o funnel plot ser simétrico;
- Enquanto ajusta o TE, reduz a variância IC estreito.

Trim and Fill

- Procedimento iterativo para ajustar assimetria causada pelos estudos pequenos, produzindo estimativa não-enviesada;

1. Trim

- Remove os estudos com N muito pequeno e recalcula o TE a cada iteração até o funnel plot ser simétrico;
- Enquanto ajusta o TE, reduz a variância 👈 IC estreito.

2. Fill

- Adiciona os estudos originais excluídos e imputa um espelho para cada efeito excluído;
- Não impacta o TE, mas corrige a variância → IC mais largo.

Funnel plot – Trim and Fill (apara e preenche)

```
#Trim and Fill
#Ajusta o modelo trim and fill
model tf <- trimfill(model data rma)
 Estimated number of missing studies on the left side: 2 (SE = 2.3028)
 Random-Effects Model (k = 13; tau^2 estimator: REML)
 tau^2 (estimated amount of total heterogeneity): 0.0000 (SE = 0.0166)
 tau (square root of estimated tau^2 value):
                                                 0.0019
 I^2 (total heterogeneity / total variability): 0.01%
 H^2 (total variability / sampling variability):
                                                 1.00
 Test for Heterogeneity:
 Q(df = 12) = 13.1239, p-val = 0.3601
 Model Results:
 estimate
                      zval pval ci.lb ci.ub
               se
  -0.3293 0.0650 -5.0687 <.0001 -0.4566 -0.2020
                                                           经验验
 Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' '1
```

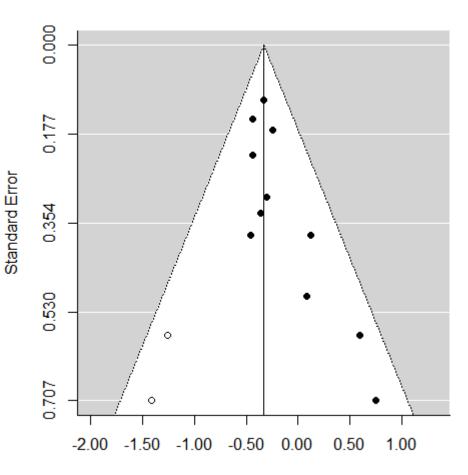
Trim and Fill

#Trim and Fill

#Ajusta o modelo trim and fill model_tf <- trimfill(model_data_rma)

#Transforma os resultados de volta a escala inicial predict(model_tf, transf = transf.ztor)

#Plota um funnel plot
funnel(model_tf)



Fisher's z Transformed Correlation Coefficient

Trim and Fill

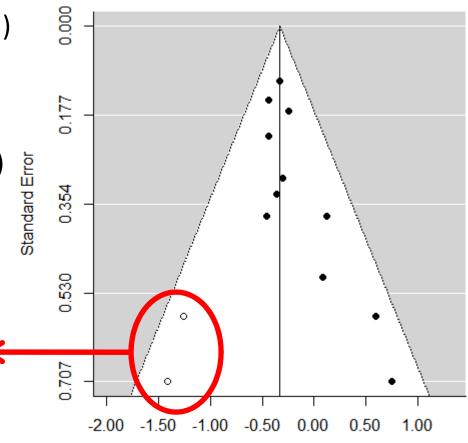
#Trim and Fill

#Ajusta o modelo trim and fill model_tf <- trimfill(model_data_rma)

#Transforma os resultados de volta a escala inicial predict(model_tf, transf = transf.ztor)

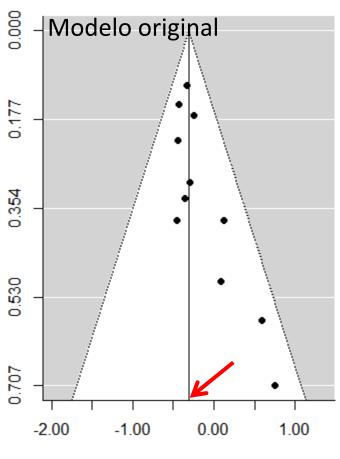
#Plota um funnel plot
funnel(model_tf)

- Estima o nº de estudos faltantes;
- Análise de sensibilidade.



