How to calculate and combine effect-sizes?

September 2024

Damien Beillouin
CIRAD-Hortsys

Pooling Effect Sizes

Pooling effect sizes is essential for synthesizing results across studies in meta-analysis. Different models can be employed based on the data structure and assumptions.

Fixed-Effect Model

The **Fixed-Effect Model** assumes that all studies estimate the same underlying effect size.

Key Features:

- Assumes homogeneity of effect sizes across studies.
- Only one common effect size is estimated.

Formula:

$$\hat{\theta}_k = \theta + \varepsilon_k, \quad \varepsilon_k \sim \mathcal{N}(0, s_k^2)$$

Fixed-Effect Model

Weighting:

$$\bar{\theta}_w = \frac{\sum_{i=1}^{k} w_i \hat{\theta}_i}{\sum_{i=1}^{k} w_i}$$

Where:

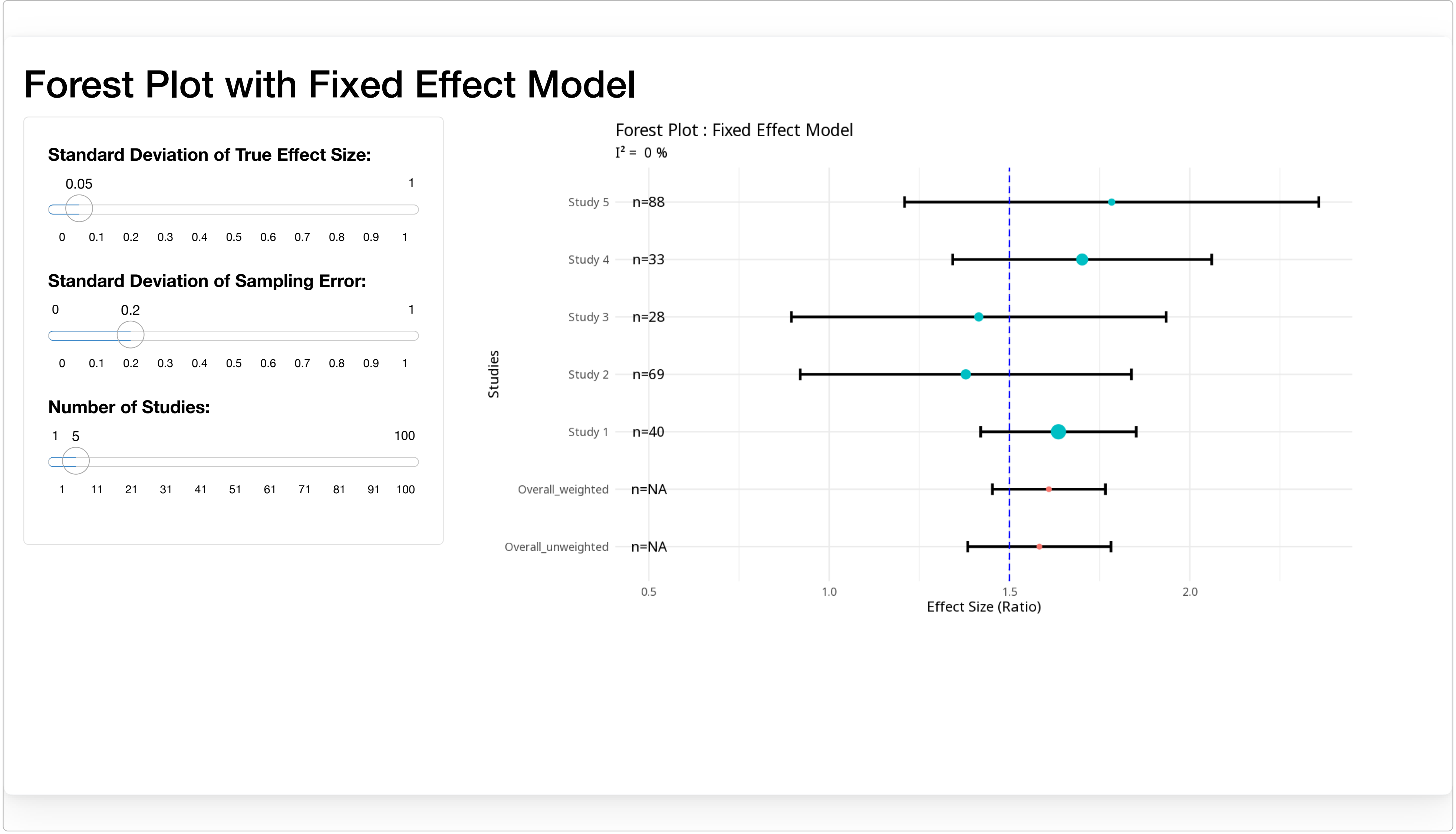
- bar_theta_w is the **weighted average effect size** across (k) studies.
- hat_theta_i is the **estimated effect size** for the (i)-th study.
- w_i is the **weight** for the (i)-th study, calculated as (w_i =).
- v_i is the **variance** of the (i)-th study.

