

# Modelos para Meta-Análise

Revisão Sistemática e Meta-Análise

Marcelo M. Weber & Nicholas A. C. Marino

[github.com/nacmarino/maR](https://github.com/nacmarino/maR)

# Recapitulando

- **Meta-Análise:** "é a análise estatística de uma ampla coleção de resultados de estudos com o propósito de integrar a evidência disponível". (*Glass, 1976*)
- É essencial extrairmos uma métrica de tamanho de efeito e também a sua variância, para que o peso de cada estudo seja proporcional à sua precisão.
- Com estas informações, queremos chegar a um valor que defina a direção e/ou magnitude de um efeito/padrão/processo, a incerteza associada a este valor e fontes que expliquem esta heterogeneidade.

# Modelos para Meta-Análise

- Existem diversos tipos de modelo para serem usados, cuja escolha:
  - **Reflete o objetivo da meta-análise:** estimar um efeito vs explorar heterogeneidade;
  - **Define como as estimativas de cada estudo serão combinadas:** todos os estudos são iguais vs estudos diferem em sua 'qualidade';
  - **Descreve as fontes de heterogeneidade entre estudos:** erro amostral vs variação entre estudos;
  - **Determina o tipo de inferência que pode ser feita:** população de estudos analisada vs todos os estudos, até os desconhecidos.
- Desta forma, a escolha do modelo reflete o(s) objetivo(s) da sua meta-análise;
- No geral, os modelos de meta-análise se dividem em dois grandes grupos: *fixed-effects models* e *random-effects models*.