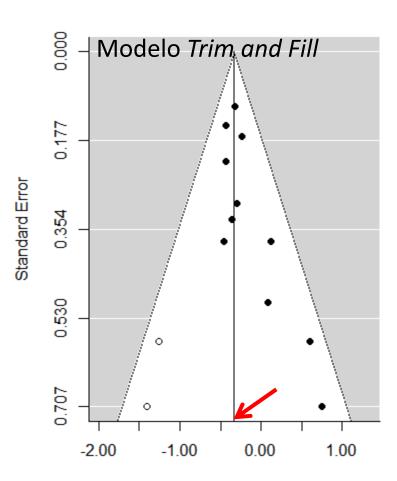
Fisher's z Transformed Correlation Coefficient

Trim and Fill



Fisher's z Transformed Correlation Coefficient

$$r = -0.31$$
; IC = [-0.44, -0.18] $r = -0.33$; IC = [-0.45, -0.20]



Fisher's z Transformed Correlation Coefficient

$$r = -0.33$$
; $IC = [-0.45, -0.20]$

Trim and Fill

- Agora compare a significância da assimetria do modelo original com o modelo *trim and fill*

#Calculo da significância da assimetria do modelo original ranktest(model_data_rma)

#Calculo da significância da assimetria do modelo trimed and filled ranktest(model_tf)

Trim and Fill

- Agora compare a significância da assimetria do modelo original com o modelo *trim and fill*

#Calculo da significância da assimetria do modelo original ranktest(model_data_rma)

#Calculo da significância da assimetria do modelo trimed and filled ranktest(model_tf)

Modelo	Kendall's tau	P-value
model_data_rma	0.5505	0.0191
Model_tf	0.1438	0.4998

Trim and Fill

- Quanto impacto tem o viés e qual seria a estimativa de TE na ausência de viés?
- Três classificações:
- Impacto trivial se todos os estudos relevantes foram incluídos, o TE permaneceria inalterado;
- 2. Impacto modesto se todos os estudos relevantes foram incluídos, o TE pode mudar, mas o achado-chave (que o efeito tem ou não importância) permaneceria provavelmente inalterado;
- **3.** Impacto substancial se todos os estudos relevantes foram incluídos, o achado-chave pode mudar.

Trim and Fill

- Interpretação um pouco diferente;
- Análise de sensibilidade: Qual é melhor estimativa de um tamanho de efeito não-enviesado?
- Se a mudança é trivial ou modesta, teremos maior confiança em dizer que o efeito é válido e robusto;
- Se a mudança é brusca, buscar possíveis fontes de variação (estudos experimentais vs. observacionais, escalas)

Rosenthal's Fail-safe N (1979)

Funnel plot: O viés tem algum impacto no efeito observado? FSN: O viés pode ser totalmente responsável pelo efeito observado?

 "Quantos estudos com um TE médio = 0 que não foram localizados (N) seriam necessários para negar a significância de um TE observado?"

Rosenthal's Fail-safe N (1979)

Funnel plot: O viés tem algum impacto no efeito observado? FSN: O viés pode ser totalmente responsável pelo efeito observado?

- "Quantos estudos com um TE médio = 0 que não foram localizados (N) seriam necessários para negar a significância de um TE observado?"
- É uma análise de sensibilidade:
- se precisamos de poucos estudos (5 ou 10) para "anular" o efeito, então devemos nos preocupar que o efeito é na verdade nulo;
- se um grande nº de estudos (20.000) é necessário para anular o TE observado, então nós estamos rejeitando H_0 corretamente.

Rosenthal's Fail-safe N (1979)

- "Quantos estudos com um TE médio = 0 que não foram localizados (N) seriam necessários para negar a significância de um TE observado?"
- Rosenthal sugere que uma meta-análise é robusta quando N > 5k + 10,

k = nº de estudos incluído na meta-análise.

