Practical Considerations: Using Robust Standard Errors in Meta-regression

Elizabeth Tipton
Northwestern University, USA

Presented at the Joint Cochrane Campbell Colloquium, Keystone Colorado, October 2010

Outline

1. Choosing the right model
   – Correlated Effects
   – Hierarchical Model
2. Different types of regression coefficients
   – Independent effects meta-regression (I-MR) vs. Dependent effects meta-regression (D-MR)
3. Efficient weight estimation: What about ρ?

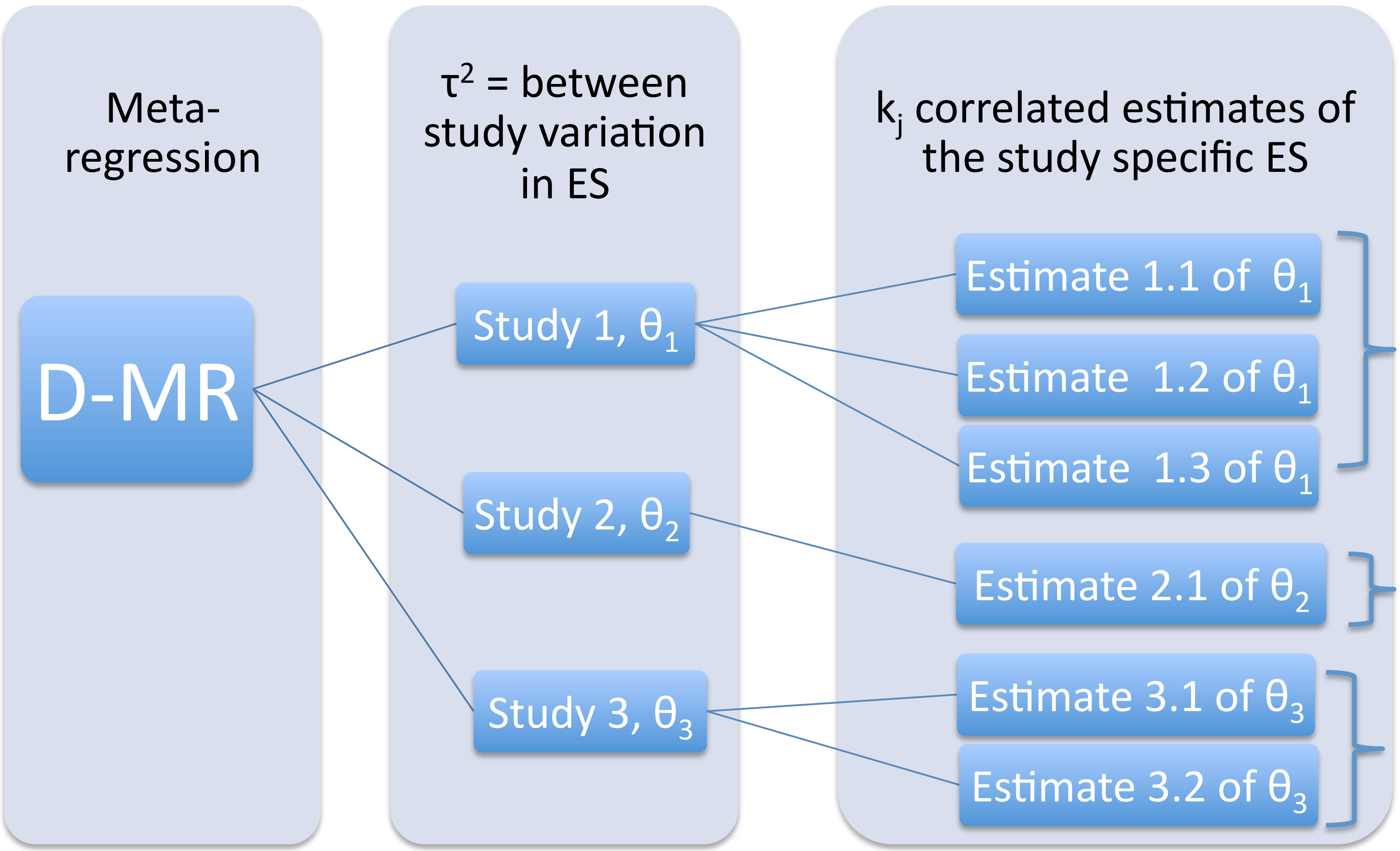
1. Choosing the right model

Where does the dependency come from?

1. If from sampling errors, then use the **correlated effects model**.

- e.g. errors that arise because the same people are used to calculate multiple effect sizes
  - e.g. Multiple measures are collected on each person in the study.
  - e.g. The same control group is used for multiple treatment contrasts.
- This model assumes
  - there is between study random variation (τ^2); and
  - every within-study effect size is an estimate of the same underlying study-specific effect size parameter; and
  - the within-study correlation is induced by sampling error.

