

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.



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 (2×2 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

Performing a fixed-effect meta-analysis	65
Summary points	67
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	
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 τ^2	114
The I^2 statistic	117

Comparing the measures of heterogeneity	119
Confidence intervals for τ^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 τ^2	164
Random effects with pooled estimate of τ^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

Statistical power for subgroup analyses and meta-regression	210
Summary points	211

PART 5: COMPLEX DATA STRUCTURES

22 OVERVIEW	215
23 INDEPENDENT SUBGROUPS WITHIN A STUDY	217
Introduction	217
Combining across subgroups	218
Comparing subgroups	222
Summary points	223
24 MULTIPLE OUTCOMES OR TIME-POINTS WITHIN A STUDY	225
Introduction	225
Combining across outcomes or time-points	226
Comparing outcomes or time-points within a study	233
Summary points	238
25 MULTIPLE COMPARISONS WITHIN A STUDY	239
Introduction	239
Combining across multiple comparisons within a study	239
Differences between treatments	240
Summary points	241
26 NOTES ON COMPLEX DATA STRUCTURES	243
Introduction	243
Summary effect	243
Differences in effect	244

PART 6: OTHER ISSUES

27 OVERVIEW	249
28 VOTE COUNTING – A NEW NAME FOR AN OLD PROBLEM	251
Introduction	251
Why vote counting is wrong	252
Vote counting is a pervasive problem	253
Summary points	255
29 POWER ANALYSIS FOR META-ANALYSIS	257
Introduction	257
A conceptual approach	257
In context	261
When to use power analysis	262

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

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

Forest plots	366
Sensitivity analysis	368
Summary points	369
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

List of Tables

Table 3.1	Roadmap of formulas in subsequent chapters	19
Table 5.1	Nomenclature for 2×2 table of outcome by treatment	33
Table 5.2	Fictional data for a 2×2 table	33
Table 8.1	Impact of sample size on variance	52
Table 8.2	Impact of study design on variance	54
Table 14.1	Dataset 1 – Part A (basic data)	88
Table 14.2	Dataset 1 – Part B (fixed-effect computations)	88
Table 14.3	Dataset 1 – Part C (random-effects computations)	88
Table 14.4	Dataset 2 – Part A (basic data)	93
Table 14.5	Dataset 2 – Part B (fixed-effect computations)	93
Table 14.6	Dataset 2 – Part C (random-effects computations)	93
Table 14.7	Dataset 3 – Part A (basic data)	98
Table 14.8	Dataset 3 – Part B (fixed-effect computations)	98
Table 14.9	Dataset 3 – Part C (random-effects computations)	98
Table 16.1	Factors affecting measures of dispersion	119
Table 18.1	Dataset 1 – Part D (intermediate computations)	136
Table 18.2	Dataset 1 – Part E (variance computations)	136
Table 18.3	Dataset 2 – Part D (intermediate computations)	140
Table 18.4	Dataset 2 – Part E (variance computations)	140
Table 18.5	Dataset 3 – Part D (intermediate computations)	144
Table 18.6	Dataset 3 – Part E (variance computations)	144
Table 19.1	Fixed effect model – computations	152
Table 19.2	Fixed-effect model – summary statistics	155
Table 19.3	Fixed-effect model – ANOVA table	158
Table 19.4	Fixed-effect model – subgroups as studies	159
Table 19.5	Random-effects model (separate estimates of τ^2) – computations	165
Table 19.6	Random-effects model (separate estimates of τ^2) – summary statistics	167
Table 19.7	Random-effects model (separate estimates of τ^2) – ANOVA table	169
Table 19.8	Random-effects model (separate estimates of τ^2) – subgroups as studies	171
Table 19.9	Statistics for computing a pooled estimate of τ^2	173
Table 19.10	Random-effects model (pooled estimate of τ^2) – computations	173