Rosenthal's Fail-safe N (1979)

```
> fail <- fsn(yi = model_data$yi, vi = model_data$vi, type = "Rosenthal", alpha = 0.05)
> fail
Fail-safe N Calculation Using the Rosenthal Approach
Observed Significance Level: 0.0011
Target Significance Level: 0.05
Fail-safe N: 28

N > 5k + 10 → ausência de viés
```

$$k = 11 \rightarrow 5*11 + 10 = 65$$

$$N = 28 \rightarrow N < 5k + 10$$

28 > 65? → detecção de viés

Rosenthal's Fail-safe N (1979)

- > Críticas:
- 1. Foca na significância estatística (p<0.05) ao invés de significância substancial;
- 2. Assume que o TE médio nos estudos "escondidos" é zero para anular o efeito (quando pode ser nulo ou negativo);
- Método bem conhecido (sugerido por Koricheva e dispensando por Borenstein);
- Valor histórico.

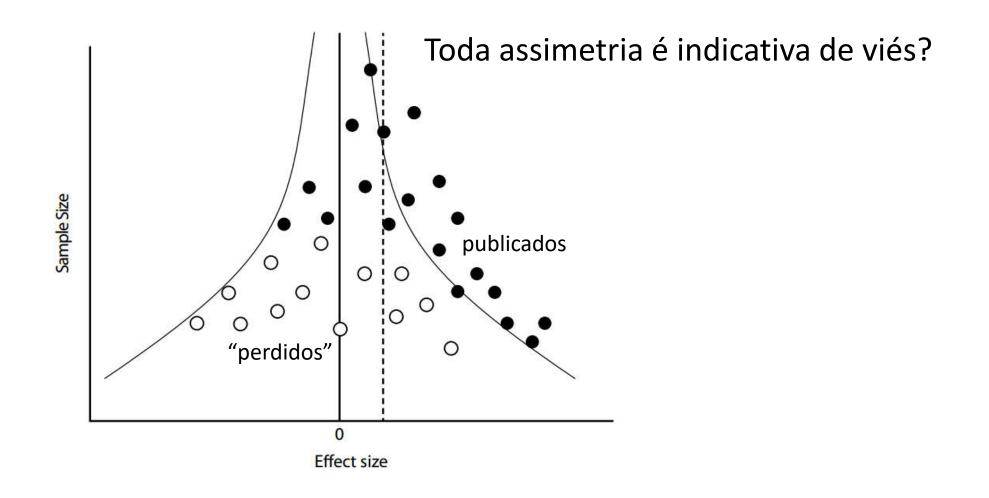
Orwin's Fail-safe N (1983)

- Considera as críticas anteriores;
- Pergunta: Quantos estudos "perdidos" poderiam trazer o
 TE estimado para um nível específico que não seja zero,
 mas que não seja um efeito substancial?

Orwin's Fail-safe N (1983)

- Considera as críticas anteriores;
- Pergunta: Quantos estudos "perdidos" poderiam trazer o
 TE estimado para um nível específico que não seja zero,
 mas que não seja um efeito substancial?
- Pesquisador define a priori um valor que representa o menor efeito possível de importância substancial e pergunta quantos estudos seriam necessários para trazer o efeito abaixo desse limiar? Ex.: r = 0.20, ORR=1.05.
- Normalmente Orwin's FSN < Rosenthal's FSN.

Identificação errônea de viés de publicação



Identificação errônea de viés de publicação

- Heterogeneidade do estudo: assimetria pode surgir sempre que um conjunto heterogêneo de TE é combinado;
- Estudos observacionais vs. experimentais: TE de estudos observacionais tem a ser menores do que estudos experimentais;
- 3. Heterogeneidade dependente de táxon ou sistema: TE podem ser maiores em alguns sistemas/táxons do que outros.

- Dentro de estudos:
- 1. Mais de um TE por estudo.

