CHAPTER 20

Meta-Regression

INTRODUCTION

In primary studies we use regression, or multiple regression, to assess the relationship between one or more covariates (moderators) and a dependent variable. Essentially the same approach can be used with meta-analysis, except that the covariates are at the level of the study rather than the level of the subject, and the dependent variable is the effect size in the studies rather than subject scores. We use the term *meta-regression* to refer to these procedures when they are used in a meta-analysis.

The differences that we need to address as we move from primary studies to meta-analysis for regression are similar to those we needed to address as we moved from primary studies to meta-analysis for subgroup analyses. These include the need to assign a weight to each study and the need to select the appropriate model (fixed versus random effects). Also, as was true for subgroup analyses, the R^2 index, which is used to quantify the proportion of variance explained by the covariates, must be modified for use in meta-analysis.

With these modifications, however, the full arsenal of procedures that fall under the heading of multiple regression becomes available to the meta-analyst. We can work with sets of covariates, such as three variables that together define a treatment, or that allow for a nonlinear relationship between covariates and the effect size. We can enter covariates into the analysis using a pre-defined sequence and assess the impact of any set, over and above the impact of prior sets, to control for confounding variables. We can incorporate both categorical (for example, dummy-coded) and continuous variables as covariates. We can use these procedures both to assess the

Introduction to Meta-Analysis. Michael Borenstein, L. V. Hedges, J. P. T. Higgins and H. R. Rothstein
 ISBN: 978-0-470-05724-7

impact of covariates and also to predict the effect size in studies with specific characteristics.

Multiple regression incorporates a wide array of procedures, and we cannot cover these fully in this volume. Rather, we assume that the reader is familiar with multiple regression in primary studies, and our goal here is to show how the same techniques used in primary studies can be applied to meta-regression.

As is true in primary studies, where we need an appropriately large ratio of *subjects* to covariates in order for the analysis be to meaningful, in meta-analysis we need an appropriately large ratio of *studies* to covariates. Therefore, the use of meta-regression, especially with multiple covariates, is not a recommended option when the number of studies is small. In primary studies some have recommended a ratio of at least ten subjects for each covariate, which would correspond to ten studies for each covariate in meta-regression. In fact, though, there are no hard and fast rules in either case.

FIXED-EFFECT MODEL

As we did when discussing subgroup analysis, we start with the fixed-effect model, which is simpler, and then move on to the random-effects model, which is generally more appropriate.

The BCG data set

Various researchers have published studies that assessed the impact of a vaccine, known as BCG, to prevent the development of tuberculosis (TB). With the re-emergence of TB in the United States in recent years (including many drug-resistant cases), researchers needed to determine whether or not the BCG vaccine should be recommended. For that reason, Colditz *et al*. (1994) reported a meta-analysis of these studies, and Berkey *et al*. (1995) showed how meta-regression could be used in an attempt to explain some of the variance in treatment effects.

The forest plot is shown in Figure 20.1. The effect size is the risk ratio, with a risk ratio of 0.10 indicating that the vaccine reduced the risk of TB by 90%, a risk ratio of 1.0 indicating no effect, and risk ratios higher than 1.0 indicating that the vaccine increased the risk of TB. Studies are sorted from most effective to least effective. As always for an analysis of risk ratios, the analysis was performed using log transformed values and then converted back to risk ratios for presentation.

Using a fixed-effect analysis the risk ratio for the 13 studies is 0.650 with a confidence interval of 0.601 to 0.704, which says that the vaccine decreased the risk of TB by at least 30% and possibly by as much as 40%. In log units the risk ratio is -0.430 with a standard error of 0.040. The Z-value is -10.625 ($p < 0.0001$), which allows us to reject the null hypothesis of no effect.

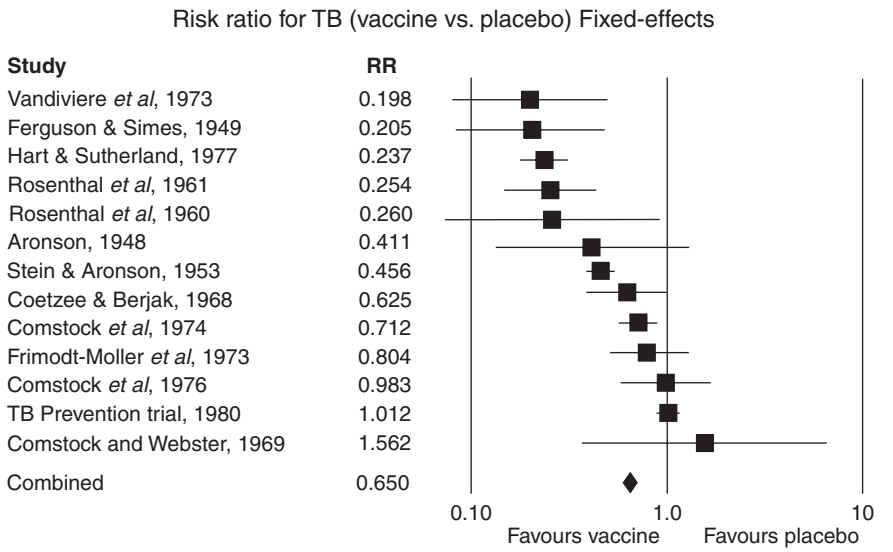

Figure 20.1 Fixed-effect model – forest plot for the BCG data.