Table 19.11	Random-effects model (pooled estimate of τ^2) – summary statistics	175
Table 19.12	Random-effects model (pooled estimate of τ^2) – ANOVA table	178
Table 19.13	Random-effects model (pooled estimate of τ^2) – subgroups as studies	179
Table 20.1	The BCG dataset	190
Table 20.2	Fixed-effect model – Regression results for BCG	190
Table 20.3	Fixed-effect model – ANOVA table for BCG regression	191
Table 20.4	Random-effects model – regression results for BCG	197
Table 20.5	Random-effects model – test of the model	198
Table 20.6	Random-effects model – comparison of model (latitude) versus the null model	202
Table 23.1	Independent subgroups – five fictional studies	218
Table 23.2	Independent subgroups – summary effect	219
Table 23.3	Independent subgroups – synthetic effect for study 1	220
Table 23.4	Independent subgroups – summary effect across studies	220
Table 24.1	Multiple outcomes – five fictional studies	226
Table 24.2	Creating a synthetic variable as the mean of two outcomes	227
Table 24.3	Multiple outcomes – summary effect	230
Table 24.4	Multiple outcomes – Impact of correlation on variance of summary effect	231
Table 24.5	Creating a synthetic variable as the difference between two outcomes	233
Table 24.6	Multiple outcomes – difference between outcomes	235
Table 24.7	Multiple outcomes – Impact of correlation on the variance of difference	237
Table 33.1	HIV as function of circumcision (by subgroup)	304
Table 33.2	HIV as function of circumcision – by study	305
Table 33.3	HIV as a function of circumcision – full population	306
Table 33.4	HIV as a function of circumcision – by risk group	306
Table 33.5	HIV as a function of circumcision/risk group – full population	307
Table 34.1	Simple example of a genetic association study	314
Table 36.1	Streptokinase data – calculations for meta-analyses of p -values	329
Table 37.1	Nomenclature for 2×2 table of events by treatment	331
Table 37.2	Mantel-Haenszel – odds ratio	333
Table 37.3	Mantel-Haenszel – variance of summary effect	334
Table 37.4	One-step – odds ratio and variance	338
Table 38.1	Fictional data for psychometric meta-analysis	346
Table 38.2	Observed (attenuated) correlations	346
Table 38.3	Unattenuated correlations	347

List of Figures

Figure 1.1	High-dose versus standard-dose of statins (adapted from Cannon <i>et al.</i> , 2006)	4
Figure 2.1	Impact of streptokinase on mortality (adapted from Lau <i>et al.</i> , 1992)	10
Figure 4.1	Response ratios are analyzed in log units	31
Figure 5.1	Risk ratios are analyzed in log units	34
Figure 5.2	Odds ratios are analyzed in log units	36
Figure 6.1	Correlations are analyzed in Fisher's z units	42
Figure 7.1	Converting among effect sizes	46
Figure 8.1	Impact of sample size on variance	53
Figure 8.2	Impact of study design on variance	54
Figure 10.1	Symbols for true and observed effects	62
Figure 11.1	Fixed-effect model – true effects	64
Figure 11.2	Fixed-effect model – true effects and sampling error	64
Figure 11.3	Fixed-effect model – distribution of sampling error	65
Figure 12.1	Random-effects model – distribution of true effects	70
Figure 12.2	Random-effects model – true effects	70
Figure 12.3	Random-effects model – true and observed effect in one study	71
Figure 12.4	Random-effects model – between-study and within-study variance	72
Figure 13.1	Fixed-effect model – forest plot showing relative weights	78
Figure 13.2	Random-effects model – forest plot showing relative weights	78
Figure 13.3	Very large studies under fixed-effect model	80
Figure 13.4	Very large studies under random-effects model	80
Figure 14.1	Forest plot of Dataset 1 – fixed-effect weights	89
Figure 14.2	Forest plot of Dataset 1 – random-effects weights	89
Figure 14.3	Forest plot of Dataset 2 – fixed-effect weights	94
Figure 14.4	Forest plot of Dataset 2 – random-effects weights	94
Figure 14.5	Forest plot of Dataset 3 – fixed-effect weights	99
Figure 14.6	Forest plot of Dataset 3 – random-effects weights	99
Figure 16.1	Dispersion across studies relative to error within studies	108
Figure 16.2	Q in relation to df as measure of dispersion	110

Figure 16.3	Flowchart showing how T^2 and I^2 are derived from Q and df	111
Figure 16.4	Impact of Q and number of studies on the p -value	113
Figure 16.5	Impact of excess dispersion and absolute dispersion on T^2	115
Figure 16.6	Impact of excess and absolute dispersion on T	116
Figure 16.7	Impact of excess dispersion on I^2	118
Figure 16.8	Factors affecting T^2 but not I^2	120
Figure 16.9	Factors affecting I^2 but not T^2	121
Figure 17.1	Prediction interval based on population parameters μ and τ^2	130
Figure 17.2	Prediction interval based on sample estimates M^* and T^2	130
Figure 17.3	Simultaneous display of confidence interval and prediction interval	131
Figure 17.4	Impact of number of studies on confidence interval and prediction interval	132
Figure 18.1	Forest plot of Dataset 1 – random-effects weights with prediction interval	136
Figure 18.2	Forest plot of Dataset 2 – random-effects weights with prediction interval	140
Figure 18.3	Forest plot of Dataset 3 – random-effects weights with prediction interval	144
Figure 19.1	Fixed-effect model – studies and subgroup effects	151
Figure 19.2	Fixed-effect – subgroup effects	155
Figure 19.3	Fixed-effect model – treating subgroups as studies	159
Figure 19.4	Flowchart for selecting a computational model	163
Figure 19.5	Random-effects model (separate estimates of τ^2) – studies and subgroup effects	164
Figure 19.6	Random-effects model (separate estimates of τ^2) – subgroup effects	167
Figure 19.7	Random-effects model (separate estimates of τ^2) – treating subgroups as studies	170
Figure 19.8	Random-effects model (pooled estimate of τ^2) – studies and subgroup effects	172
Figure 19.9	Random-effects model (pooled estimate of τ^2) – subgroup effects	176
Figure 19.10	Random-effects model (pooled estimate of τ^2) – treating subgroups as studies	179
Figure 19.11	A primary study showing subjects within groups	180
Figure 19.12	Random-effects model – variance within and between subgroups	182
Figure 19.13	Proportion of variance explained by subgroup membership	182
Figure 20.1	Fixed-effect model – forest plot for the BCG data	189
Figure 20.2	Fixed-effect model – regression of log risk ratio on latitude	193