  and with weights equal to wi=1/vi

Fixed-Effect Model

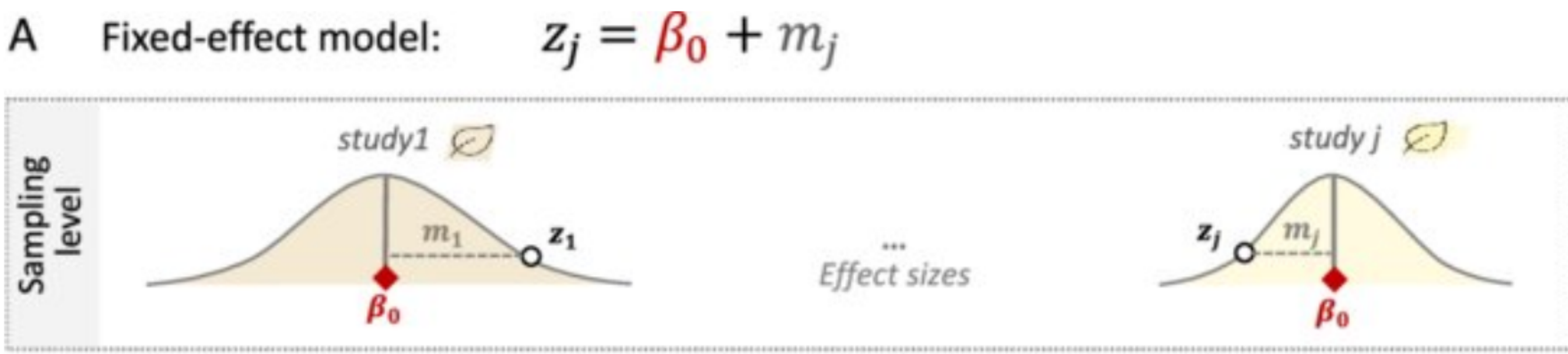
Let's say: trueRR=1.5; and

$$\hat{\theta}_k = \theta + \varepsilon_k$$



How to calculate and combine effect-sizes?

Fixed-Effect Model



Limitations of Fixed-Effect Models

- The outcome of interest could have been measured in many ways.
- The type of treatment may not have been exactly the same.
- The intensity and duration of treatment could differ.
- The target population of the studies may not have been exactly the same for each study.
- The control groups used may have been different.

Random-Effects Model

The **Random-Effects Model** accounts for variability among studies, assuming that effect sizes vary due to different study conditions.

Formula:

$$\hat{\theta}_k = \mu + \zeta_k + \epsilon_k; \quad \zeta_k \sim N(0, \mathscr{T}_k^2); \quad \epsilon_k \sim N(0, s_k^2)$$

- hat_theta_k: Estimated effect size for study (k).
- mu : The overall mean (or meta-analytic mean).
- ζ_k: Random effect for study (k),$\mathcal{T}$^2 is the between-study variance.
- epsilon_k: error term, sk2is the within study variance

Random Model Estimation

- The weights are defined as:

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

Where:

- (_k): Estimated effect size for study (k).
- ( ): The overall mean (or meta-analytic mean).
- (_k (0, ^2)): Random effect for study (k), where (^2) is the between-study variance.
- (_k (0, v_k)): Within-study error term for study (k), where (v_k) is the variance of the effect size estimate for study (k).