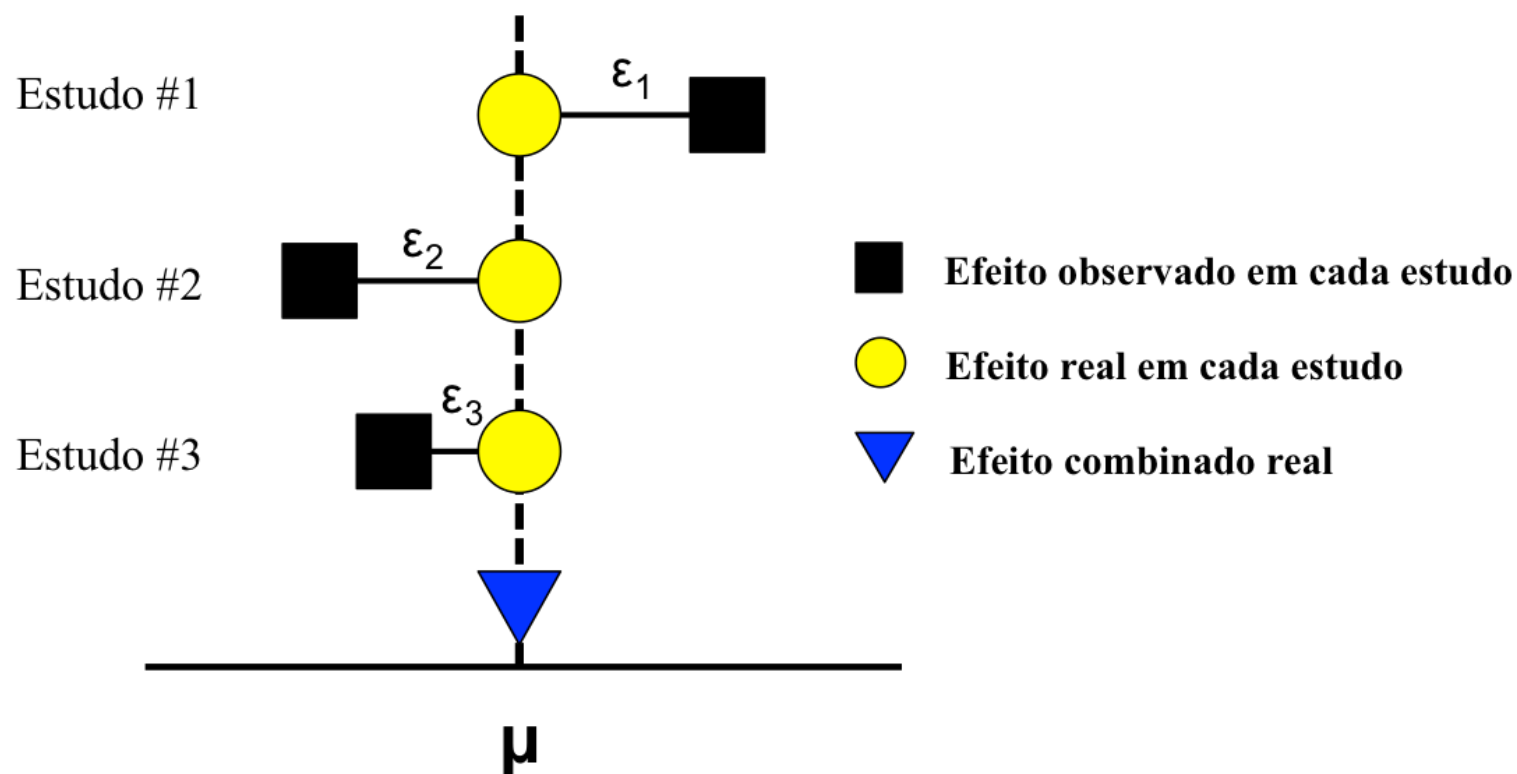
# Fixed-effects model

- Usado para estimar a natureza de um efeito;
- Baseado no pressuposto de normalidade;
- Em um fixed-effects model, nós assumimos que:
  - Todos os estudos pertencem a uma mesma população e estão medindo a mesma coisa.
  - A variabilidade existente é explicada unicamente por erro na amostragem em cada estudo;
  - Varibilidade causada por diferenças entre estudos é ínfima ou inexistente.

# Fixed-effects model

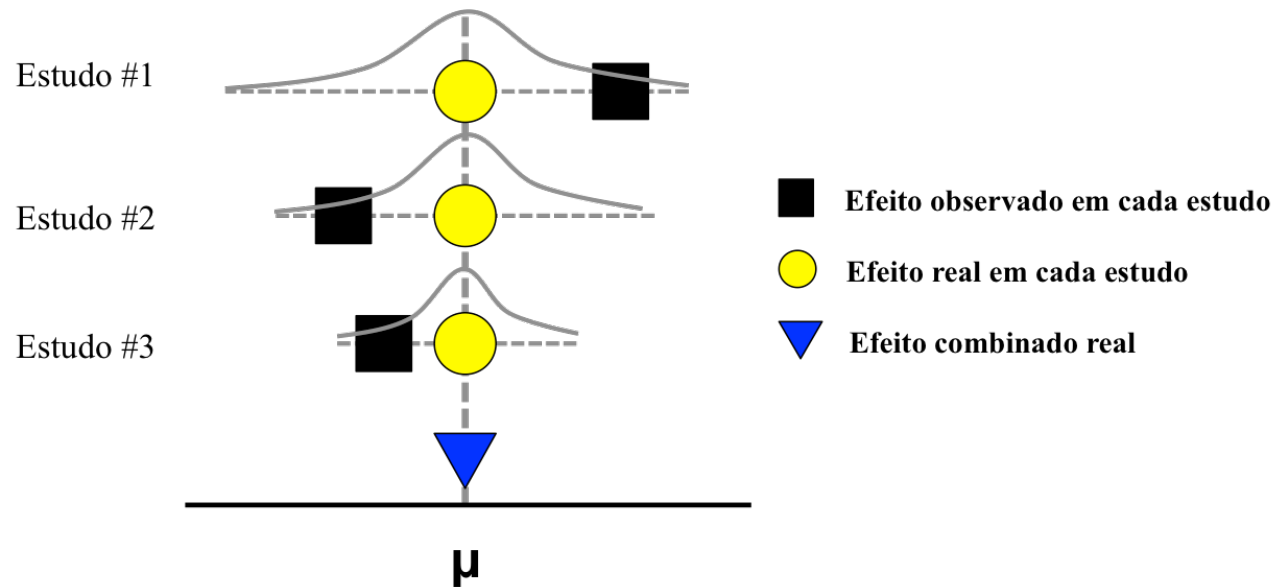
- Exemplo de fixed-effects model:

$$\mu = T_i + \epsilon_i$$



# Fixed-effects model

- A incerteza na estimativa de cada estudo está contida dentro da variabilidade do efeito comum que estes estão medindo.
- Generalizações são restritas aos estudos incluídos na meta-análise.