At least as interesting, however, is the substantial variation in the treatment effects, which ranged from a risk ratio of 0.20 (an 80% reduction in risk) to 1.56 (a 56% *increase* in risk). While some of this variation is due to within-study error, some of it reflects variation in the true effects. The Q-statistic is 152.233 with $df = 12$ and $p < 0.0001$. T^2 is 0.309, which yields a prediction interval of approximately 0.14 to 1.77, meaning that the true effect (risk ratio) in the next study could fall anywhere in the range of 0.14 to 1.77. The value of I^2 is 92.12, which means that 92% of the observed variance comes from real differences between studies and, as such, can potentially be explained by study-level covariates.

The next issue for the authors was to try and explain some of this variation. There was reason to believe that the drug was more effective in colder climates. This hypothesis was based on the theory that persons in colder climates were less likely to have a natural immunity to TB. It was also based on the expectation that the drug would be more potent in the colder climates (in warmer climates the heat would cause the drug to lose potency, and direct exposure to sunlight could kill some of the bacteria that were required for the vaccine to work properly).

In the absence of better predictor variables (such as the actual storage conditions used for the vaccine) Berkey *et al*. (1995) used absolute distance from the equator as a surrogate for climate (i.e. geographical regions in the Northern US would be colder than those in the tropics), and used regression to look for a relationship between *Distance* and treatment effect. Given the *post hoc* nature of this analysis, a positive finding would probably not be definitive, but would suggest a direction for additional research. (See also Sutton *et al*., 2000; Egger *et al*., 2001.)

Assessing the impact of the slope

Table 20.1 shows the data for each study (events and sample size, effect size and latitude). Table 20.2 shows the results for a meta-regression using absolute latitude to predict the log risk ratio.

The regression coefficient for latitude is −0.0292, which means that every one degree of latitude corresponds to a decrease of 0.0292 units in effect size. In this case, the effect size is the log risk ratio, and (given the specifics of this example) this corresponds to a more effective vaccination.

If we were working with a primary study and wanted to test the coefficient for significance we might use a *t*-test of the form

$$t = \frac{B}{SE_B}, \tag{20.1}$$

In meta-analysis the coefficient for any covariate (B) and its standard error are based on groups of *studies* rather than groups of *subjects* but the same logic applies. Historically, in meta-regression the test is based on the *Z*-distribution, and that is the

Table 20.1 The BCG dataset.

	Vaccinated		Control					
	TB	Total	TB	Total	*RR*	ln*RR*	V_{lnRR}	Latitude
Vandiviere et al, 1973	8	2545	10	629	0.198	−1.621	0.223	19
Ferguson & Simes, 1949	6	306	29	303	0.205	−1.585	0.195	55
Hart & Sutherland, 1977	62	13598	248	12867	0.237	−1.442	0.020	52
Rosenthal *et al*, 1961	17	1716	65	1665	0.254	−1.371	0.073	42
Rosenthal *et al*, 1960	3	231	11	220	0.260	−1.348	0.415	42
Aronson, 1948	4	123	11	139	0.411	−0.889	0.326	44
Stein & Aaronson, 1953	180	1541	372	1451	0.456	−0.786	0.007	44
Coetzee & Berjak, 1968	29	7499	45	7277	0.625	−0.469	0.056	27
Comstock *et al*, 1974	186	50634	141	27338	0.712	−0.339	0.012	18
Frimodt-Moller *et al*, 1973.	33	5069	47	5808	0.804	−0.218	0.051	13
Comstock *et al*, 1976	27	16913	29	17854	0.983	−0.017	0.071	33
TB Prevention Trial, 1980	505	88391	499	88391	1.012	0.012	0.004	13
Comstock & Webster, 1969	5	2498	3	2341	1.562	0.446	0.533	33

Table 20.2 Fixed-effect model – Regression results for BCG.

	Fixed effect, Z-Distribution					
	Point estimate	Standard error	95% Lower	95% Upper	*Z*-value	*p*-Value
Intercept	0.34356	0.08105	0.18471	0.50242	4.23899	0.00002
Latitude	−0.02924	0.00265	−0.03444	−0.02404	−11.02270	0.00000

Table 20.3 Fixed-effect model – ANOVA table for BCG regression.

Analysis of variance			
	Q	*df*	*p*-Value
Model (Q_{model})	121.49992	1	0.00000
Residual (Q_{resid})	30.73309	11	0.00121
Total (Q)	152.23301	12	0.00000

approach presented here (however, see notes at the end of this chapter about other approaches). The statistic to test the significance of the slope is

$$Z = \frac{B}{SE_B}. \tag{20.2}$$

Under the null hypothesis that the coefficient is zero, Z would follow the normal distribution.

In the running example the coefficient for latitude is -0.02924 with standard error 0.00265, so

$$Z = \frac{-0.02924}{0.00265} = -11.0227.$$

The two-tailed p-value corresponding to $Z = -11.0227$ is < 0.00001. This tells us that the slope is probably not zero, and the vaccination is more effective when the study is conducted at a greater distance from the equator.