Heterogeneity : I^2 Statistic

- **Definition**:
  - I^2 (I-squared) is a measure of the percentage of variation across studies that is attributable to heterogeneity rather than chance.
- **Importance**:
  - Understanding I^2 helps researchers assess the consistency of results across studies.
  - It informs decisions on whether to use fixed or random effects models in meta-analyses.
- **Calculation**:
  - I^2 is calculated from the Q statistic:

$$I^2 = 100\% \times \frac{\tau^2}{\tau^2 + s_k^2}$$

Where: $\mathcal{T}$^2 is the between-study variance. sk2is the within study variance

I^2 Statistic

- **Considerations**:
  - High I^2 values do not always imply that meta-analysis is inappropriate.
  - Subgroup analyses and meta-regressions can help explore sources of heterogeneity.
- **Interpretation**:
  - **0%**: No heterogeneity. All studies have similar effects.
  - **1-25%**: Low heterogeneity. Minor variations among studies.
  - **25-50%**: Moderate heterogeneity. Some differences in effects.
  - **50-75%**: Substantial heterogeneity. Notable variations among studies.
  - **75-100%**: Considerable heterogeneity. Major differences in effects.

Mixed Model Estimation

Let's say: trueRR=1.5; and $\hat{(\theta\_k)}=\mu+\zeta\_k+\varepsilon\_k$

How to calculate and combine effect-sizes?