# Fixed-effects models

- Cada estudo então contribui com uma estimativa do efeito real, que deve ser ponderado pela sua imprecisão ao medir este efeito.
- No geral, o peso de cada estudo ( $w_i$ ) em uma meta-análise é proporcional ao inverso de sua variância ( $s_i$ ): quanto maior variância de um estudo, menor o seu peso ao estimar o efeito real.

$$w_i = \frac{1}{s_i}$$

- Neste sentido, o modelo de meta-análise é semelhante à uma *weighted regression*: o efeito real é calculado com base em cada medida de effect size ( $ES_i: d_i, LRR_i, r_i, \dots$ ) e o peso associado à ele ( $w_i$ ).

$$\mu = \frac{\sum ES_i \cdot w_i}{\sum w_i} \text{ (onde } \mu, \text{ é o efeito comum entre todos os estudos)}$$

# Fixed-effects models

- A variância em torno de  $T$  ( $V_T$ ) é dada por:

$$V_T = \frac{1}{\sum w_i}$$

- Com isto, é possível calcular o intervalo de confiança ao redor de  $\mu$ :

$$95\% \text{ CI} = \mu \pm 1.96 \sqrt{V_T}$$

- Também podemos testes de significância com o valor estimado de  $\mu$  (que segue uma distribuição normal):

$$Z = \frac{\mu}{\sqrt{V_T}}$$

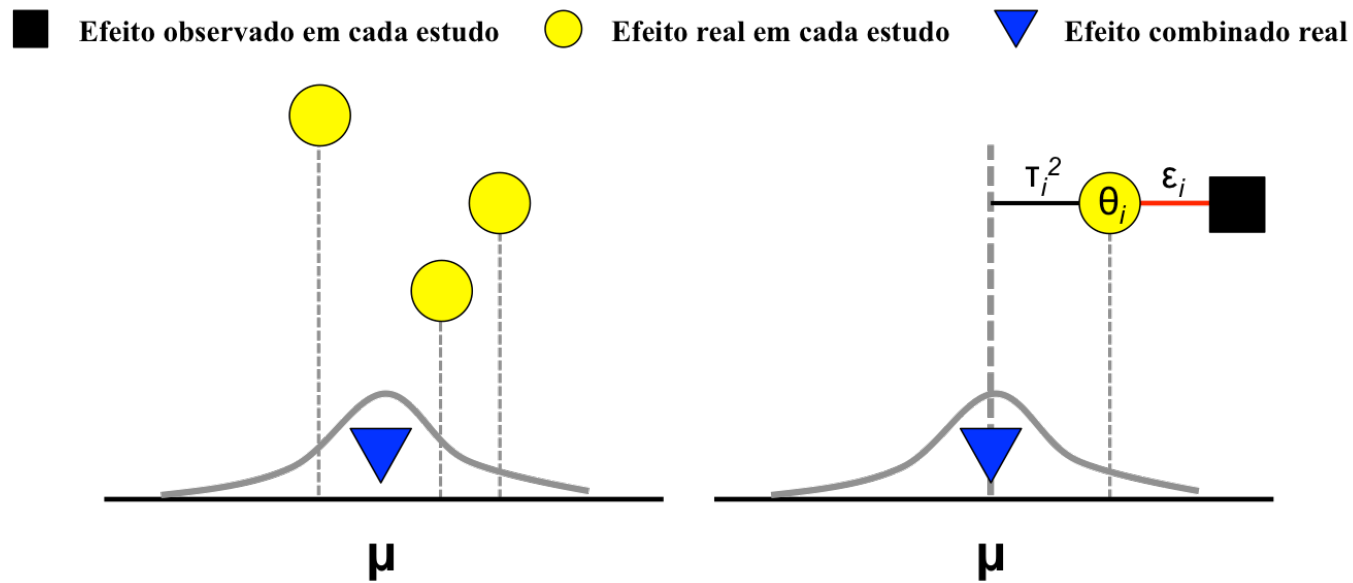


# Random-effects model

- Usado para estimar a natureza de um efeito;
- Baseado no pressuposto de normalidade;
- Em um random-effects model:
  - Estudos pertencem a diferentes subpopulações que compõem uma população maior: cada estudo está estimando o verdadeiro efeito em sua subpopulação.
  - A variabilidade existente é explicada por erro na amostragem em cada estudo e por diferenças entre estudos.