The Z-test can be used to test the statistical significance of any single coefficient but when we want to assess the impact of several covariates simultaneously we need to use the Q-test. (This is analogous to the situation in primary studies where we use a t-test to assess the impact of one coefficient but an F-test to assess the impact of two or more.)

As we did when working with analysis of variance we can divide the sum of squares into its component parts, and create an analysis of variance table as follows.

As before, Q is defined as a weighted sum of squares, which we can partition into its component parts. Q reflects the total dispersion of studies about the grand mean. Q_{resid} reflects the distance of studies from the regression line. Q_{model} is the dispersion explained by the covariates.

Each Q statistic is evaluated with respect to its degrees of freedom, as follows.

- Q is 152.2330, with 12 degrees of freedom and $p < 0.00001$ (this is the same value presented for the initial meta-analysis with no covariates). This means that the amount of total variance is more than we would expect based on within-study error.
- Q_{model} is 121.4999 with 1 degree of freedom and $p < 0.00001$. This means that the relationship between latitude and treatment effect is stronger than we would expect by chance.

- Q_{resid} is 30.7331 with 11 degrees of freedom and $p < 0.0001$. This means that even with latitude in the model, some of the between-studies variance (reflected by the distance between the regression line and the studies) remains unexplained.

Q_{model} here is analogous to Q_{bet} for subgroup analysis, and Q_{resid} is analogous to Q_{within} for subgroup analysis. If the covariates are coded to represent subgroups, then Q_{model} will be identical to Q_{bet}, and Q_{resid} will be identical to Q_{within}.

The *Z*-test and the *Q*-test

In this example there is only one covariate and so we have the option of using either the *Z*-test or the *Q*-test to test its relationship with effect size. It follows that the two tests should yield the same results, and they do. The *Z*-value is −11.0227, with a corresponding p-value of <0.0001. The *Q*-value is 121.4999 with a corresponding *p*-value of <0.0001 (with 1 *df*). Finally, Q should be equal to Z^2 (since *Q* squares each difference while *Z* does not) and in fact -11.0227^2 is equal to 121.4999.

When we have more than one covariate the *Q* statistic serves as an omnibus test of the hypothesis that all the *B*'s are zero. The *Z*-test can be used to test any coefficient, with the others held constant.

Quantify the magnitude of the relationship

The *Z*-test, like all tests of significance, speaks to the question of statistical, rather than substantive, significance. Therefore, in addition to reporting the test of significance, one should always report the magnitude of the relationship. Here, the relationship of latitude to effect (expressed as a log risk ratio) is

$$\ln(RR) = 0.3435 - 0.0292(X)$$

where *X* is the absolute latitude. Figure 20.2 shows the plot of log risk ratio on latitude.

In the graph, each study is represented by a circle that shows the actual coordinates (observed effect size by latitude) for that study. The size (specifically, the area) of each circle is proportional to that study's weight in the analysis. Since this analysis is based on the fixed-effect model, the weight is simply the inverse of the within-study variance for each study.

The center line shows the predicted values. A study performed relatively close to the equator (such as the study performed in Madras, India, latitude 13) would have an expected effect near zero (corresponding to a risk ratio of 1.0, which means that the vaccination has no impact on TB). By contrast, a study at latitude 55 (Saskatchewan) would have an expected effect near −1.20 (corresponding to a risk ratio near 0.30, which means that the vaccination is expected to decrease the risk of TB by about 70%).

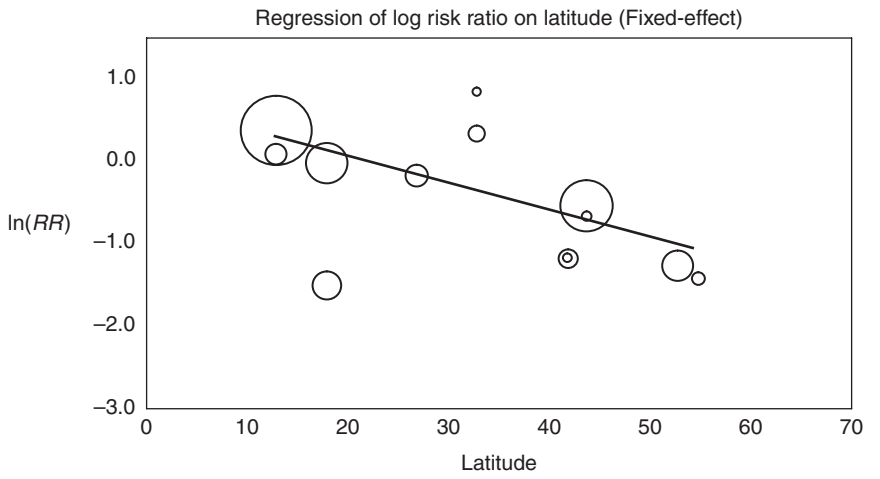

Figure 20.2 Fixed-effect model – regression of log risk ratio on latitude.

The 95% confidence interval for B is given by

$$LL_B = B - 1.96 \times SE_B \tag{20.3}$$

and

$$UL_B = B + 1.96 \times SE_B. \tag{20.4}$$

In the BCG example

$$LL_B = (-0.0292) - 1.96 \times 0.0027 = -0.0344$$

and

$$UL_B = (-0.0292) + 1.96 \times 0.0027 = -0.0240.$$

In words, the true coefficient could be as low as −0.0344 and as high as −0.0240. These limits can be used to generate confidence intervals on the plot.