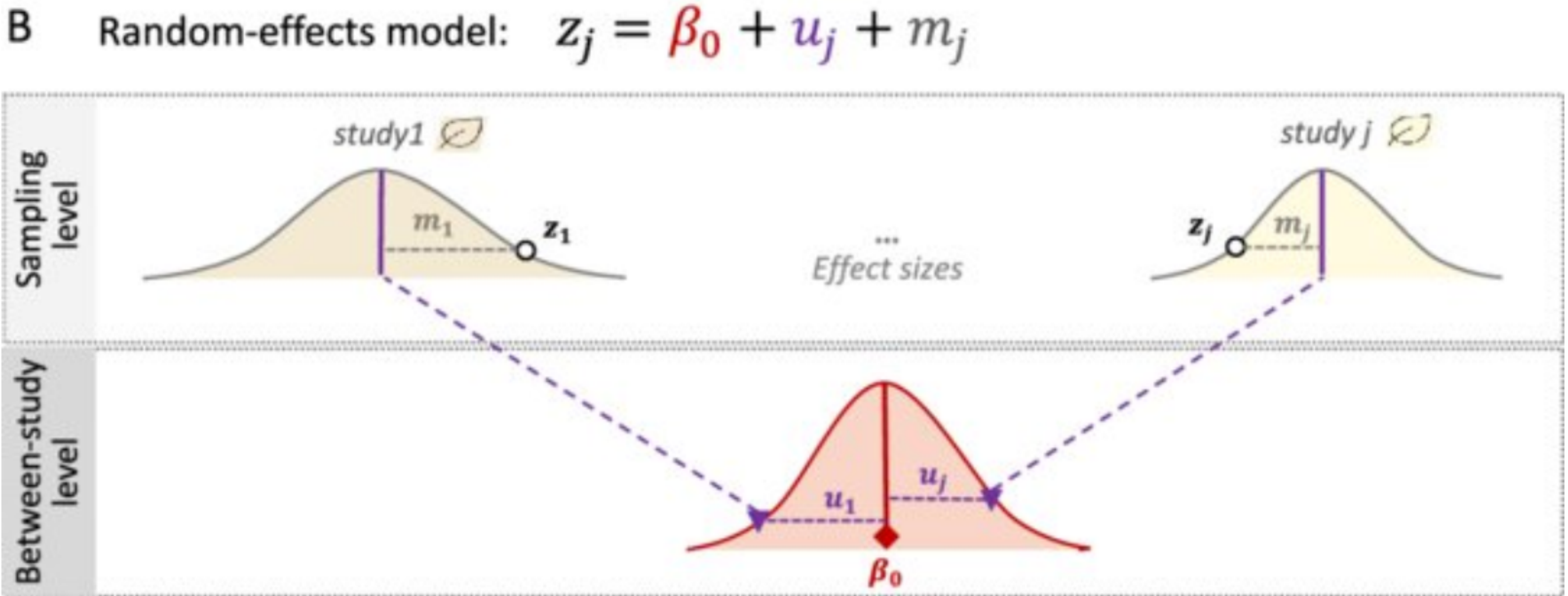
Methods for Estimating τ^2

- **DerSimonian-Laird (DL)**: Widely used but underestimates with high heterogeneity.
- **Restricted Maximum Likelihood (REML)**: Less biased, recommended for small samples.
- **Hedges & Olkin (HO)**: Adjusted moment estimator.
- **Paule-Mandel (PM)**: Performs well under high heterogeneity.
- **Bayesian Estimation**: Provides a distribution for τ^2, useful with prior knowledge.

Understanding Knapp-Hartung Adjustments

- The **Knapp-Hartung Adjustments** are applied to improve the accuracy of **confidence intervals** in random-effects meta-analyses.
- They adjust the **standard errors** of the pooled effect size to account for the uncertainty in estimating between-study variance (τ^2).
- More **conservative** than traditional methods, resulting in **wider confidence intervals**, especially when the number of studies is small or heterogeneity is high.
- Recommended for **meta-analyses with few studies** (< 10).
- Useful when between-study heterogeneity is suspected but difficult to quantify accurately.
- Can prevent **false positives** by reducing the risk of overestimating precision.

Mixed Model Estimation



How to calculate and combine effect-sizes?

Three-Level Meta-Analytic Model

Overview

Three-level meta-analysis extends the traditional random-effects model by incorporating an additional level to account for dependencies in the data, such as multiple effect sizes within studies. This approach models the variability at three levels:

1. **Level 1**: Sampling variability within effect sizes.
2. **Level 2**: Variability between effect sizes within studies.
3. **Level 3**: Variability between studies.

Purpose of the Model

The three-level model is ideal for datasets with hierarchical structures, allowing for accurate estimation of variance components and avoiding underestimation of standard errors. This model is frequently used when there are multiple outcomes per study or when primary studies include repeated measures.

Three-Level Meta-Analytic Model

Formula

$$\hat{\theta}_{ij} = \mu + u_k + \zeta_{kj} + \epsilon_{kj}$$

$$u_j \sim N(0, \tau_1^2); \quad \zeta_{ij} \sim N(0, \tau_2^2); \quad \epsilon_{ij} \sim N(0, s_{kj}^2)$$

Where:

- hat_theta_kj: Estimated effect size for the (k)-th effect size in the (j)-th study.
- mu: The overall mean effect size.
- u_k: Random effect for study (k), with (_1^2) as the between-study variance.
- zeta_{kj}: Random effect for effect size (k) within study (j)
- epsilon_kj: Error term, with (v_{kj}) representing the sampling variance.

Three level random effect model: weights

. - Weights for individual effect sizes are defined as:

$$w_{ij} = \frac{1}{\tau_1^2 + \tau_2^2 + v_{ij}}$$

Three level random effect model

Let’s say RR=1.5; $\hat{\theta}_{k} = \mu + \zeta_{(2)jk} + \zeta_{(3)j} + \varepsilon_{jk}$

Forest Plot with Three-Level Meta-Analytical Model

Standard Deviation of True Effect Size:

0.05 (range 0–1)

0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1

Standard Deviation of Sampling Error:

0.2 (range 0–1)

0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1

Standard Deviation of Group Effect:

0.15 (range 0–1)

0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1

Number of Studies:

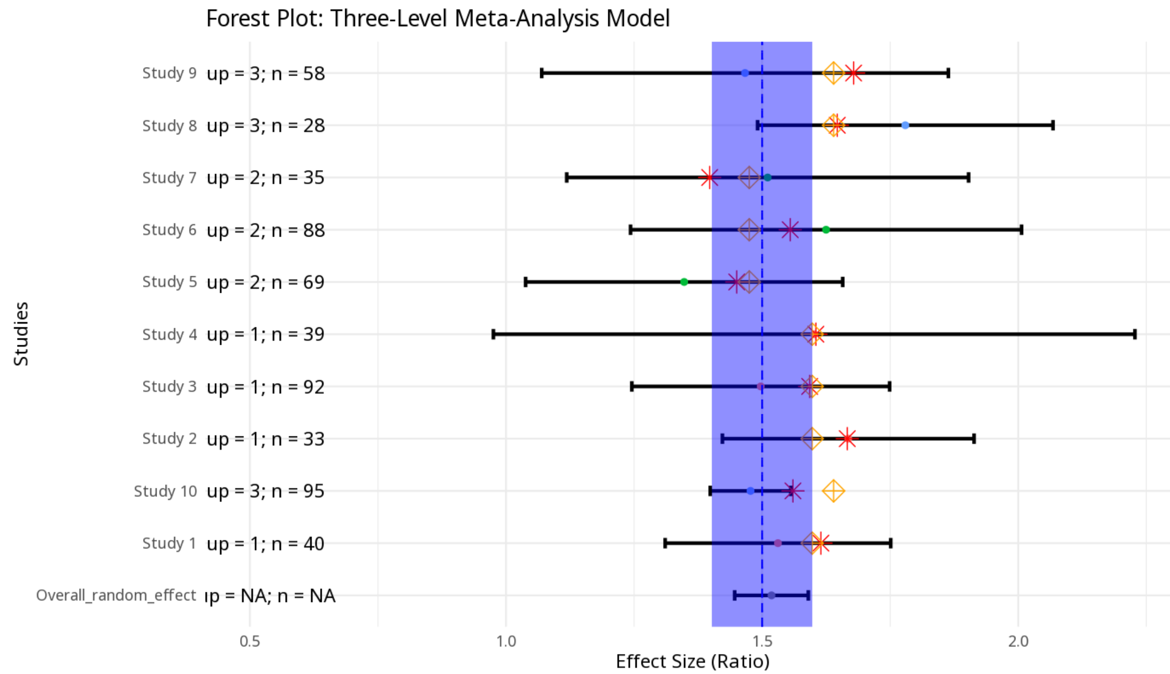
10 (range 3–15)

3 5 7 9 11 13 15

Number of Groups:

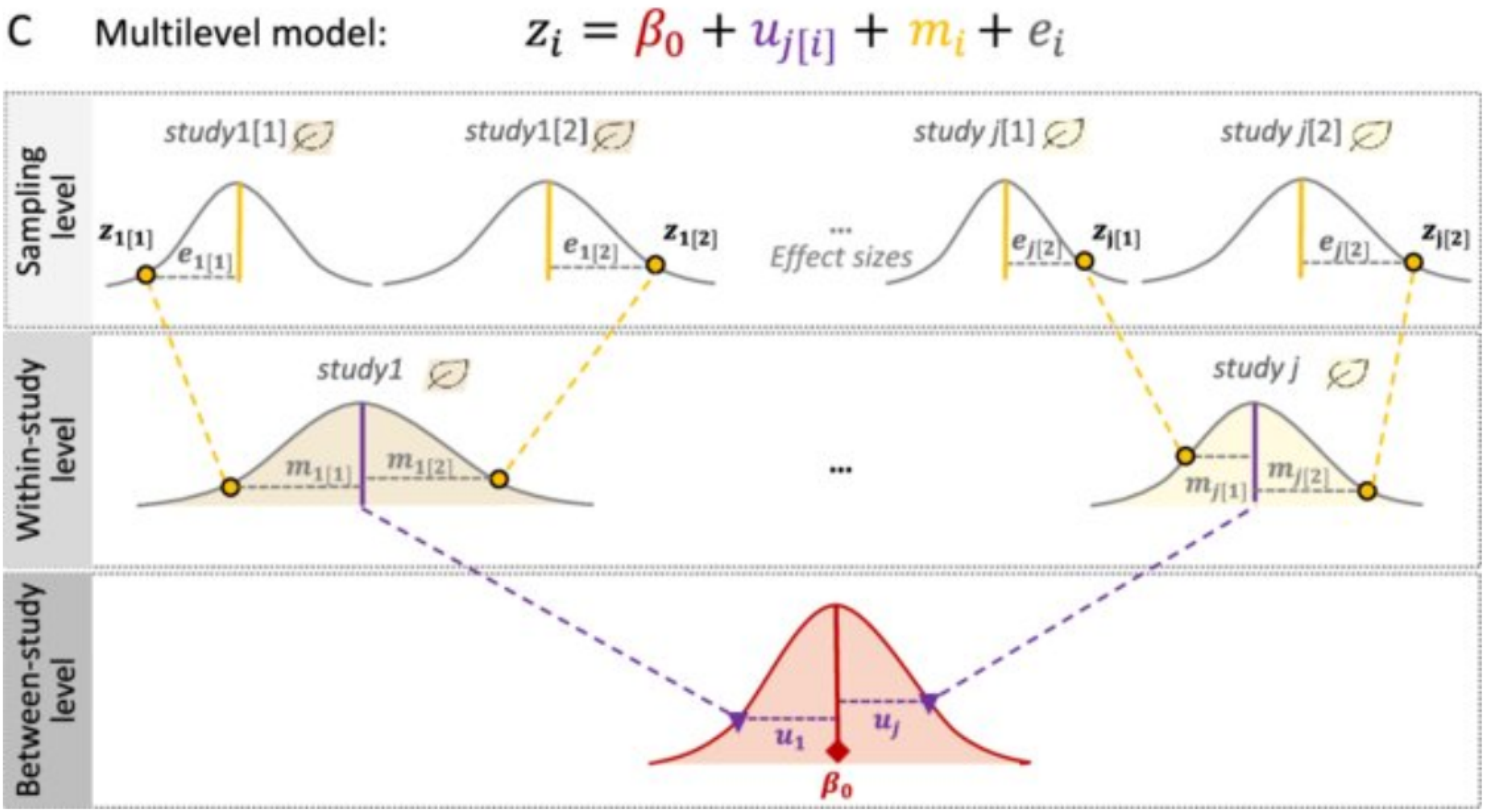
3 (range 2–5)

