Contents

List of Tables xv
List of Figures xix
Acknowledgements xxv
Preface xxvii
Preface to the Second Edition xxxv
Website xxxvii

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 11
Consistency of effects 12
Summary points 13

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 31

5 EFFECT SIZES BASED ON BINARY DATA (2 × 2 TABLES) 33
Introduction 33
Risk ratio 33
Odds ratio 35
Risk difference 37
Choosing an effect size index 38
Summary points 38

6 EFFECT SIZES BASED ON CORRELATIONS 39
Introduction 39
Computing *r* 39
Other approaches 40
Summary points 41

7 CONVERTING AMONG EFFECT SIZES 43
Introduction 43
Converting from the log odds ratio to *d* 44
Converting from *d* to the log odds ratio 45
Converting from *r* to *d* 45
Converting from *d* to *r* 46
Summary points 47

8 FACTORS THAT AFFECT PRECISION 49
Introduction 49
Factors that affect precision 50
Sample size 50
Study design 51
Summary points 53

9 CONCLUDING REMARKS 55

PART 3: FIXED-EFFECT VERSUS RANDOM-EFFECTS MODELS

10 OVERVIEW 59
Introduction 59
Nomenclature 60

11 FIXED-EFFECT MODEL 61
Introduction 61
The true effect size 61
Impact of sampling error 61
Performing a fixed-effect meta-analysis 63
Summary points 64

12 RANDOM-EFFECTS MODEL
Introduction
The true effect sizes
Impact of sampling error
Performing a random-effects meta-analysis
Summary points

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

14 WORKED EXAMPLES (PART 1)
Introduction
Worked example for continuous data (Part 1)
Worked example for binary data (Part 1)
Worked example for correlational data (Part 1)
Summary points

PART 4: HETEROGENEITY

15 OVERVIEW
Introduction
Nomenclature
Worked examples

16 IDENTIFYING AND QUANTIFYING HETEROGENEITY
Introduction
Isolating the variation in true effects
Computing Q
Estimating τ^2
The I^2 statistic
Comparing the measures of heterogeneity
Confidence intervals for τ^2
Confidence intervals (or uncertainty intervals) for I^2
Summary points

17 PREDICTION INTERVALS 119
Introduction 119
Prediction intervals in primary studies 119
Prediction intervals in meta-analysis 121
Confidence intervals and prediction intervals 123
Comparing the confidence interval with the prediction interval 123
Summary points 125

18 WORKED EXAMPLES (PART 2) 127
Introduction 127
Worked example for continuous data (Part 2) 127
Worked example for binary data (Part 2) 131
Worked example for correlational data (Part 2) 134
Summary points 138

19 AN INTUITIVE LOOK AT HETEROGENEITY 139
Introduction 139
Motivating example 140
The Q-value and the p-value do not tell us how much the effect size varies 141
The confidence interval does not tell us how much the effect size varies 142
The I^2 statistic does not tell us how much the effect size varies 142
What I^2 tells us 142
The I^2 index vs. the prediction interval 145
The prediction interval 145
Prediction interval is clear, concise, and relevant 147
Computing the prediction interval 147
How to use I^2 149
How to explain heterogeneity 149
How much does the effect size vary across studies? 150
Caveats 150
Conclusion 150
Further reading 151
Summary points 151
The meaning of I^2 in Figure 19.2 151

20 CLASSIFYING HETEROGENEITY AS LOW, MODERATE, OR HIGH 155
Introduction 155
Interest should generally focus on an index of absolute heterogeneity 155
The classifications lead themselves to mistakes of interpretation 158
Classifications focus attention in the wrong direction 158
Summary points 158

PART 5: EXPLAINING HETEROGENEITY

21 SUBGROUP ANALYSES 161
Introduction 161
Fixed-effect model within subgroups 163
Computational models 172
Random effects with separate estimates of τ^2 174
Random effects with pooled estimate of τ^2 181
The proportion of variance explained 189
Mixed-effects model 192
Obtaining an overall effect in the presence of subgroups 193
Summary points 195

22 META-REGRESSION 197
Introduction 197
Fixed-effect model 198
Fixed or random effects for unexplained heterogeneity 203
Random-effects model 206
Summary points 212

23 NOTES ON SUBGROUP ANALYSES AND META-REGRESSION 213
Introduction 213
Computational model 213
Multiple comparisons 216
Software 216
Analyses of subgroups and regression analyses are observational 217
Statistical power for subgroup analyses and meta-regression 218
Summary points 219

PART 6: PUTTING IT ALL IN CONTEXT

24 LOOKING AT THE WHOLE PICTURE 223
Introduction 223
Methylphenidate for adults with ADHD 226
Impact of GLP-1 mimetics on blood pressure 228
Augmenting clozapine with a second antipsychotic 228
Conclusions 231
Caveats 231
Summary points 232

25 LIMITATIONS OF THE RANDOM-EFFECTS MODEL 233
Introduction 233
Assumptions of the random-effects model 234

A textbook case 234
When studies are pulled from the literature 235
A useful fiction 237
Transparency 238
A narrowly defined universe 238
Two important caveats 239
In context 239
Extreme cases 240
Summary points 241

26 KNAPP–HARTUNG ADJUSTMENT 243
Introduction 243
Adjustment is rarely employed in simple analyses 243
Adjusting the standard error 244
The Knapp–Hartung adjustment for other effect size indices 246
t distribution vs. *Z* distribution 247
Limitations of the Knapp–Hartung adjustment 248
Summary points 249

PART 7: COMPLEX DATA STRUCTURES

27 OVERVIEW 253

28 INDEPENDENT SUBGROUPS WITHIN A STUDY 255
Introduction 255
Combining across subgroups 255
Comparing subgroups 260
Summary points 260