FIXED OR RANDOM EFFECTS FOR UNEXPLAINED HETEROGENEITY

In Part 3 we discussed the difference between a fixed-effect model and a random-effects model. Under the fixed-effect model we assume that the true effect is the same in all studies. By contrast, under the random-effects model we allow that the true effect may vary from one study to the next.

When we were working with a single population the difference in models translated into one true effect versus a distribution of true effects for all studies. When we were working with subgroups (analysis of variance) it translated into one true effect versus a distribution of effects for all studies within a subgroup (for example, studies that used Intervention *A* or studies that used Intervention *B*).

For meta-regression it translates into one true effect versus a distribution of effects for studies with the same predicted value (for example, two studies that share the same value on all covariates). This is shown schematically in Figure 20.3 and Figure 20.4.

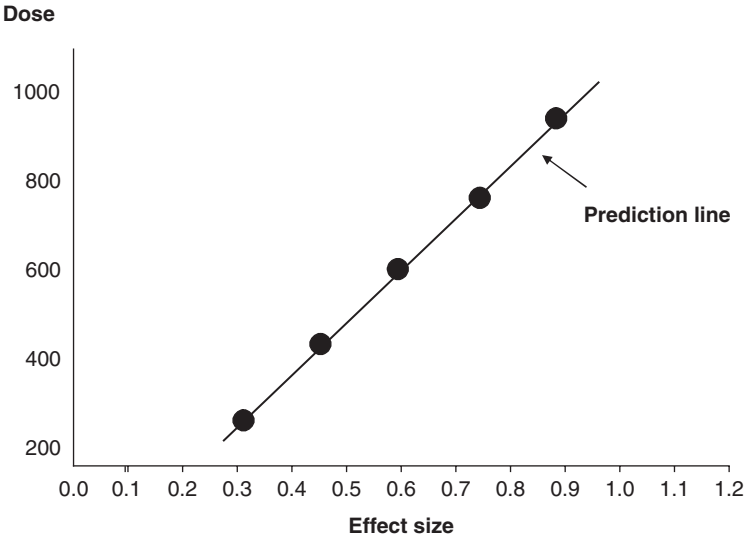

Figure 20.3 Fixed-effect model – population effects as function of covariate.

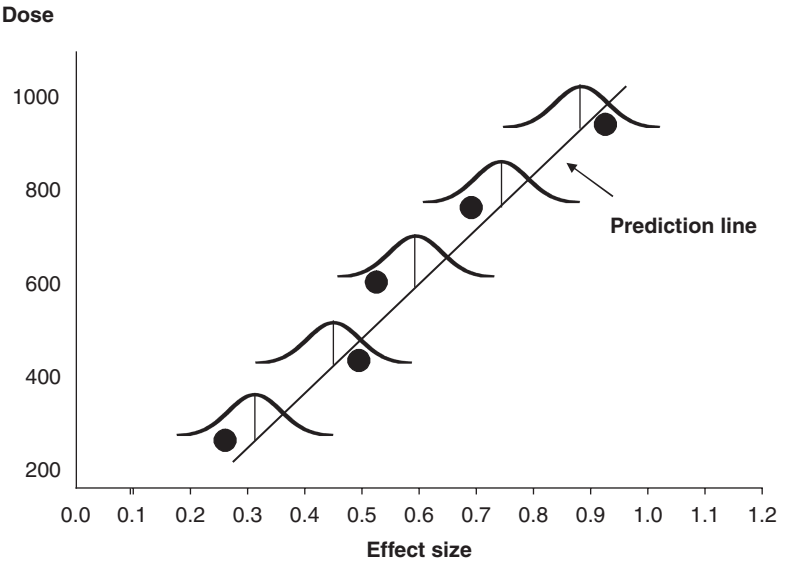

Figure 20.4 Random-effects model – population effects as a function of covariate.

Under the fixed-effect model, for any set of covariate values (that is, for any predicted value) there is one population effect size (represented by a circle in Figure 20.3). Under the random-effects model, for any predicted value there is a distribution of effect sizes (in Figure 20.4 the distribution is centered over the predicted value but the population effect size can fall to the left or right of center). (In both figures we assume that the prediction is perfect, so that the true effect (or mean effect) falls directly on the prediction line.)

As always, the selection of a method should follow the logic of how the studies were selected. When we introduced the idea of fixed versus random effects for a single population, the example we used for the fixed-effect model was a pharmaceutical company that planned a series of ten trials that were identical for all intents and purposes (page 83). When we moved on to analysis of variance we extended the same example, and assumed that five cohorts would be used to test placebo versus *Drug A*, while five would be used to test placebo versus *Drug B* (page 161). We can extend this example, and use it to create an example where a fixed-effect analysis would be appropriate for meta-regression. As before, we'll assume a total of 10 studies, five for placebo versus *Drug A* and five for placebo versus *Drug B*. This time we'll assume that each study used either 10, 20, 40, 80, or 160 mg. of the drug and we will use regression to assess the impact of Drug and Dose. The fixed-effect model makes sense here because the studies are known to be identical on all other factors.