# Random-effects model

- Em um random-effects model, temos dois componentes de variância:
  - Variância dentro dos estudos:  $\epsilon_i$  (**within-study variance**)
  - Variância entre os estudos:  $\tau_i^2$  (**between-study variance**)
- Você é capaz de explorar onde existe maior variância nos resultados: dentro ou entre estudos - em outras palavras: *consistência* ou *contingência*.



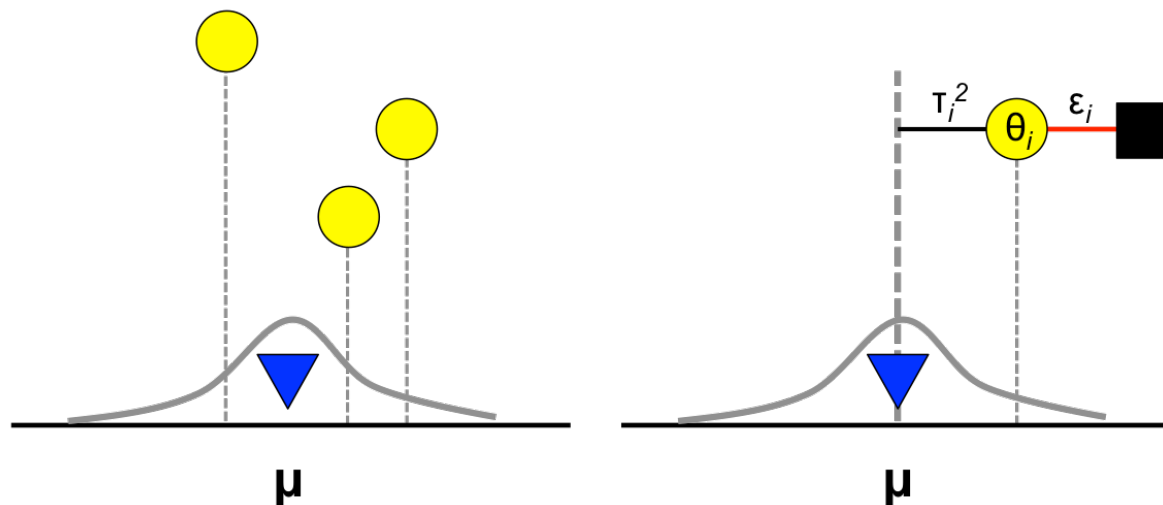
# Random-effects model

- Em um random-effects model:

$$\tau_i = \theta_i + \epsilon_i \text{ (efeito real em cada estudo)}$$

$$\mu = \tau_i + \tau_i^2 \text{ (efeito combinado real)}$$

■ Efeito observado em cada estudo    ● Efeito real em cada estudo    ▼ Efeito combinado real



# Random-effects model

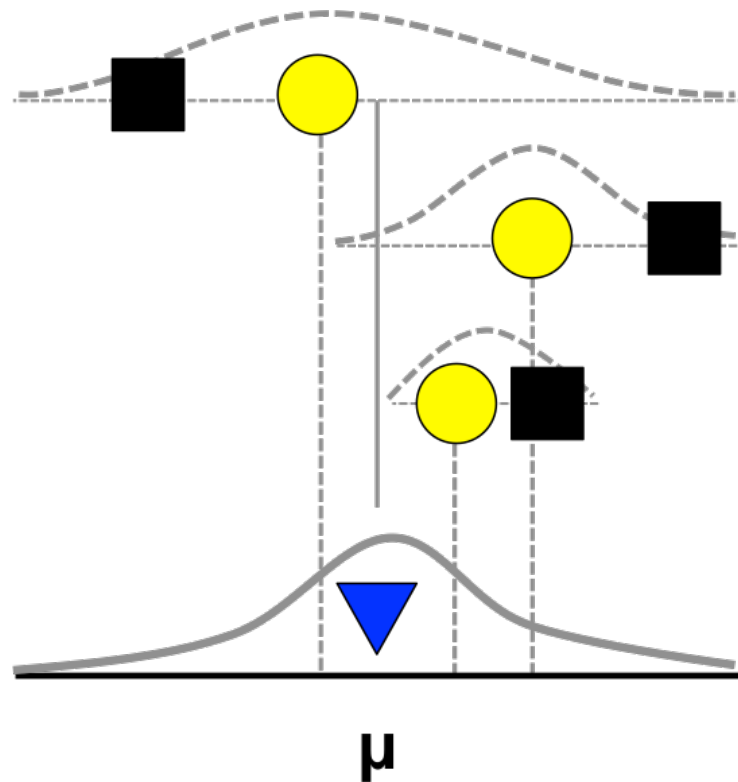
- O efeito comum que os estudos estão medindo contém a variabilidade entre estudos e também dentro de cada estudo.
- Isto afeta diretamente a estimativa do peso de cada estudo no modelo da meta-análise:

$$w_i = \frac{1}{s_i + \tau^2}$$

- Aqui,  $\tau^2$  representa o componente da heterogeneidade total que é originária entre estudos (resolução dos cálculos para chegar a este valor está no Box 9.2, Capítulo 9 do livro de Meta-Análise).

# Random-effects model

- Generalizações podem ser feitas aos estudos incluídos na meta-análise e também aqueles que são desconhecidos.



# Fixed- vs Random-effects model

- Qual modelo utilizar?
  - RE é mais realístico à meta-análises em ecologia, uma vez que é impossível replicar perfeitamente os experimentos e sistemas de estudos são naturalmente diferentes;
  - No entanto, RE requer a estimativa de um parâmetro adicional relacionado à variância entre estudos, que pode ter estimativa enviesada por pequeno tamanho amostral ou número de estudos - uso de REML reduz um pouco este problema;
  - Por conta disso, intervalos de confiança estimados com RE models tendem a ser mais amplos do que aqueles estimados com FE models.

# Fixed- vs Random-effects model

- Note que caso o  $\tau^2$  seja 0 (ou muito próximo à zero), a estimativa do peso de cada estudo é reduzida àquela de um fixed-effects model;
- Por conta disso, se você precisa, o ideal é tentar um random-effects model e caso o valor de  $\tau^2$  seja baixo aí sim adotar um fixed-effects model.

Modelo	Estimate	SE	CI.Inferior	CI.Superior
Fixed-effects model	-0.43	0.04	-0.51	-0.35
Random-effects model	-0.71	0.18	-1.07	-0.36

# Mixed-effects models

- É utilizado quando queremos não só estimar um efeito, mas também explorar a heterogeneidade ao redor dele;
- Acomoda uma fração aleatória da heterogeneidade (random-effects) - causada por variância entre- e dentro- dos estudos -, assim como uma fração fixa (característica dos estudos que queremos explorar);
- Neste sentido, ao invés de 'jogar' toda a heterogeneidade entre estudos para um termo de variância, parte dela é atribuída a *moderadores* (veremos mais sobre isso na próxima aula);
- No R, basta usarmos a função `rma` no `metafor` e especificar a variável que queremos analisar com o argumento `mods`;
- Você também pode buscar explorar a heterogeneidade ao redor da estimativa do efeito utilizando um fixed-effects model: basta especificar o argumento `method = "FE"` na mesma estrutura do modelo abaixo.