1.1 Correlated Effects Model

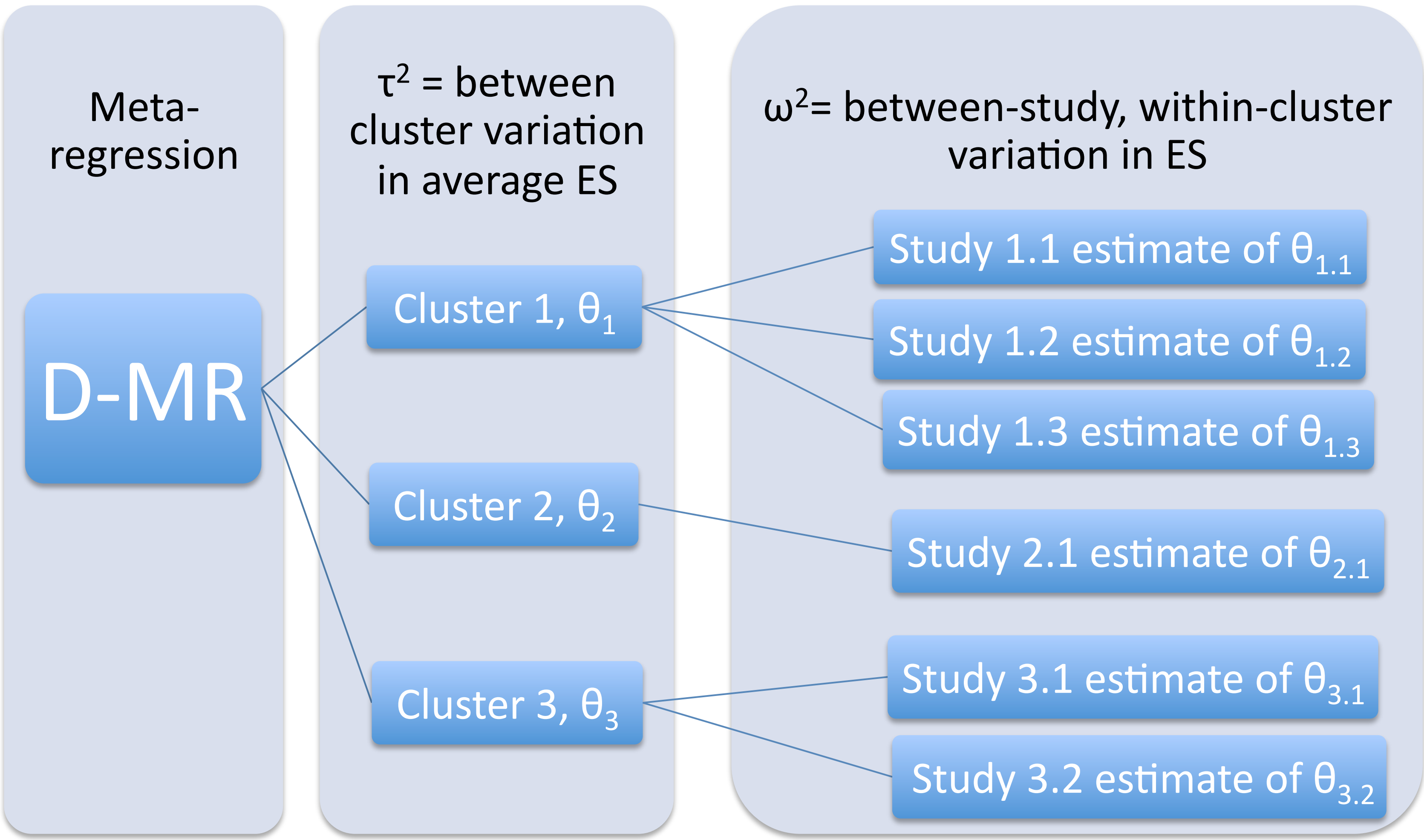


1.2 Choosing the right model

2. If from parameter variance, then use the **hierarchical model**.

– Variation that arise from parameters:
  - e.g. Studies nested within clusters of researchers.
  - e.g. Studies nested within research themes.
– This model assumes
  - there is between-cluster random variation in average study effect sizes (τ^2); and
  - there is within-cluster random variation in effect size parameters across studies (ω^2); and
  - there are no sampling error induced correlations.