- Dentro de estudos:
- 1. Mais de um TE por estudo.
- Entre estudos:
- 1. Estudos diferentes conduzidos na mesma área;

- Dentro de estudos:
- 1. Mais de um TE por estudo.
- Entre estudos:
- 1. Estudos diferentes conduzidos na mesma área;
- 2. Estudos diferentes realizados com a mesma espécie;

- Dentro de estudos:
- 1. Mais de um TE por estudo.
- Entre estudos:
- 1. Estudos diferentes conduzidos na mesma área;
- 2. Estudos diferentes realizados com a mesma espécie;
- 3. Estudos diferentes conduzidos pelos mesmos pesquisadores;

- Dentro de estudos:
- 1. Mais de um TE por estudo.
- Entre estudos:
- 1. Estudos diferentes conduzidos na mesma área;
- 2. Estudos diferentes realizados com a mesma espécie;
- 3. Estudos diferentes conduzidos pelos mesmos pesquisadores;
- 4. Relações filogenéticas entre espécies.

Como a não-independência pode ser abordada?

 Excluir múltiplas estimativas e/ou focar apenas em uma única variável resposta;

Como a não-independência pode ser abordada?

- Excluir múltiplas estimativas e/ou focar apenas em uma única variável resposta;
- 2. Assumir (erroneamente) que todos os efeitos são independentes;

Como a não-independência pode ser abordada?

- Excluir múltiplas estimativas e/ou focar apenas em uma única variável resposta;
- 2. Assumir (erroneamente) que todos os efeitos são independentes;
- 3. Uso de modelos hierárquicos multivariados (modelo multinível ou aninhado)

Α	В	С	D	E
Study	N	correlation	group	author
Study_1	50	-0.409 A		Diniz
Study_2	38	-0.236	-0.236 A	
Study_3	12	-0.346	Α	Diniz
Study_4	14	-0.291	Α	Diniz
Study_5	10	-0.431	Α	Diniz
Study_6	89	-0.316	Α	Diniz
Study_7	7	0.084	В	Araujo
Study_8	10	0.119	В	Araujo
Study_9	5	0.637	В	Araujo
Study_10	24	-0.411	Α	Diniz
Study_11	6	0.536	В	Araujo
Study_12	9	-0.6	В	Diniz
Study_13	44	-0.513	Α	Diniz

```
hier <- read.table("data_hier.txt", header = T, sep = '\t')

#Calcular os tamanhos de efeito e as variâncias
dat <-escalc(measure = "ZCOR", ri = hier$correlation, ni = hier$N, data
= hier, method="REML")

#Criar um modelo não-hierárquico
res.dat.NH <- rma(yi = dat$yi, vi, data = dat, method="REML")
```

```
> res.dat.NH
Random-Effects Model (k = 13; tau<sup>2</sup> estimator: REML)
tau^2 (estimated amount of total heterogeneity): 0.0000 (SE = 0.0146)
tau (square root of estimated tau^2 value):
                                            0.0014
I^2 (total heterogeneity / total variability): 0.00%
H^2 (total variability / sampling variability): 1.00
Test for Heterogeneity:
Q(df = 12) = 11.1314, p-val = 0.5177
Model Results:
                              pval ci.lb ci.ub
estimate
                     zval
              se
 -0.3543 0.0599 -5.9172 <.0001 -0.4716 -0.2369
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#Criando o modelo hierárquico multivariado
res.dat <- rma.mv(yi = dat$yi, vi, random = ~1|author, data = dat, method
= "REML")
      > res.dat
      Multivariate Meta-Analysis Model (k = 13; method: REML)
      Variance Components:
                 estim sqrt nlvls fixed factor
      sigma^2 0.1930 0.4393 2
                                        no
                                            author
      Test for Heterogeneity:
      Q(df = 12) = 11.1314, p-val = 0.5177
      Model Results:
      estimate
                          zval
                                  pval ci.lb ci.ub
                   se
       -0.1002 0.3334 -0.3005 0.7638 -0.7537 0.5533
      Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' '1
```

Modelo	estimate	se	pval	ci.lb	ci.ub
res.dat.NH	-0.3543	0.0599	<.0001	-0.4716	-0.2369
res.dat	-0.1002	0.3334	0.7638	-0.7537	0.5533

Resumo

- 1. O que é viés e as suas causas;
- 2. Como podemos reportar o viés: graficamente e/ou estatisticamente;
- 3. Método trim and fill e fail-safe number;
- 4. TE não-independentes;
- 5. Uso de modelos hierárquicos multivariados.