Figure 20.3	Fixed-effect model – population effects as function of covariate	194
Figure 20.4	Random-effects model – population effects as a function of covariate	194
Figure 20.5	Random-effects model – forest plot for the BCG data	197
Figure 20.6	Random-effects model – regression of log risk ratio on latitude	199
Figure 20.7	Between-studies variance (T^2) with no covariate	201
Figure 20.8	Between-studies variance (T^2) with covariate	201
Figure 20.9	Proportion of variance explained by latitude	202
Figure 23.1	Creating a synthetic variable from independent subgroups	219
Figure 28.1	The p -value for each study is > 0.20 but the p -value for the summary effect is < 0.02	252
Figure 29.1	Power for a primary study as a function of n and δ	267
Figure 29.2	Power for a meta-analysis as a function of number studies and δ	269
Figure 29.3	Power for a meta-analysis as a function of number studies and heterogeneity	272
Figure 30.1	Passive smoking and lung cancer – forest plot	282
Figure 30.2	Passive smoking and lung cancer – funnel plot	283
Figure 30.3	Passive smoking and lung cancer – funnel plot with imputed studies	287
Figure 30.4	Passive smoking and lung cancer – cumulative forest plot	288
Figure 32.1	Estimating the effect size versus testing the null hypothesis	298
Figure 32.2	The p -value is a poor surrogate for effect size	300
Figure 32.3	Studies where p -values differ but effect size is the same	300
Figure 32.4	Studies where p -values are the same but effect sizes differ	301
Figure 32.5	Studies where the more significant p -value corresponds to weaker effect size	301
Figure 33.1	HIV as function of circumcision – by study	304
Figure 33.2	HIV as function of circumcision – in three sets of studies	308
Figure 36.1	Effect size in four fictional studies	328
Figure 41.1	Forest plot using lines to represent the effect size	367
Figure 41.2	Forest plot using boxes to represent the effect size and relative weight	367
Figure 42.1	Impact of streptokinase on mortality – forest plot	372
Figure 42.2	Impact of streptokinase on mortality – cumulative forest plot	373
Figure 43.1	Forest plot of five fictional studies and a new trial (consistent effects)	382

Figure 43.2	Forest plot of five fictional studies and a new trial (heterogeneous effects)	383
Figure 44.1	CMA – data entry screen for 2×2 tables	395
Figure 44.2	CMA – analysis screen	396
Figure 44.3	CMA – high resolution forest plot	397
Figure 44.4	RevMan – data entry screen for 2×2 tables	398
Figure 44.5	RevMan – analysis screen	399
Figure 44.6	Stata macros – data entry screen for 2×2 tables	401
Figure 44.7	Stata macros – analysis screen	401
Figure 44.8	Stata macros – high resolution forest plot	402

How a Meta-Analysis Works

Introduction
Individual studies
The summary effect
Heterogeneity of effect sizes

INTRODUCTION

Figure 1.1 illustrates a meta-analysis that shows the impact of high dose versus standard dose of statins in preventing death and myocardial infarction (MI). This analysis is adapted from one reported by Cannon *et al.* and published in the *Journal of the American College of Cardiology* (2006).

Our goal in presenting this here is to introduce the various elements in a meta-analysis (the effect size for each study, the weight assigned to each effect size, the estimate of the summary effect, and so on) and show where each fits into the larger scheme. In the chapters that follow, each of these elements will be explored in detail.

INDIVIDUAL STUDIES

The first four rows on this plot represent the four studies. For each, the study name is shown at left, followed by the effect size, the relative weight assigned to the study for computing the summary effect, and the *p*-value. The effect size and weight are also shown schematically.

