Promoting Transparency & Replicability in Meta-Analysis

Emily E. Tanner-Smith, PhD
October 24th, 2019
OPRE Methods Meeting
“Most of us were trained to analyze complex relationships among variables in the primary analysis of research data. But at the higher level, where variance, nonuniformity and uncertainty are no less evident, we too often substitute literary exposition for quantitative rigor. The proper integration of research requires the same statistical methods that are applied in primary data analysis.”

Gene V. Glass (1976: 6)
What is Meta-Analysis?

**Research synthesis**: process by which findings from two or more empirical research studies are summarized (synthesized) to address a particular research question.

**Systematic reviews**: systematic identification, appraisal, and synthesis of findings from two or more studies.

**Meta-analysis**: statistical (quantitative) synthesis of numerical findings from two or more studies.
<table>
<thead>
<tr>
<th>Study</th>
<th>Intervention effect</th>
<th>SMD</th>
<th>95%-CI</th>
<th>Weight (fixed)</th>
<th>Weight (random)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>-0.08</td>
<td>[-0.34; 0.19]</td>
<td>12.5%</td>
<td>7.0%</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>-0.11</td>
<td>[-0.32; 0.11]</td>
<td>18.6%</td>
<td>7.1%</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>-0.00</td>
<td>[-0.80; 0.79]</td>
<td>1.4%</td>
<td>5.2%</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>0.21</td>
<td>[0.02; 0.44]</td>
<td>16.2%</td>
<td>7.1%</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>0.21</td>
<td>[0.07; 0.49]</td>
<td>11.3%</td>
<td>7.0%</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>0.82</td>
<td>[0.50; 1.15]</td>
<td>8.2%</td>
<td>6.9%</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>0.15</td>
<td>[0.35; 0.66]</td>
<td>3.4%</td>
<td>6.3%</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>0.04</td>
<td>[-0.38; 0.29]</td>
<td>7.6%</td>
<td>6.8%</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>0.36</td>
<td>[0.20; 0.92]</td>
<td>2.7%</td>
<td>6.1%</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>0.46</td>
<td>[0.06; 0.85]</td>
<td>5.6%</td>
<td>6.7%</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td>0.18</td>
<td>[0.07; 0.21]</td>
<td>5.6%</td>
<td>6.7%</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td>1.09</td>
<td>[0.41; 1.76]</td>
<td>1.9%</td>
<td>5.7%</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td>3.38</td>
<td>[2.39; 4.38]</td>
<td>0.9%</td>
<td>6.7%</td>
</tr>
<tr>
<td>14</td>
<td></td>
<td>0.47</td>
<td>[-0.27; 1.22]</td>
<td>1.8%</td>
<td>5.4%</td>
</tr>
<tr>
<td>15</td>
<td></td>
<td>-0.36</td>
<td>[-1.24; 0.52]</td>
<td>1.1%</td>
<td>4.9%</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td>0.79</td>
<td>[-0.07; 1.64]</td>
<td>1.2%</td>
<td>5.0%</td>
</tr>
<tr>
<td>17</td>
<td></td>
<td>0.14</td>
<td>[-2.64; 2.36]</td>
<td>0.1%</td>
<td>1.5%</td>
</tr>
</tbody>
</table>

**Fixed effect model**

<table>
<thead>
<tr>
<th>Effect size</th>
<th>Weight (fixed)</th>
<th>Weight (random)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.19</td>
<td>100.0%</td>
<td>--</td>
</tr>
</tbody>
</table>

**Random effects model**

<table>
<thead>
<tr>
<th>Effect size</th>
<th>Weight (fixed)</th>
<th>Weight (random)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.38</td>
<td>100.0%</td>
<td>--</td>
</tr>
</tbody>
</table>

**Prediction interval**

Heterogeneity \( I^2 = 81\% \), \( \tau^2 = 0.4085 \), \( p < 0.01 \)

\[-1.03; 1.79\]
Petrosino et al. (2013)
Scared Straight
Meta-Analysis
Best Practice Guidelines
• AHRQ (2014)
• APA MARS (2018)
• Cochrane MECIR (2019)
• PRISMA

Critical Appraisal Tools
• AMSTAR (2007)
• JBI Tool (2017)
• ROBIS (2016)

Key Considerations
• Protocol
• Research Questions
• Inclusion/Exclusion Criteria
• Literature Search Strategies
• Data Extraction Methods
• Statistical Modelling Approaches
• Heterogeneity Assessments
• Risk of Bias & Study Quality Assessments
• Sensitivity Analyses
• Coverage of Evidence
• Practical Relevance
Common Registries

- Campbell Collaboration
- Cochrane Collaboration
- Figshare
- Open Science Framework
- PROSPERO
Meta-Analytic Research Questions
<table>
<thead>
<tr>
<th><strong>Broad Scope</strong></th>
<th><strong>Narrow Scope</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Example</strong></td>
<td>To assess the effectiveness of brief alcohol interventions for reducing binge drinking among college students</td>
</tr>
<tr>
<td></td>
<td>To assess the effectiveness of brief personalized feedback interventions for reducing binge drinking among college students attending football games</td>
</tr>
<tr>
<td><strong>Pros</strong></td>
<td>• Can address broader questions about when, why, and for whom a program works</td>
</tr>
<tr>
<td></td>
<td>• Can model variability in effects across settings, contexts, program components, other study features (e.g., comparative effectiveness)</td>
</tr>
<tr>
<td></td>
<td>• Can address specific questions about “what works”</td>
</tr>
<tr>
<td></td>
<td>• Straightforward interpretation</td>
</tr>
<tr>
<td></td>
<td>• Less resource intensive</td>
</tr>
<tr>
<td><strong>Cons</strong></td>
<td>• More resource intensive</td>
</tr>
<tr>
<td></td>
<td>• Synthesizing heterogeneous evidence can yield uninformative mean effect sizes</td>
</tr>
<tr>
<td></td>
<td>• Complexity of interpretation</td>
</tr>
<tr>
<td></td>
<td>• Small number of studies synthesized</td>
</tr>
<tr>
<td></td>
<td>• Limited external validity</td>
</tr>
<tr>
<td></td>
<td>• Inability to model variability in effects</td>
</tr>
</tbody>
</table>
Inclusion & Exclusion Criteria

- Participants
  - Patients
  - Population
- Intervention
  - Program
  - Policy
  - Practice
- Comparison
  - Control
- Study Design
- Outcomes
- Time Stamp
- Other:
  - Geography
  - Language
  - Quality
Systematic Literature Searching

• Comprehensive Coverage
• Diverse Sources
• Transparent Documentation
Statistical Considerations

• Missing data handling
• Modelling strategies
• Heterogeneity & variability
• Study quality & risks of bias
• Small study & reporting biases
• Sensitivity analyses & robustness checks
Applicability & Utility of Findings

• Statistical vs. clinical/practical significance

• Coverage of evidence and generalizability considerations
References & Resources


PROSPERO. International prospective register of systematic reviews. Centre for Reviews and Dissemination, University of York. Available at: [https://www.crd.york.ac.uk/prospero/](https://www.crd.york.ac.uk/prospero/)


Whiting, P., ..., & the ROBIS group. (2016). ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *Journal of Clinical Epidemiology, 69*, 225-234. doi:10.1016/j.jclinepi.2015.06.005
Contact Information

Emily E. Tanner-Smith, Ph.D.
University of Oregon
Associate Professor, Counseling Psychology and Human Services
Associate Dean for Research, College of Education
etanners@uoregon.edu
https://appliedstats.uoregon.edu/