Random and Fixed Effects Models in Meta-analysis

Terri Pigott, C2 Methods Editor & co-Chair
Professor, Loyola University Chicago
tpigott@luc.edu

Fixed and Random Effects Models in Meta-analysis

• How do we choose among fixed and random effects models when conducting a meta-analysis?
• Common question asked by reviewers working on systematic reviews that include a meta-analysis
Our goal today

• Provide a description of fixed and random effects models
• Outline the underlying assumptions of these two models in order to clarify the choices a reviewer has in a meta-analysis
• Discuss how to estimate key parameters in the model
• Introduce issues for random and mixed-effects basic meta-analysis and moderator analyses
• Share resources for these methods for reviewers

An advertisement

• Much of my own thinking on these issues has been clarified by the following sources:


Our choice between the two models depends on:

- Our assumption about how the effect sizes vary in our meta-analysis
- The two models are based on different assumptions about the nature of the variation among effect sizes in our research synthesis

**Fixed-Effect model**

- Borenstein et al. adopt the wording of fixed effect (no “s”) here because in a fixed effect model, we assume that the effect sizes in our meta-analysis differ only because of sampling error and they all share a common mean
- Our effect sizes differ from each other because each study used a different sample of participants – and that is the only reason for the differences among our estimates
In a fixed-effect model

- Note that the effect size from each study estimate a single common mean – the fixed-effect
- We know that each study will give us a different effect size, but each effect size is an estimate of a common mean, designated in the prior picture as $\theta$
In a random effects model

- We assume two components of variation:
  - Sampling variation as in our fixed-effect model assumption
  - Random variation because the effect sizes themselves are sampled from a population of effect sizes
- In a random effects model, we know that our effect sizes will differ because they are sampled from an unknown distribution
- Our goal in the analysis will be to estimate the mean and variance of the underlying population of effect sizes
In a random effects model

• We see in the picture that each distribution has its own mean that is sampled from the underlying population distribution of effect sizes
• That underlying population distribution also has its own variance, $\tau^2$, commonly called the variance component
• Thus, each effect size has two components of variation, one due to sampling error, and one from the underlying distribution

Some notation for shorthand

$\nu_i^2$ is the sampling variance for our effect size
$\tau^2$ is the variance component,
the variance of our effect size distribution
$T_i$ is our effect size (SMD, odds-ratio, correlation, etc)
$\mu$ is the mean of the underlying effect size distribution
In fixed effects, we can write our model as

\[ T_i = \theta + e_i, \text{ where} \]
\[ e_i \sim N(0, \nu_i^2) \]

So, each effect size estimates a single mean effect, \( \theta \), and differs from this mean effect by sampling error.

In random effects, we can write our model as

\[ T_i = \mu + e_i + \xi_i, \text{ where} \]
\[ e_i \sim N(0, \nu_i^2) \text{ and } \xi_i \sim N(0, \tau^2) \]

Each effect size differs from the underlying population mean, \( \mu \), due to both sampling error and the underlying population variance.
Given the prior slides, how do we choose?

- The **fixed-effect model** assumes only sampling error as the source of variation among effect sizes
- This assumption is plausible when our studies are close replications of one another, using the same procedures, measures, etc.

Choosing a model (continued)

- In the **random effects model**, we assume that our effect sizes are sampled from an underlying population of effect sizes
- We already assume that our studies will differ not only because there are different participants, but also because of differences in the way studies were conducted
- Thus, we often choose random effects models because we anticipate variation among our studies for a number of reasons
Implications of the choice

- Recall that all our analyses in a meta-analysis are weighted by the inverse of the variance of the effect size, i.e., by the precision of the effect size estimate
- Because we have more variation assumed in a random effects model, our weights for each study will be more equal to one another
- In other words, in a fixed effect model, we will more heavily weight larger studies. In a random effects model, the larger studies will not be weighted as heavily

Example of a fixed effect analysis from RevMan

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Experimental Events</th>
<th>Control Events</th>
<th>Weight</th>
<th>Odds Ratio</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>01 Leschi 2001</td>
<td>70</td>
<td>211</td>
<td>63</td>
<td>53.5%</td>
<td>1.00 [0.70, 1.61]</td>
</tr>
<tr>
<td>02 Hengstler 1997</td>
<td>31</td>
<td>37</td>
<td>38</td>
<td>22.4%</td>
<td>0.50 [0.31, 1.12]</td>
</tr>
<tr>
<td>03 Hengstler</td>
<td>19</td>
<td>50</td>
<td>40</td>
<td>14.6%</td>
<td>0.74 [0.49, 2.96]</td>
</tr>
<tr>
<td>04 Hengstler</td>
<td>9</td>
<td>43</td>
<td>40</td>
<td>9.4%</td>
<td>0.12 [0.05, 0.33]</td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td>364</td>
<td>372</td>
<td>100.0%</td>
<td>0.79 [0.58, 1.07]</td>
<td></td>
</tr>
</tbody>
</table>

Total events: 123
Heterogeneity: Chi² = 10.16, df = 3 (P = 0.0094); I² = 59%
Test for overall effect Z = 1.56 (P = 0.12)
So, where are we now?