Effect size

The effect size, a value which reflects the magnitude of the treatment effect or (more generally) the strength of a relationship between two variables, is the unit of currency in a meta-analysis. We compute the effect size for each study, and then

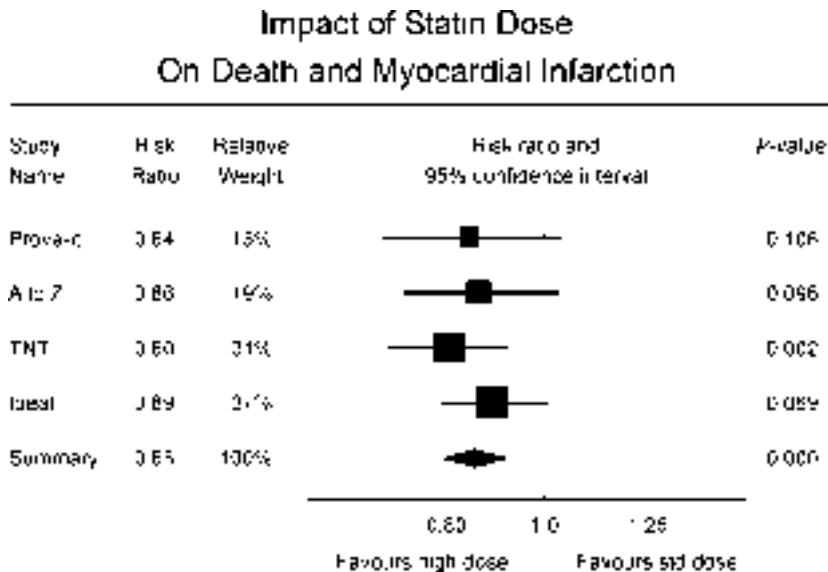


Figure 1.1 High-dose versus standard-dose of statins (adapted from Cannon *et al.*, 2006).

work with the effect sizes to assess the consistency of the effect across studies and to compute a summary effect.

The effect size could represent the impact of an intervention, such as the impact of medical treatment on risk of infection, the impact of a teaching method on test scores, or the impact of a new protocol on the number of salmon successfully returning upstream. The effect size is not limited to the impact of interventions, but could represent *any relationship* between two variables, such as the difference in test scores for males versus females, the difference in cancer rates for persons exposed or not exposed to second-hand smoke, or the difference in cardiac events for persons with two distinct personality types. In fact, what we generally call an *effect size* could refer simply to the estimate of a single value, such as the prevalence of Lyme disease.

In this example the effect size is the risk ratio. A risk ratio of 1.0 would mean that the risk of death or MI was the same in both groups, while a risk ratio less than 1.0 would mean that the risk was lower in the high-dose group, and a risk ratio greater than 1.0 would mean that the risk was lower in the standard-dose group.

The effect size for each study is represented by a square, with the location of the square representing both the direction and magnitude of the effect. Here, the effect size for each study falls to the left of center (indicating a benefit for the high-dose group). The effect is strongest (most distant from the center) in the *TNT* study and weakest in the *Ideal* study.

Note. For measures of effect size based on ratios (as in this example) a ratio of 1.0 represents no difference between groups. For measures of effect based on differences (such as mean difference), a difference of 0.0 represents no difference between groups.

Precision

In the schematic, the effect size for each study is bounded by a confidence interval, reflecting the precision with which the effect size has been estimated in that study. The confidence interval for the last study (*Ideal*) is noticeably narrower than that for the first study (*Prove-it*), reflecting the fact that the *Ideal* study has greater precision. The meaning of precision and the factors that affect precision are discussed in Chapter 8.

Study weights

The solid squares that are used to depict each of the studies vary in size, with the size of each square reflecting the weight that is assigned to the corresponding study when we compute the summary effect. The *TNT* and *Ideal* studies are assigned relatively high weights, while somewhat less weight is assigned to the *A to Z* study and still less to the *Prove-it* study.

As one would expect, there is a relationship between a study's precision and that study's weight in the analysis. Studies with relatively good precision (*TNT* and *Ideal*) are assigned more weight while studies with relatively poor precision (*Prove-it*) are assigned less weight. Since precision is driven primarily by sample size, we can think of the studies as being weighted by sample size.

However, while precision is one of the elements used to assign weights, there are often other elements as well. In Part 3 we discuss different assumptions that one can make about the distribution of effect sizes across studies, and how these affect the weight assigned to each study.

p-values

For each study we show the *p*-value for a test of the null. There is a necessary correspondence between the *p*-value and the confidence interval, such that the *p*-value will fall under 0.05 if and only if the 95% confidence interval does not include the null value. Therefore, by scanning the confidence intervals we can easily identify the statistically significant studies. The role of *p*-values in the analysis, as well as the relationship between *p*-values and effect size, is discussed in Chapter 32.

In this example, for three of the four studies the confidence interval crosses the null, and the *p*-value is greater than 0.05. In one (the *TNT* study) the confidence interval does not cross the null, and the *p*-value falls under 0.05.

THE SUMMARY EFFECT

One goal of the synthesis is usually to compute a summary effect. Typically we report the effect size itself, as well as a measure of precision and a *p*-value.