As before, we note that this example is not representative of most systematic reviews. In the vast majority of cases, especially when the studies are performed by different researchers and then culled from the literature, it is more plausible that the impact of the covariates captures some, *but not all*, of the true variation among effects. In this case, it is the random-effects model that reflects the nature of the distribution of true effects, and should therefore be used in the analyses. Also as before, if the study design suggests that the random-effects model is appropriate, then this model should be selected *a priori*. It is a mistake to start with the fixed-effect model and move on to random effects only if the test for heterogeneity is statistically significant.

Since the meaning of a summary effect size is different for fixed versus random effects, the null hypothesis being tested also differs. Both test a null hypothesis of no linear relationship between the covariates and the effect size. The difference is that under the fixed-effect model that effect size is the common effect size for all studies with a given value of the covariates. Under the random-effects model that effect size is the mean of the true effect sizes for all studies with a given value of the covariates. This is important, because the different null hypotheses reflect different assumptions about the sources of error. This means that different error terms are used to compute tests of significance and confidence intervals under the two models.

Computationally, the difference between fixed effect and random effects is in the definition of the variance. Under the fixed-effect model the variance is

the variance within studies, while under the random-effects model it is the variance within studies plus the variance between studies (τ^2). This holds true for one population, for multiple subgroups, and for regression, but the mechanism used to estimate τ^2 depends on the context. When we are working with a single population, τ^2 reflects the dispersion of true effects across all studies, and is therefore computed for the full set of studies. When we are working with subgroups, τ^2 reflects the dispersion of true effects within a subgroup, and is therefore computed within subgroups. When we are working with regression, τ^2 reflects the dispersion of true effects for studies with the same predicted value (that is, the same value on the covariates) and is therefore computed for each point on the prediction slope. As a practical matter, of course, most points on the slope have only a single study, and so this computation is less transparent than that for the single population (or subgroups) but the concept is the same. The computational details are handled by software and will not be addressed here.

The practical implications of using a random-effects model rather than a fixed-effect model for regression are similar to those that applied to a single population and to subgroups. First, the random-effects model will lead to more moderate weights being assigned to each study. As compared with a fixed-effect model, the random-effects model will assign more weight to small studies and less weight to large studies. Second, the confidence interval about each coefficient (and slope) will be wider than it would be under the fixed-effect model. Third, the *p*-values corresponding to each coefficient and to the model as a whole are less likely to meet the criterion for statistical significance.

As always, the selection of a model must be based on the context and characteristics of the studies. In particular, if there is heterogeneity in true effects that is not explained by the covariates, then the random-effects model is likely to be more appropriate.

RANDOM-EFFECTS MODEL

We return now to the BCG example and apply the random-effects model (Figure 20.5).

Using a random effects analysis the risk ratio for the 13 studies is 0.490 with a confidence interval of 0.345 to 0.695, which says that the *mean effect* of the vaccine was to decrease the risk of TB by at least 30% and possibly by as much as 65% (see Figure 20.5). In log units the risk ratio is −0.714 with a standard error of 0.179, which yields a *Z*-value of -3.995 ($p < 0.001$) which allows us to reject the null hypothesis of no effect.

At least as interesting, however, is the substantial variation in the treatment effects, which ranged from a risk ratio of 0.20 (vaccine *reduces* the risk by 80%) to 1.56 (vaccine *increases* the risk by 56%). The relevant statistics were presented when we discussed the fixed-effect model.

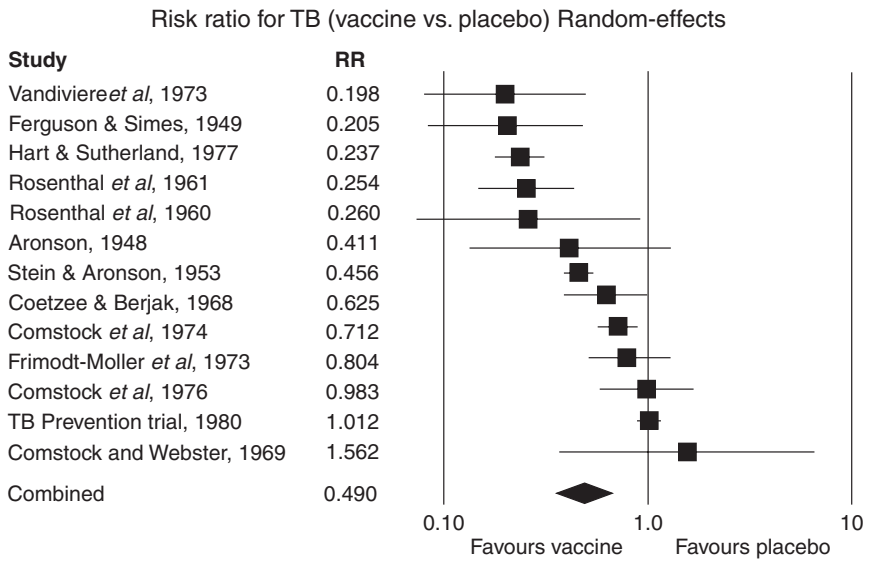

Figure 20.5 Random-effects model – forest plot for the BCG data.

Assessing the impact of the slope

A meta-regression using the random-effect model (method of moments) yields the results shown in Table 20.4.