• We have discussed the underlying assumptions of fixed effect and random effects models.
• We have also talked about how we choose between the two models.
• Now we will talk about what the basic analysis and moderator analysis under a fixed effect and random effects model would entail.
Goals in a basic fixed effect meta-analysis

• In a fixed effect model, we will be most interested in estimating our common effect, $\theta$, and its standard error
• We will also want to know if there is heterogeneity present by computing $Q$
• HOWEVER, WE WILL NOT decide on the basis of a significant $Q$ that we should really do a random effects model
• WE WILL make a decision about fixed effect versus random effects models because of our substantive knowledge of the area of the systematic review

Goals in a random effects model

• We also want to estimate a mean effect size, but now this is the mean effect size from the underlying population, $\mu$
• We also want to estimate the variance of the underlying effect size, $\tau^2$
• We will test heterogeneity by testing whether $\tau^2$ is different from 0.
• Biggest difficulty: how to estimate $\tau^2$
Estimating the variance component, $\tau^2$

- There are two main methods for computing $\tau^2$
  - Variously called the method of moments, the DerSimonian/Laird estimate
  - Restricted maximum likelihood
- The method of moments estimator is easy to compute and is based on the value of $Q$, the homogeneity statistic
- Restricted maximum likelihood requires an iterative solution

Method of moments estimator

$$\tau^2 = \frac{Q - (k - 1)}{c},$$

where

- $Q$ is the homogeneity statistic,
- $k$ is the number of effect sizes, and
- $c$ is based on the fixed effect weights, $w_i$, and is

$$c = \sum w_i - \frac{\sum w_i^2}{\sum w_i}$$

This is the estimator used in many meta-analysis programs such as RevMan and CMA
Restricted maximum likelihood estimator

- Many statisticians do not like the method of moments estimator
- Can estimate $\tau^2$ using HLM, SAS Proc Mixed, or R
- I can give you sample programs for these estimators

Brief discussion of random effects moderator analyses

- In planning our meta-analysis, we have extracted from each study information about the methods, measures, etc that we believe are related to potential variation in effect sizes
- We have done this a priori because we have substantive knowledge of the area, and we have a logic model guiding our systematic review
- Let us suppose we want to see if our effect sizes differ depending on the age of our participants: elementary school (age 5 – 11), middle school (age 12-14), high school (14-18)
Categorical model (analogue to one-way ANOVA)

- We have three categories: elementary school, middle school, and high school
- We want to see if the mean effect sizes for studies grouped by age of student differ from each other
- Question: What are our assumptions about $\tau^2$?

Our options for $\tau^2$

- Option 1: We assume that each group has its own underlying distribution of effect sizes, so that we are really estimating $\mu_{elem}$, $\mu_{middle}$, and $\mu_{high}$ as well as three separate variance components, $\tau^2_{elem}$, $\tau^2_{middle}$, and $\tau^2_{high}$.
- Problem with Option 1: Usually we don’t have a large number of studies within each group, and our estimates of the variance components will not be precise without sufficient numbers of studies (no matter how we estimate the variance component)
Option 2

- Assume that we have a common variance component $\tau^2$
- We will estimate our variance component in our initial stages of the meta-analysis, and we will use that variance component for our analyses
- Thus, each effect size will be weighted by inverse of the sum of its sampling variance and the variance component, and this will be the weight used in the ANOVA model
- We can also use this weight in meta-regression

Example from Sirin (2005)

<table>
<thead>
<tr>
<th>Type of achievement test</th>
<th>n</th>
<th>Weighted Mean</th>
<th>SD of mean effect size</th>
<th>Lower 95% CI</th>
<th>Upper 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>State</td>
<td>5</td>
<td>0.614</td>
<td>0.146</td>
<td>0.327</td>
<td>0.900</td>
</tr>
<tr>
<td>Standardized</td>
<td>6</td>
<td>0.265</td>
<td>0.129</td>
<td>0.012</td>
<td>0.517</td>
</tr>
<tr>
<td>TOTAL</td>
<td>11</td>
<td>0.417</td>
<td>0.096</td>
<td>0.228</td>
<td>0.606</td>
</tr>
</tbody>
</table>
Some concluding remarks

• Given the anticipated differences among studies in a typical social science systematic review, we usually choose random effects models.

• From the C2 Methods Editor (me): make sure to provide a rationale for your choice in the protocol, based on your substantive knowledge of the area of the review.

• Many other issues not touched on in this talk and see the references following this slide.

References