2 3 4 5



How to calculate and combine effect-sizes?

Three level random effect model



How to calculate and combine effect-sizes?

Subgroup Analyses

Definition

- **Subgroup Analysis** involves dividing studies into distinct groups based on specific characteristics to assess whether effect sizes vary across these groups.

Purpose

- To identify factors that influence the effect of an intervention and to better understand heterogeneity among study results.

When to Conduct Subgroup Analyses

Criteria for Subgroup Analysis

- **Significant Heterogeneity:** High I^2 values suggest further investigation is warranted.
- **Hypothesis Testing:** When researchers have specific hypotheses regarding differences in effects.
- **Interest in Specific Populations:** To evaluate how different demographics (e.g., age, gender) respond to an intervention.

Methodology of Subgroup Analyses

Statistical Approach

- Analyze subsets of data separately and compare effect sizes across groups.
- Use a meta-analytic model for each subgroup, maintaining consistency with overall methodology.

Formula

- For subgroup analyses, the overall model can be modified to include group-specific parameters:

$$\hat{\theta}_{g,k} = \mu_g + \zeta_{g,k} + \epsilon_{g,k}$$

Where: - (*{g,k}): Estimated effect size for study (k) in group (g). - (g): Overall mean effect for group (g). - ({g,k}): Random effect for study (k) in group (g). - (* {g,k}): Error term for study (k) in group (g).

Meta-Regression

Definition

- **Meta-Regression** extends meta-analysis by examining how study-level covariates influence effect sizes.

Purpose

- To explore relationships between study characteristics (e.g., sample size, intervention duration) and observed effects.

Meta-Regression Methodology

Statistical Approach

- The meta-regression model includes covariates that may explain variability in effect sizes:

$$\hat{\theta}_k = \mu + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p + \zeta_k + \epsilon_k$$

Where: - (x_i): Study-level covariate. - (_i): Coefficient indicating the relationship between (x_i) and effect size.

Benefits of Meta-Regression

Advantages

- **Identify Predictors of Variability:** Helps clarify which factors influence intervention effectiveness.
- **Informs Future Research Directions:** Insights from meta-regression can guide future studies and hypotheses.

Challenges in Meta-Regression

Considerations

- **Sample Size Requirements:** Requires sufficient studies with varying characteristics for robust results.
- **Complexity of Interpretation:** Interactions among covariates can complicate understanding.