1.3 Hierarchical Model



2. Within- vs. between- study models

In detecting the effect of an intervention, there are two types of models and analyses:

1. Within-study model:

For example:

- Within each study, people are randomly assigned to:
  - Exercise 6 times a week (Treatment A); or
  - Exercise 3 times a week (Treatment B); or
  - Not exercise (Control).
- The ES for comparing Treatments B and A then is simply

$$ES_{(\text{Trt B vs Trt A})} = ES_{(\text{Trt A vs C})} - ES_{(\text{Trt B vs C})}$$

- In a MR, each study contributes one such comparison. The MR estimates an average ***causal*** ES for exercising 6 (vs 3) times per week.

2.1 Within- vs. between- study models

2. Between-study effects:

For example:

- In some studies, individuals are randomly assigned to:
  - Exercise 6 times a week (Treatment A); or
  - Not exercise (Control).
- In other studies, individuals are randomly assigned to:
  - Exercise 3 times a week (Treatment B); or
  - Not exercise (Control).
- The ES for exercising 6 (vs 3) times a week is calculated using MR. Note that this average effect is ***NOT causal***.

The idea of within- and between-study effects plays an important role in D-MR.

2.2 Regression Coefficients: I-MR

Let T_j be the effect size and X_j be the length to follow-up:

$$T_j = \beta_0 + X_j\beta_1 + \cdots$$

Note: here each study contributes one value of X_j.

The coefficients β_0 and β_1 can be interpreted as:

- β_0 = the average effect size when $X_j = 0$
  - e.g. the average effect size in studies in which the intervention just occurred.
- β_1 = the effect of a 1-unit increase in X_j on T_j
  - e.g. the ES change from moving from a study in which the intervention just occurred to a study in which the ES was measured at a follow-up 1 month later.

2.3 Regression Coefficients: D-MR

For a fixed study (j=1), now assume there are multiple outcomes. This study has its own regression equation:

$$T_{i1} = \beta_{01} + X_{i1}\beta_{11} + \ldots$$

Note: here each outcome contributes one value of X_{i1}.

<u>The coefficients β_{01} and β_{11} can be interpreted as:</u>

- β_{01} = the average effect size when $X_{i1} = 0$
  - e.g. the average effect size for units in the study (j=1) when the intervention just occurred.
- β_{11} = the effect of a 1-unit increase in X_{i1} on T_{i1}
  - e.g. the effect size change for units in the study at the time of intervention and at follow-up 1 month later.

2.4 Between and within: D-MR

In D-MR these two types of regression occur in one analysis:

- Within Study: $T_{ij} = \beta_{0j} + X_{ij}\beta_2 + \cdots$
- Between Study: $\beta_{0j} = \beta_0 + X_{\bullet j}\beta_1 + \cdots$

Here there are two different relationships between X and T:

- The **between-study** effect of $X_{\bullet j}$ on $T_{\bullet j}$
  - Note: this effect *is* found in I-MR
- The **within-study** effect of X_{ij} on T_{ij}
  - Note: this effect *is NOT* found in I-MR

These two regressions are combined into one analysis and model:

$$T_{ij} = \beta_0 + X_{ij}\beta_2 + X_{\bullet j}\beta_1 + \cdots$$

2.5 The effect of centering: D-MR

In D-MR, *how X is centered*:

- Affects the interpretation of the coefficients; and
- Allows within- and between- study effects to be properly separated.

The best way to center is call **Group Mean Centering:**

$$X^{c}_{ij} = X_{ij} - X_{\bullet j}$$

Where $X_{\bullet j}$ is the mean value of X_{ij} in group j (and where group is either study or cluster).

2.6 Different models for X: D-MR

Model 1: $T_{ij} = \beta_0 + X^c_{ij}\beta_2 + \ldots$

- Here only the within-effect of X^c_{ij} is of interest.
- β_2 = the effect of a 1-unit increase in X_{ij} on T_{ij}

Model 2: $T_{ij} = \beta_0 + X_{\bullet j}\beta_1 + \ldots$

- Here only the between-effect of $X_{.j}$ is of interest.
- β_1 = the effect of a 1-unit increase in $X_{.j}$ on $T_{.j}$

Model 3: $T_{ij} = \beta_0 + X_{\bullet j}\beta_1 + X^c_{ij}\beta_2 + \ldots$

- Here both the within and between effects of X are of interest.
- β_2 and β_1 are as above, and their estimates are independent.

2.7 What can go wrong?

What if you don't center X_{ij}?

$$T_{ij} = \beta_0 + X_{ij}\beta_2 + \ldots$$

- The effect you *think* you are modeling (β_2) is the within- effect, BUT
- What you have *actually* modeled is a weighted combination of the within- and between- effects.
- This makes interpreting β_2 *very difficult*.

→ **This is why group centering is preferred.**

2.8 Interpretation issues

Including $\beta_2 X^c_{ij}$ in addition to $\beta_1 X_{.j}$ in the model allows a *new* type of effect to be modeled.

However, in some D-MR, there may only be a few clusters with values of X that vary *within* each cluster.

