Introduction to Meta-Analysis

Michael Borenstein
Biostat, Inc, New Jersey, USA.

Larry V. Hedges
Northwestern University, Evanston, USA.

Julian P.T. Higgins
MRC, Cambridge, UK.

Hannah R. Rothstein
Baruch College, New York, USA.

WILEY
A John Wiley and Sons, Ltd., Publication
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of Tables</td>
<td>xiii</td>
</tr>
<tr>
<td>List of Figures</td>
<td>xv</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>xix</td>
</tr>
<tr>
<td>Preface</td>
<td>xxi</td>
</tr>
<tr>
<td>Web site</td>
<td>xxix</td>
</tr>
</tbody>
</table>

### PART 1: INTRODUCTION

1 **HOW A META-ANALYSIS WORKS**
   - Introduction                                                      | 3   |
   - Individual studies                                                | 3   |
   - The summary effect                                                | 5   |
   - Heterogeneity of effect sizes                                     | 6   |
   - Summary points                                                    | 7   |

2 **WHY PERFORM A META-ANALYSIS**
   - Introduction                                                      | 9   |
   - The streptokinase meta-analysis                                   | 10  |
   - Statistical significance                                          | 11  |
   - Clinical importance of the effect                                 | 12  |
   - Consistency of effects                                            | 12  |
   - Summary points                                                    | 14  |

### PART 2: EFFECT SIZE AND PRECISION

3 **OVERVIEW**
   - Treatment effects and effect sizes                               | 17  |
   - Parameters and estimates                                          | 18  |
   - Outline of effect size computations                               | 19  |

4 **EFFECT SIZES BASED ON MEANS**
   - Introduction                                                      | 21  |
   - Raw (unstandardized) mean difference $D$                           | 21  |
   - Standardized mean difference, $d$ and $g$                         | 25  |
   - Response ratios                                                   | 30  |
   - Summary points                                                    | 32  |
5  EFFECT SIZES BASED ON BINARY DATA (2 × 2 TABLES) 33
   Introduction 33
   Risk ratio 34
   Odds ratio 36
   Risk difference 37
   Choosing an effect size index 38
   Summary points 39

6  EFFECT SIZES BASED ON CORRELATIONS 41
   Introduction 41
   Computing $r$ 41
   Other approaches 43
   Summary points 43

7  CONVERTING AMONG EFFECT SIZES 45
   Introduction 45
   Converting from the log odds ratio to $d$ 47
   Converting from $d$ to the log odds ratio 47
   Converting from $r$ to $d$ 48
   Converting from $d$ to $r$ 48
   Summary points 49

8  FACTORS THAT AFFECT PRECISION 51
   Introduction 51
   Factors that affect precision 52
   Sample size 52
   Study design 53
   Summary points 55

9  CONCLUDING REMARKS 57

PART 3: FIXED-EFFECT VERSUS RANDOM-EFFECTS MODELS

10 OVERVIEW 61
   Introduction 61
   Nomenclature 62

11 FIXED-EFFECT MODEL 63
   Introduction 63
   The true effect size 63
   Impact of sampling error 63
12 RANDOM-EFFECTS MODEL 69
| Introduction | 69 |
| The true effect sizes | 69 |
| Impact of sampling error | 70 |
| Performing a random-effects meta-analysis | 72 |
| Summary points | 74 |

13 FIXED-EFFECT VERSUS RANDOM-EFFECTS MODELS 77
| Introduction | 77 |
| Definition of a summary effect | 77 |
| Estimating the summary effect | 78 |
| Extreme effect size in a large study or a small study | 79 |
| Confidence interval | 80 |
| The null hypothesis | 83 |
| Which model should we use? | 83 |
| Model should not be based on the test for heterogeneity | 84 |
| Concluding remarks | 85 |
| Summary points | 85 |

14 WORKED EXAMPLES (PART 1) 87
| Introduction | 87 |
| Worked example for continuous data (Part 1) | 87 |
| Worked example for binary data (Part 1) | 92 |
| Worked example for correlational data (Part 1) | 97 |
| Summary points | 102 |

PART 4: HETEROGENEITY 105

15 OVERVIEW 105
| Introduction | 105 |
| Nomenclature | 106 |
| Worked examples | 106 |

16 IDENTIFYING AND QUANTIFYING HETEROGENEITY 107
| Introduction | 107 |
| Isolating the variation in true effects | 107 |
| Computing $Q$ | 109 |
| Estimating $\tau^2$ | 114 |
| The $I^2$ statistic | 117 |
### Comparing the measures of heterogeneity

- Confidence intervals for $\tau^2$  
- Confidence intervals (or uncertainty intervals) for $I^2$  
- Summary points

### Prediction Intervals

- Introduction  
- Prediction intervals in primary studies  
- Prediction intervals in meta-analysis  
- Confidence intervals and prediction intervals  
- Comparing the confidence interval with the prediction interval  
- Summary points

### Worked Examples (Part 2)

- Introduction  
- Worked example for continuous data (Part 2)  
- Worked example for binary data (Part 2)  
- Worked example for correlational data (Part 2)  
- Summary points

### Subgroup Analyses

- Introduction  
- Fixed-effect model within subgroups  
- Computational models  
- Random effects with separate estimates of $\tau^2$  
- Random effects with pooled estimate of $\tau^2$  
- The proportion of variance explained  
- Mixed-effects model  
- Obtaining an overall effect in the presence of subgroups  
- Summary points