This has the same format as Table 20.2, showing the coefficients for predicting the log risk ratio from latitude and related statistics. We use an asterisk (*) to indicate that these statistics are based on the random-effects model. With that distinction, the interpretation of the slope(s) is the same as that for the fixed-effect model. Concretely,

$$Z^* = \frac{B^*}{SE_{B^*}}. \tag{20.5}$$

Under the null hypothesis that the coefficient is zero, Z^* would follow the normal distribution.

In the running example the coefficient for latitude is -0.0923 with standard error 0.00673, so

$$Z = \frac{-0.02923}{0.00673} = -4.3411.$$

Table 20.4 Random-effects model – regression results for BCG.

	Random effects, *Z*-Distribution					
	Point estimate	Standard error	95% Lower	95% Upper	*Z*-value	*p*-Value
Intercept	0.25954	0.23231	−0.19577	0.71486	1.11724	0.26389
Latitude	−0.02923	0.00673	−0.04243	−0.01603	−4.34111	0.00001

Table 20.5 Random-effects model – test of the model.

Test of the model:
Simultaneous test that all coefficients (excluding intercept) are zero
$Q^*_{model} = 18.8452$, $df = 1$, $p = 0.00001$
Goodness of fit: Test that unexplained variance is zero
$T^2 = 0.063$, $SE = 0.055$, $Q_{resid} = 30.733$, $df = 11$, $p = 0.00121$

(In this example the slope happens to be almost identical under the fixed-effect and random-effects models, but this is not usually the case.) The two-tailed *p*-value corresponding to $Z^* = -4.3411$ is 0.00001. This tells us that the slope is probably not zero, and the vaccination is more effective when the study is conducted at a greater distance from the equator.

Under the null hypothesis that none of the covariates 1 to *p* is related to effect size, Q^*_{model} would be distributed as chi-squared with degrees of freedom equal to *p*. In the running example, $Q^*_{model} = 18.8452$, $df = 1$, and $p = 0.00001$ (see Table 20.5).

In this example there is only one covariate (latitude) and so we have the option of using either the *Z*-test or the *Q*-test to assess the impact of this covariate. It follows that the two tests should yield the same results, and they do. The *Z*-value is −4.3411, with a corresponding p-value of 0.00001. The *Q*-value is 18.8452 with a corresponding *p*-value of 0.00001. Finally, Q^*_{model} should be equal to Z^{*2} and in fact 18.8452 equals -4.3411^2.

The goodness of fit test addresses the question of whether there is heterogeneity that is not explained by the covariates. Q_{resid} can also be used to estimate (and test) the variance, τ^2, of this unexplained heterogeneity. This Q_{resid} is the weighted residual *SS* from the regression using fixed-effect weights (see Table 20.3)

Q^*_{model} here is analogous to Q^*_{bet} for subgroup analysis, and Q_{resid} is analogous to Q_{within} for subgroup analysis. If the covariates represent subgroups, then Q^*_{model} is identical to Q^*_{bet} and Q_{resid} is identical to Q_{within}. If there are no predictors then Q_{resid} here is the same as *Q* for the original meta-analysis.

When working with meta-regression with the fixed-effect model we were able to partition the total variance into a series of components, with Q_{model} plus Q_{resid} summing to *Q*. This was possible with the fixed-effect model because the weight assigned to each study was determined solely by the within-study error, and was therefore the same for all three sets of calculations. By contrast, under the random-effects model the weight assigned to each study incorporates between-studies variance also, and this varies from one set of calculations to the next. Therefore, the variance components are not additive. For that reason, we display an analysis of variance table for the fixed-effect analysis, but not for the random-effects analysis.

Quantify the magnitude of the relationship

The relationship of latitude to effect (expressed as a log risk ratio) is

$$\ln(RR) = 0.2595 - 0.0292(X)$$

where *X* is the absolute latitude. We can plot this in Figure 20.6.

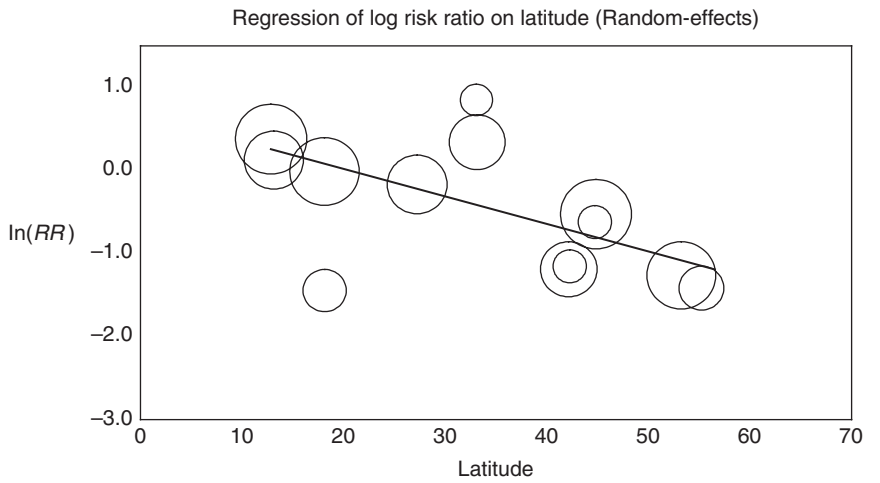

Figure 20.6 Random-effects model – regression of log risk ratio on latitude.