This leads to 2 issues:

1. The estimate of β_2 (associated with X^c_{ij}) will be imprecise (i.e. have a large standard error).
2. The types of clusters in which X_{ij} varies may be different (i.e. not representative) of clusters in which X_{ij} does not vary.

2.9 Conclusions

1. When using a covariate, ask if the effect of interest is *between-* ($X_{\bullet j}$) or *within-* (X_{ij}^{C}) studies.
2. Make sure to *group-center* your within-study variables.
3. Check your data to see if X_{ij}^{C} varies in many studies and if you think these studies are representative of all studies.

3. Weighting issues: ρ unknown

Recall that in general,

1. The robust standard error estimator ***does not*** require information on the true correlation in the data. Additionally, the estimator works for *any weights.*
2. The **most efficient weights** are inverse-variance weights,
   - i.e for any covariance matrix Σ, $W = \Sigma^{-1}$

In the **hierarchical model**, these weights can be estimated fairly easily.

In the **correlated effects** case, while the variances are known, the covariances between the estimates are NOT known.

3.1 Correlated effects model

One way to estimate the efficient weights is to assume a simplified correlation structure.

Assume that within each study j,

1. The correlation between all the pairs of effect sizes is a constant ρ;
2. This correlation is the same in all studies; and
3. The k_j sampling variances within the study are approximately equal, with average $V_{\bullet j}$.

Then the (approximately) efficient weights can be shown to be:

$$W_{ij} = 1/\{(V_{\bullet j} + \tau^2)[1+(k_j-1)\rho]\}$$

3.2 Where ρ occurs in weights

In these weights, the unknown correlation ρ occurs twice:

1. In the estimator of τ^2:

$$\hat{\tau}^2 = \frac{Q_E - m + \mathrm{tr}\left(\mathbf{V}\sum_{j=1}^{m}\frac{w_j}{k_j}\mathbf{X}_j'\mathbf{X}_j\right) + \rho\,\mathrm{tr}\left(\mathbf{V}\sum_{j=1}^{m}\frac{w_j}{k_j}\left[\mathbf{X}_j'\mathbf{J}_j\mathbf{X}_j - \mathbf{X}_j'\mathbf{X}_j\right]\right)}{\sum_{j=1}^{m}k_j w_j - \mathrm{tr}\left(\mathbf{V}\sum_{j=1}^{m}w_j^2\mathbf{X}_j'\mathbf{J}_j\mathbf{X}_j\right)}$$

where $\mathbf{J}_j$ is a k_j x k_j matrix of ones and $\mathbf{V} = (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}$

2. In the multiplier: $[1+(k_j-1)\rho]$.

3.3 Three approaches

Three strategies for dealing with ρ in calculating efficient weights can be used:

1. Sensitivity approach

2. Conservative approach

3. External information approach

3.4 Sensitivity approach

1. Sensitivity approach:
   – Run the model with various values of ρ in (0,1).

This approach allows you to see how **robust** the results are to weights based on different values of ρ.

3.5 Conservative approach

2. Conservative approach:
   – Assume $\rho = 1$.

The weights become: $W_{ij} = 1/\{k_j(V_{\bullet j} + \tau^2)\}$.

Each study gets total weight $\Sigma W_{ij} = 1/(V_{\bullet j} + \tau^2)$.

These are **conservative** in that a study does not receive additional weight because it has multiple measures.

3.6 External information approach

3. External information approach:

   – e.g. test reliability measures, information from a study that reports the correlations, information from the design of the test.

   – In the **binary outcomes** case (e.g. log odds ratio), an upper bound on ρ can be calculated from estimates of the treatment and control proportions, (π_C, π_T).

3.7 Combined approaches

In practice, a combination of these three approaches will often be used.

Example: *HTJ recommendation*

- Estimate τ^2 using a sensitivity approach; but
- For the multiplier, use a conservative strategy.

Or, the same strategy can be used for both estimating both τ^2 and the multiplier.

Overall Conclusions

1. Make sure you choose the proper model for the type of dependencies in your data.

2. For each covariate X_{ij} in your model, remember that you can include
    1. The group-centered within-study variable ($X^c_{ij} = X_{ij} – X_{\cdot j}$), and/or
    2. The average ($X_{\cdot j}$).

3. When using the correlated effects model with efficient weights:
    1. If you have information on ρ, use it!
    2. If the T_{ij} are functions of proportions, use this information to get an upper bound on ρ.
    3. If you have no information on ρ:
        - Use a sensitivity approach for estimating τ^2
        - Assume ρ=1 in your weights, i.e. $W_{ij} = 1/k_j[V_{\cdot j} + \tau^2]$

Thank you!

For more information:

http://www.northwestern.edu/ipr/qcenter/RVE-meta-analysis.html