### Meta-Regression

- Introduction  
- Fixed-effect model  
- Fixed or random effects for unexplained heterogeneity  
- Random-effects model  
- Summary points

### Notes on Subgroup Analyses and Meta-Regression

- Introduction  
- Computational model  
- Multiple comparisons  
- Software  
- Analyses of subgroups and regression analyses are observational
PART 5: COMPLEX DATA STRUCTURES

22 OVERVIEW

23 INDEPENDENT SUBGROUPS WITHIN A STUDY
   Introduction
   Combining across subgroups
   Comparing subgroups
   Summary points

24 MULTIPLE OUTCOMES OR TIME-POINTS WITHIN A STUDY
   Introduction
   Combining across outcomes or time-points
   Comparing outcomes or time-points within a study
   Summary points

25 MULTIPLE COMPARISONS WITHIN A STUDY
   Introduction
   Combining across multiple comparisons within a study
   Differences between treatments
   Summary points

26 NOTES ON COMPLEX DATA STRUCTURES
   Introduction
   Summary effect
   Differences in effect

PART 6: OTHER ISSUES

27 OVERVIEW

28 VOTE COUNTING – A NEW NAME FOR AN OLD PROBLEM
   Introduction
   Why vote counting is wrong
   Vote counting is a pervasive problem
   Summary points

29 POWER ANALYSIS FOR META-ANALYSIS
   Introduction
   A conceptual approach
   In context
   When to use power analysis
### Part 7: Issues Related to Effect Size

#### 30 Publication Bias

- **Introduction**
- **The problem of missing studies**
- **Methods for addressing bias**
- **Illustrative example**
- **The model**
- **Getting a sense of the data**
- **Is there evidence of any bias?**
- **Is the entire effect an artifact of bias?**
- **How much of an impact might the bias have?**
- **Summary of the findings for the illustrative example**
- **Some important caveats**
- **Small-study effects**
- **Concluding remarks**
- **Summary points**

#### 31 Overview

- **Introduction**
- **Relationship between p-values and effect sizes**
- **The distinction is important**
- **The p-value is often misinterpreted**
- **Narrative reviews vs. meta-analyses**
- **Summary points**

#### 32 Effect Sizes Rather Than p-Values

- **Introduction**
- **Relationship between p-values and effect sizes**
- **The distinction is important**
- **The p-value is often misinterpreted**
- **Narrative reviews vs. meta-analyses**
- **Summary points**

#### 33 Simpson’s Paradox

- **Introduction**
- **Circumcision and risk of HIV infection**
- **An example of the paradox**
- **Summary points**

#### 34 Generality of the Basic Inverse-Variance Method

- **Introduction**
- **Other effect sizes**
- **Other methods for estimating effect sizes**
- **Individual participant data meta-analyses**
### PART 8: FURTHER METHODS

#### 35 OVERVIEW

#### 36 META-ANALYSIS METHODS BASED ON DIRECTION AND \( p \)-VALUES

- Introduction
- Vote counting
- The sign test
- Combining \( p \)-values
- Summary points

#### 37 FURTHER METHODS FOR DICHOTOMOUS DATA

- Introduction
- Mantel-Haenszel method
- One-step (Peto) formula for odds ratio
- Summary points

#### 38 PSYCHOMETRIC META-ANALYSIS

- Introduction
- The attenuating effects of artifacts
- Meta-analysis methods
- Example of psychometric meta-analysis
- Comparison of artifact correction with meta-regression
- Sources of information about artifact values
- How heterogeneity is assessed
- Reporting in psychometric meta-analysis
- Concluding remarks
- Summary points

### PART 9: META-ANALYSIS IN CONTEXT

#### 39 OVERVIEW

#### 40 WHEN DOES IT MAKE SENSE TO PERFORM A META-ANALYSIS?

- Introduction
- Are the studies similar enough to combine?
- Can I combine studies with different designs?
- How many studies are enough to carry out a meta-analysis?
- Summary points

#### 41 REPORTING THE RESULTS OF A META-ANALYSIS

- Introduction
- The computational model
Contents

42 CUMULATIVE META-ANALYSIS 371

Introduction 371
Why perform a cumulative meta-analysis? 373
Summary points 376

43 CRITICISMS OF META-ANALYSIS 377

Introduction 377
One number cannot summarize a research field 378
The file drawer problem invalidates meta-analysis 378
Mixing apples and oranges 379
Garbage in, garbage out 380
Important studies are ignored 381
Meta-analysis can disagree with randomized trials 381
Meta-analyses are performed poorly 384
Is a narrative review better? 385
Concluding remarks 386
Summary points 386

PART 10: RESOURCES AND SOFTWARE

44 SOFTWARE 391

Introduction 391
The software 392
Three examples of meta-analysis software 393
Comprehensive Meta-Analysis (CMA) 2.0 395
RevMan 5.0 398
Stata macros with Stata 10.0 400
Summary points 403

45 BOOKS, WEB SITES AND PROFESSIONAL ORGANIZATIONS 405

Books on systematic review methods 405
Books on meta-analysis 405
Web sites 406

REFERENCES 409
INDEX 415