29 MULTIPLE OUTCOMES OR TIME-POINTS WITHIN A STUDY 263
Introduction 263
Combining across outcomes or time-points 264
Comparing outcomes or time-points within a study 270
Summary points 275

30 MULTIPLE COMPARISONS WITHIN A STUDY 277
Introduction 277
Combining across multiple comparisons within a study 277
Differences between treatments 278
Summary points 279

31 NOTES ON COMPLEX DATA STRUCTURES 281
Introduction 281
Summary effect 281
Differences in effect 282

PART 8: OTHER ISSUES

32 OVERVIEW 287

33 VOTE COUNTING – A NEW NAME FOR AN OLD PROBLEM 289
Introduction 289
Why vote counting is wrong 290
Vote counting is a pervasive problem 291
Summary points 293

34 POWER ANALYSIS FOR META-ANALYSIS 295
Introduction 295
A conceptual approach 295
In context 299
When to use power analysis 300
Planning for precision rather than for power 301
Power analysis in primary studies 301
Power analysis for meta-analysis 304
Power analysis for a test of homogeneity 309
Summary points 312

35 PUBLICATION BIAS 313
Introduction 313
The problem of missing studies 314
Methods for addressing bias 316
Illustrative example 317
The model 317
Getting a sense of the data 318
Is there evidence of any bias? 320
How much of an impact might the bias have? 320
Summary of the findings for the illustrative example 324
Conflating bias with the small-study effect 325
Using logic to disentangle bias from small-study effects 326
These methods do not give us the 'correct' effect size 327
Some important caveats 327
Procedures do not apply to studies of prevalence 328
The model for publication bias is simplistic 328
Concluding remarks 329
Putting it all together 330
Summary points 330

PART 9: ISSUES RELATED TO EFFECT SIZE

36 OVERVIEW 335

37 EFFECT SIZES RATHER THAN *p*-VALUES **337**
Introduction 337
Relationship between *p*-values and effect sizes 337
The distinction is important 339
The *p*-value is often misinterpreted 340
Narrative reviews vs. meta-analyses 341
Summary points 342

38 SIMPSON'S PARADOX **343**
Introduction 343
Circumcision and risk of HIV infection 343
An example of the paradox 345
Summary points 348

39 GENERALITY OF THE BASIC INVERSE-VARIANCE METHOD **349**
Introduction 349
Other effect sizes 350
Other methods for estimating effect sizes 353
Individual participant data meta-analyses 354
Bayesian approaches 355
Summary points 357

PART 10: FURTHER METHODS

40 OVERVIEW **361**

41 META-ANALYSIS METHODS BASED ON DIRECTION AND *p*-VALUES **363**
Introduction 363
Vote counting 363
The sign test 363
Combining *p*-values 364
Summary points 368

42 FURTHER METHODS FOR DICHOTOMOUS DATA **369**
Introduction 369
Mantel–Haenszel method 369
One-step (Peto) formula for odds ratio 373
Summary points 376

43 PSYCHOMETRIC META-ANALYSIS **377**
Introduction 377
The attenuating effects of artifacts 378
Meta-analysis methods 380
Example of psychometric meta-analysis 381
Comparison of artifact correction with meta-regression 384

Sources of information about artifact values 384
How heterogeneity is assessed 385
Reporting in psychometric meta-analysis 386
Concluding remarks 386
Summary points 387

PART 11: META-ANALYSIS IN CONTEXT

44 OVERVIEW 391

45 WHEN DOES IT MAKE SENSE TO PERFORM A META-ANALYSIS? 393
Introduction 393
Are the studies similar enough to combine? 394
Can I combine studies with different designs? 395
How many studies are enough to carry out a meta-analysis? 399
Summary points 400

46 REPORTING THE RESULTS OF A META-ANALYSIS 401
Introduction 401
The computational model 402
Forest plots 402
Sensitivity analysis 404
Summary points 405

47 CUMULATIVE META-ANALYSIS 407
Introduction 407
Why perform a cumulative meta-analysis? 409
Summary points 412

48 CRITICISMS OF META-ANALYSIS 413
Introduction 413
One number cannot summarize a research field 414
The file drawer problem invalidates meta-analysis 414
Mixing apples and oranges 415
Garbage in, garbage out 416
Important studies are ignored 417
Meta-analysis can disagree with randomized trials 417
Meta-analyses are performed poorly 420
Is a narrative review better? 420
Concluding remarks 422
Summary points 422

49 COMPREHENSIVE META-ANALYSIS SOFTWARE 425
Introduction 425
Features in CMA 426

Teaching elements 427
Documentation 427
Availability 427
Acknowledgments 427
Motivating example 428
Data entry 428
Basic analysis 429
What is the *average* effect size? 430
How much does the effect size vary? 430
Plot showing distribution of effects 431
High-resolution plot 432
Subgroup analysis 433
Meta-regression 435
Publication bias 438
Explaining results 439

50 HOW TO EXPLAIN THE RESULTS OF AN ANALYSIS 443
Introduction 443
The overview 444
The mean effect size 444
Variation in effect size 444
Notations 444
Impact of resistance exercise on pain 445
Correlation between letter knowledge and word recognition 450
Statins for prevention of cardiovascular events 455
Bupropion for smoking cessation 460
Mortality following mitral-valve procedures in elderly patients 465

PART 12: RESOURCES

51 SOFTWARE FOR META-ANALYSIS 471
Comprehensive meta-analysis 471
Metafor 471
Stata 472
Revman 472

52 WEB SITES, SOCIETIES, JOURNALS, AND BOOKS 473
Web sites 473
Professional societies 476
Journals 476
Special issues dedicated to meta-analysis 477
Books on systematic review methods and meta-analysis 477

REFERENCES 479

INDEX 491