In this Figure, each study is represented by a circle that shows the actual coordinates (observed effect size by latitude) for that study. The size (specifically, the area) of each circle is proportional to that study's weight in the analysis. Since this analysis is based on the random-effects model, the weight is the total variance (within-study plus T^2) for each study.

Note the difference from the fixed-effect graph (Figure 20.2). When using random effects, the weights assigned to each study are more similar to one another. For example, the TB prevention trial (1980) study dominated the graph under the fixed-effect model (and exerted substantial influence on the slope) while Comstock and Webster (1969) had only a trivial impact (the relative weights for the two studies are 41% and 0.3% respectively). Under random effects the two are more similar (14% and 1.6% respectively).

The center line shows the predicted values. A study performed relatively close to the equator (latitude of 10) would have an expected effect near zero (corresponding to a risk ratio of 1.0, which means that the vaccination has no impact on TB). By contrast, a study at latitude 55 (Saskatchewan) would have an expected effect near −1.50 (corresponding to a risk ratio near 0.20, which means that the vaccination decreased the risk of TB by about 80%).

The 95% confidence interval for B is given by

$$LL_{B^*} = B^* - 1.96 \times SE_{B^*} \tag{20.6}$$

and

$$UL_{B^*} = B^* + 1.96 \times SE_{B^*}. \tag{20.7}$$

In the running example

$$LL_{B^*} = (-0.0292) - 1.96 \times 0.0067 = -0.0424$$

and

$$UL_{B^*} = (-0.0292) + 1.96 \times 0.0067 = -0.0160.$$

In words, the true coefficient could be as low as −0.0424 and as high as −0.0160.

The proportion of variance explained

In Chapter 19 we introduced the notion of the proportion of variance explained by subgroup membership in a random-effects analysis. The same approach can be applied to meta-regression.

Consider Figure 20.7, which shows a set of six studies with no covariate. Since there is no covariate the prediction slope is simply the mean (the intercept, if we were to compute a regression), depicted by a vertical line. The normal distribution at the bottom of the figure reflects T^2, and is needed to explain why the dispersion *from the prediction line* (the mean) exceeds the within-study error.

Now, consider Figure 20.8, which shows the same size studies with a covariate X, and the prediction slope depicted by a line that reflects the prediction equation. The normal distribution at each point on the prediction line reflects the value of T^2, and is needed to explain why the dispersion *from the prediction line* (this time, the slope) exceeds the within study error. Because the covariate explains some of the between-studies variance, the T^2 in Figure 20.8 is smaller than the one in Figure 20.8., and the ratio of the two can be used to quantify the proportion of variance explained.

Note 1. Normally, we would plot the effect size on the Y axis and the covariate on the X axis (see, for example, Figure 20.6). Here, we have transposed the axes to maintain the parallel with the forest plot.

Note 2. For clarity, we plotted the true effects for each figure. In practice, of course, we observe estimates of the true effects, remove the portion of variance attributed to within-study error, and impute the amount of variance remaining.

In primary studies, a common approach to describing the impact of a covariate is to report the proportion of variance explained by that covariate. That index, R^2, is defined as the ratio of explained variance to total variance,

$$R^2 = \frac{\sigma^2_{explained}}{\sigma^2_{total}} \tag{20.8}$$

or, equivalently,

$$R^2 = 1 - \left(\frac{\sigma^2_{unexplained}}{\sigma^2_{total}} \right). \tag{20.9}$$

This index is intuitive as it can be interpreted as a ratio, with a range of 0 to 1, or of 0% to 100%. Many researchers are familiar with this index, and have a sense of what proportion of variance is likely to be explained by different kinds of covariates or interventions.

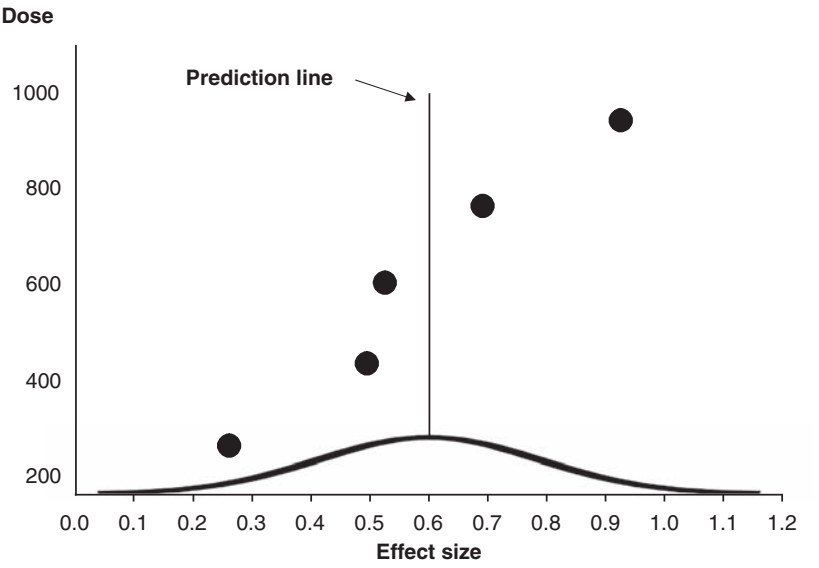

Figure 20.7 Between-studies variance (T^2) with no covariate.