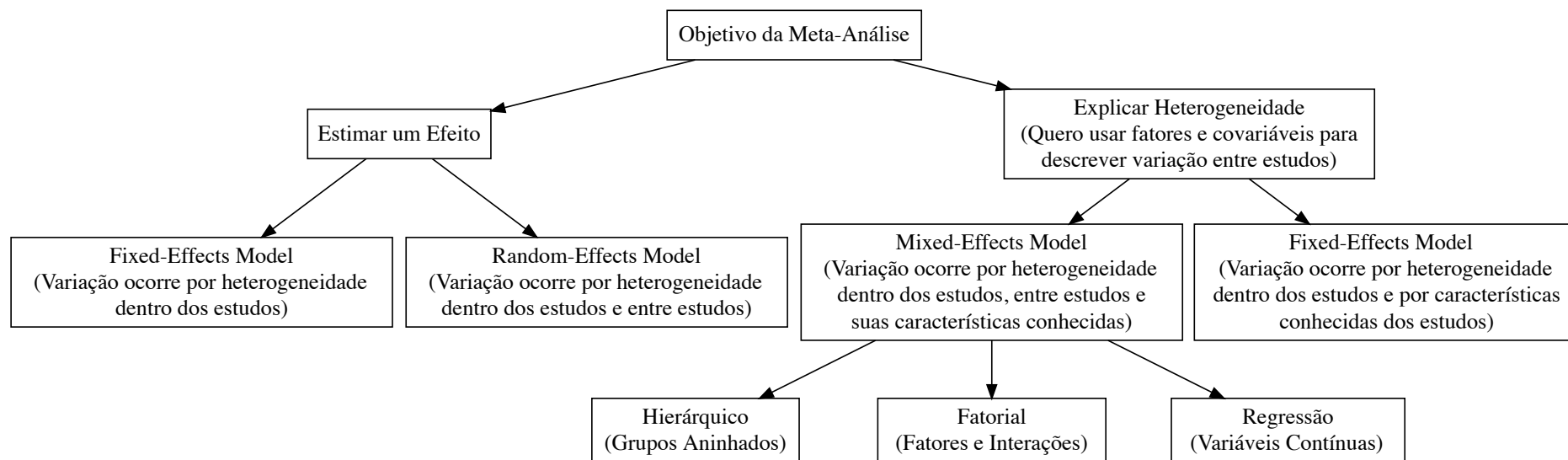
# Outros modelos mais complexos

- **Meta-Regressão:** é mais uma terminologia para representar os modelos em que a heterogeneidade é explicada por variáveis contínuas (embora algumas pessoas também usem este termo quando usam variáveis categóricas);
- **Modelo Hierárquico ou Multivariado:** usado quando estrutura é similar à uma ANOVA aninhada ou quando existe correlação entre as observações. Em alguns casos, você acaba fazendo este tipo de modelo de forma artificial para explorar heterogeneidade em subgrupos dentro dos grupos testados (no R: `rma.mv`);
- **Modelo Fatorial:** usado quando a meta-análise segue um desenho fatorial;
- **Modelo via Generalized Linear (Mixed-Effects) Models:** similar aos fixed-, random- e mixed-effects model, mas permite usar outras famílias de distribuição de variáveis aleatórias (no R: `rma.glmm`).

# Observações

- Outros pacotes para rodar os modelos de meta-análise: `nlme` e `lme4`.
- Uma alternativa ao uso do inverso da variância como o peso de cada observação é usar o *tamanho amostral* de cada estudo: estudos maiores têm mais peso do que os estudos menores, pois tendem a ser mais precisos.
- Recomenda-se que os intervalos de confiança para a estimativa do efeito também sejam calculados através do método do *bootstrap* paramétrico ou não-paramétrico (ver pacote `boot`).
  - Paramétrico: 'cria' populações de effect size baseado na estimativa do efeito e da variância pelo modelo;
  - Não-paramétrico: 'cria' populações de effect size baseado nas observações existentes no conjunto de dados utilizado na meta-análise.
- Você também pode usar o *bootstrap* para calcular intervalos de confiança para cada nível de moderadores categóricos, enquanto que para moderadores contínuos utilizar o *bootstrap* pode ser um pouco mais complicado.

# Só para ficar didático



# Resumindo

- Existem diversos tipos de modelos que podem ser usados em uma meta-análise, cada um com um tipo de característica e uso próprio;
- Não existe o melhor modelo para a sua meta-análise, mas sim o modelo que descreve melhor seus objetivos e perguntas;
- Independente do modelo que você use, é boa prática ponderar cada observação/estudo pelo inverso de sua variância.

# Literatura Recomendada

1. Nakagawa & Santos, 2012, *Evol Ecol*, Methodological issues and advances in biological meta-analysis
2. Harrison, 2011, *Methods Ecol Evol*, Getting started with meta-analysis
3. Mengersen et al, 2013, Statistical models and approaches to inference, In: *Handbook of meta-analysis in ecology and evolution* (Capítulo 8)
4. Rosenberg, 2013, Moment and least-squares based approaches to meta-analytic inference, In: *Handbook of meta-analysis in ecology and evolution* (Capítulo 9)
5. Mengersen & Schmid, 2013, Maximum likelihood approaches to meta-analysis, In: *Handbook of meta-analysis in ecology and evolution* (Capítulo 10)