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

List of Tables xiii  
List of Figures xv  
Acknowledgements xix  
Preface xxi  
Web site xxix  

## PART 1: INTRODUCTION

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

2 WHY PERFORM A META-ANALYSIS 9  
   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 17  
   Treatment effects and effect sizes 17  
   Parameters and estimates 18  
   Outline of effect size computations 19  

4 EFFECT SIZES BASED ON MEANS 21  
   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 X 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
Comparing the measures of heterogeneity 119
Confidence intervals for \( \tau^2 \) 122
Confidence intervals (or uncertainty intervals) for \( I^2 \) 124
Summary points 125

17 PREDICTION INTERVALS 127
Introduction 127
Prediction intervals in primary studies 127
Prediction intervals in meta-analysis 129
Confidence intervals and prediction intervals 131
Comparing the confidence interval with the prediction interval 132
Summary points 133

18 WORKED EXAMPLES (PART 2) 135
Introduction 135
Worked example for continuous data (Part 2) 135
Worked example for binary data (Part 2) 139
Worked example for correlational data (Part 2) 143
Summary points 147

19 SUBGROUP ANALYSES 149
Introduction 149
Fixed-effect model within subgroups 151
Computational models 161
Random effects with separate estimates of \( \tau^2 \) 164
Random effects with pooled estimate of \( \tau^2 \) 171
The proportion of variance explained 179
Mixed-effects model 183
Obtaining an overall effect in the presence of subgroups 184
Summary points 186

20 META-REGRESSION 187
Introduction 187
Fixed-effect model 188
Fixed or random effects for unexplained heterogeneity 193
Random-effects model 196
Summary points 203

21 NOTES ON SUBGROUP ANALYSES AND META-REGRESSION 205
Introduction 205
Computational model 205
Multiple comparisons 208
Software 209
Analyses of subgroups and regression analyses are observational 209
<table>
<thead>
<tr>
<th>Contents</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Planning for precision rather than for power</td>
<td>263</td>
</tr>
<tr>
<td>Power analysis in primary studies</td>
<td>263</td>
</tr>
<tr>
<td>Power analysis for meta-analysis</td>
<td>267</td>
</tr>
<tr>
<td>Power analysis for a test of homogeneity</td>
<td>272</td>
</tr>
<tr>
<td>Summary points</td>
<td>275</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>30 PUBLICATION BIAS</strong></td>
<td>277</td>
</tr>
<tr>
<td>Introduction</td>
<td>277</td>
</tr>
<tr>
<td>The problem of missing studies</td>
<td>278</td>
</tr>
<tr>
<td>Methods for addressing bias</td>
<td>280</td>
</tr>
<tr>
<td>Illustrative example</td>
<td>281</td>
</tr>
<tr>
<td>The model</td>
<td>281</td>
</tr>
<tr>
<td>Getting a sense of the data</td>
<td>281</td>
</tr>
<tr>
<td>Is there evidence of any bias?</td>
<td>283</td>
</tr>
<tr>
<td>Is the entire effect an artifact of bias?</td>
<td>284</td>
</tr>
<tr>
<td>How much of an impact might the bias have?</td>
<td>286</td>
</tr>
<tr>
<td>Summary of the findings for the illustrative example</td>
<td>289</td>
</tr>
<tr>
<td>Some important caveats</td>
<td>290</td>
</tr>
<tr>
<td>Small-study effects</td>
<td>291</td>
</tr>
<tr>
<td>Concluding remarks</td>
<td>291</td>
</tr>
<tr>
<td>Summary points</td>
<td>291</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PART 7: ISSUES RELATED TO EFFECT SIZE</strong></td>
<td>295</td>
</tr>
<tr>
<td><strong>31 OVERVIEW</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>32 EFFECT SIZES RATHER THAN p-VALUES</strong></td>
<td>297</td>
</tr>
<tr>
<td>Introduction</td>
<td>297</td>
</tr>
<tr>
<td>Relationship between p-values and effect sizes</td>
<td>297</td>
</tr>
<tr>
<td>The distinction is important</td>
<td>299</td>
</tr>
<tr>
<td>The p-value is often misinterpreted</td>
<td>300</td>
</tr>
<tr>
<td>Narrative reviews vs. meta-analyses</td>
<td>301</td>
</tr>
<tr>
<td>Summary points</td>
<td>302</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>33 SIMPSON'S PARADOX</strong></td>
<td>303</td>
</tr>
<tr>
<td>Introduction</td>
<td>303</td>
</tr>
<tr>
<td>Circumcision and risk of HIV infection</td>
<td>303</td>
</tr>
<tr>
<td>An example of the paradox</td>
<td>305</td>
</tr>
<tr>
<td>Summary points</td>
<td>308</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>34 GENERALITY OF THE BASIC INVERSE-VARIANCE METHOD</strong></td>
<td>311</td>
</tr>
<tr>
<td>Introduction</td>
<td>311</td>
</tr>
<tr>
<td>Other effect sizes</td>
<td>312</td>
</tr>
<tr>
<td>Other methods for estimating effect sizes</td>
<td>315</td>
</tr>
<tr>
<td>Individual participant data meta-analyses</td>
<td>316</td>
</tr>
</tbody>
</table>
Bayesian approaches 318
Summary points 319

PART 8: FURTHER METHODS

35 OVERVIEW 323

36 META-ANALYSIS METHODS BASED ON DIRECTION AND p-VALUES 325
   Introduction 325
   Vote counting 325
   The sign test 325
   Combining p-values 326
   Summary points 330

37 FURTHER METHODS FOR DICHOTOMOUS DATA 331
   Introduction 331
   Mantel-Haenszel method 331
   One-step (Peto) formula for odds ratio 336
   Summary points 339

38 PSYCHOMETRIC META-ANALYSIS 341
   Introduction 341
   The attenuating effects of artifacts 342
   Meta-analysis methods 344
   Example of psychometric meta-analysis 346
   Comparison of artifact correction with meta-regression 348
   Sources of information about artifact values 349
   How heterogeneity is assessed 349
   Reporting in psychometric meta-analysis 350
   Concluding remarks 351
   Summary points 351

PART 9: META-ANALYSIS IN CONTEXT

39 OVERVIEW 355

40 WHEN DOES IT MAKE SENSE TO PERFORM A META-ANALYSIS? 357
   Introduction 357
   Are the studies similar enough to combine? 358
   Can I combine studies with different designs? 359
   How many studies are enough to carry out a meta-analysis? 363
   Summary points 364

41 REPORTING THE RESULTS OF A META-ANALYSIS 365
   Introduction 365
   The computational model 366