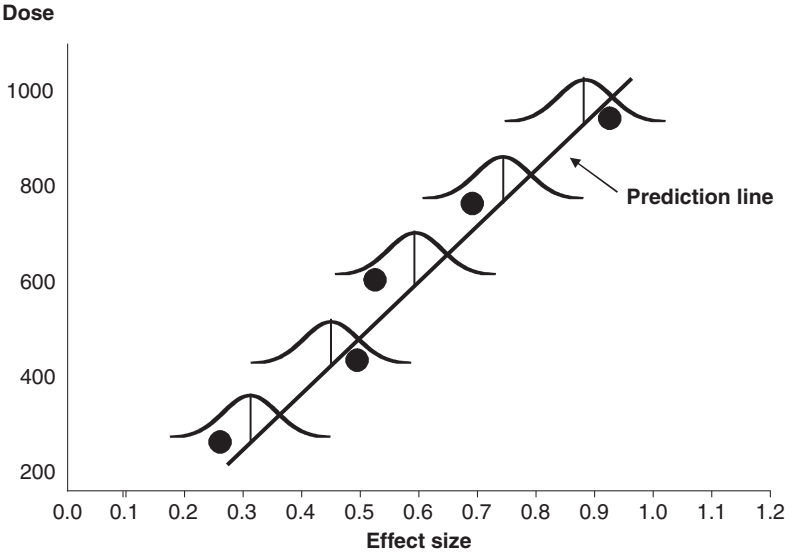

Figure 20.8 Between-studies variance (T^2) with covariate.

This index cannot be applied directly to meta-analysis. The reason is that in meta-analysis the total variance includes both variance within studies and between studies. The covariates are study-level covariates, and as such they can potentially explain only the between-studies portion of the variance. In the running illustration, even if

Table 20.6 Random-effects model – comparison of model (latitude) versus the null model.

Comparison of model with latitude versus the null model
Total between-study variance (intercept only) $T^2_{total} = 0.309$, SE = 0.230, $Q_{resid} = 152.233$, $df = 12$, $p = 0.00000$
Unexplained between-study variance (with latitude in model) $T^2_{unexplained} = 0.063$, SE = 0.055, $Q_{resid} = 30.733$, $df = 11$, $p = 0.0012$
Proportion of total between-study variance explained by the model R^2 analog = 1−(0.063/0.309) 79.50%

the *true* effect for each study fell directly on the prediction line the proportion of variance explained would not approach 1.0 because the *observed* effects would fall at some distance from the prediction line.

Therefore, rather than working with this same index we use an analogous index, defined as the *true* variance explained, as a proportion of the total *true* variance. Since the true variance is the between-studies variance, τ^2, we compute

$$R^2 = \frac{T^2_{explained}}{T^2_{total}} \tag{20.10}$$

or

$$R^2 = 1 - \left(\frac{T^2_{unexplained}}{T^2_{total}} \right). \tag{20.11}$$

In the running example T^2_{total} for the full set of studies was 0.309, and $T^2_{unexplained}$ for the equation with latitude is 0.063. This gives us

$$R^2 = 1 - \left(\frac{0.063}{0.309} \right) = 0.7950. \tag{20.12}$$

This is shown schematically for the running example (see Figure 20.9). In Figure 20.9, the top line shows that 8% of the total variance was within studies and 92%

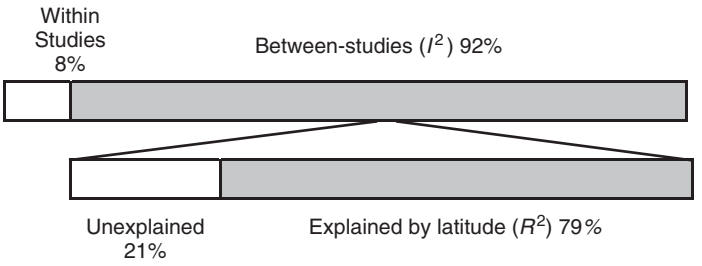

Figure 20.9 Proportion of variance explained by latitude.

was between studies (which is also the definition of I^2). The within-studies variance cannot be explained by a study-level moderator, and so is removed from the equation and we focus on the shaded part.

On the bottom line, the type of intervention is able to explain 79% of the relevant variance, leaving 21% unexplained. Critically, the 79% and 21% sum to 100%, since we are concerned only with the variance between-studies.

While the R^2 index has a range of 0 to 1 in the population, it is possible for sampling error to yield a value of R^2 that falls outside of this range. In that case, the value is set to either 0 (if the estimate falls below 0) or to 1 (if it falls above 1).

SUMMARY POINTS

- Just as we can use multiple regression in primary studies to assess the relationship between subject-level covariates and an outcome, we can use meta-regression in meta-analysis to assess the relationship between study-level covariates and effect size.
- Meta-regression may be performed under the fixed-effect or the random-effects model, but in most cases the latter is appropriate.
- In addition to testing the impact of covariates for statistical significance, it is important to quantify the magnitude of their relationship with effect size. For this purpose we can use an index based on the percent reduction in true variance, analogous to the R^2 index used with primary studies.