A Handbook of Statistical Analyses Using **R**

Brian S. Everitt and Torsten Hothorn

Meta-Analysis: Nicotine Gum and Smoking Cessation and the Efficacy of BCG Vaccine in the Treatment of Tuberculosis

12.1 Introduction

12.2 Systematic Reviews and Meta-Analysis

12.3 Analysis Using **R**

The aim in collecting the results from the randomised trials of using nicotine gum to help smokers quit was to estimate the overall *odds ratio*, the odds of quitting smoking for those given the gum, divided by the odds of quitting for those not receiving the gum. The odds ratios and corresponding confidence intervals are computed by means of the `meta.MH` function for fixed effects meta-analysis as shown here

```
R> library("rmeta")
R> data("smoking", package = "HSAUR")
R> smokingOR <- meta.MH(smoking[["tt"]], smoking[["tc"]],
+                       smoking[["qt"]], smoking[["qc"]],
+                       names = rownames(smoking))
```

and the results can be inspected via a `summary` method – see Figure˜12.1.

We shall use both the fixed effects and random effects approaches here so that we can compare results. For the fixed effects model (see Figure˜12.1) the estimated overall log-odds ratio is 0.513 with a standard error of 0.066. This leads to an estimate of the overall odds ratio of 1.67, with a 95% confidence interval as given above. For the random effects model

```
R> smokingDSL <- meta.DSL(smoking[["tt"]], smoking[["tc"]],
+                         smoking[["qt"]], smoking[["qc"]],
+                         names = rownames(smoking))
R> print(smokingDSL)
```

```
Random effects ( DerSimonian-Laird ) meta-analysis
Call: meta.DSL(ntrt = smoking[["tt"]], nctrl = smoking[["tc"]], ptrt = smoking[["qt"
    pctrl = smoking[["qc"]], names = rownames(smoking))
Summary OR= 1.75    95% CI ( 1.48, 2.07 )
Estimated random effects variance: 0.05
```

```
R> summary(smokingOR)
```

```
Fixed effects ( Mantel-Haenszel ) meta-analysis
Call: meta.MH(ntrt = smoking[["tt"]], nctrl = smoking[["tc"]], ptrt = smoking[["qt"]]
    pctrl = smoking[["qc"]], names = rownames(smoking))
------------------------------------
               OR (lower  95% upper)
Blondal89      1.85   0.99      3.46
Campbell91     0.98   0.50      1.92
Fagerstrom82   1.76   0.80      3.89
Fee82          1.53   0.77      3.05
Garcia89       2.95   1.01      8.62
Garvey00       2.49   1.43      4.34
Gross95        2.62   1.03      6.71
Hall85         2.03   0.78      5.29
Hall87         2.82   1.33      5.99
Hall96         0.87   0.46      1.64
Hjalmarson84   2.17   1.10      4.28
Huber88        6.00   2.57     14.01
Jarvis82       3.33   1.37      8.08
Jensen91       1.43   0.84      2.44
Killen84       1.33   0.43      4.15
Killen90       1.23   0.93      1.64
Malcolm80      3.52   0.85     14.54
McGovern92     1.17   0.70      1.94
Nakamura90     3.82   1.15     12.71
Niaura94       1.34   0.35      5.19
Pirie92        1.84   1.20      2.82
Puska79        1.46   0.78      2.75
Schneider85    1.71   0.52      5.62
Tonnesen88     2.12   0.93      4.86
Villa99        1.76   0.55      5.64
Zelman92       1.46   0.68      3.14
------------------------------------
Mantel-Haenszel OR =1.67 95% CI ( 1.47,1.9 )
Test for heterogeneity: X^2( 25 ) = 34.9 ( p-value 0.09 )
```

Figure 12.1 R output of the summary method for smokingOR.

the corresponding estimate is 1.751. Both models suggest that there is clear evidence that nicotine gum increases the odds of quitting. The random effects confidence interval is considerably wider than that from the fixed effects model; here the test of homogeneity of the studies is not significant implying that we might use the fixed effects results. But the test is not particularly powerful and it is more sensible to assume a priori that heterogeneity is present and so we use the results from the random effects model.

```
R> plot(smokingOR, ylab = "")
```
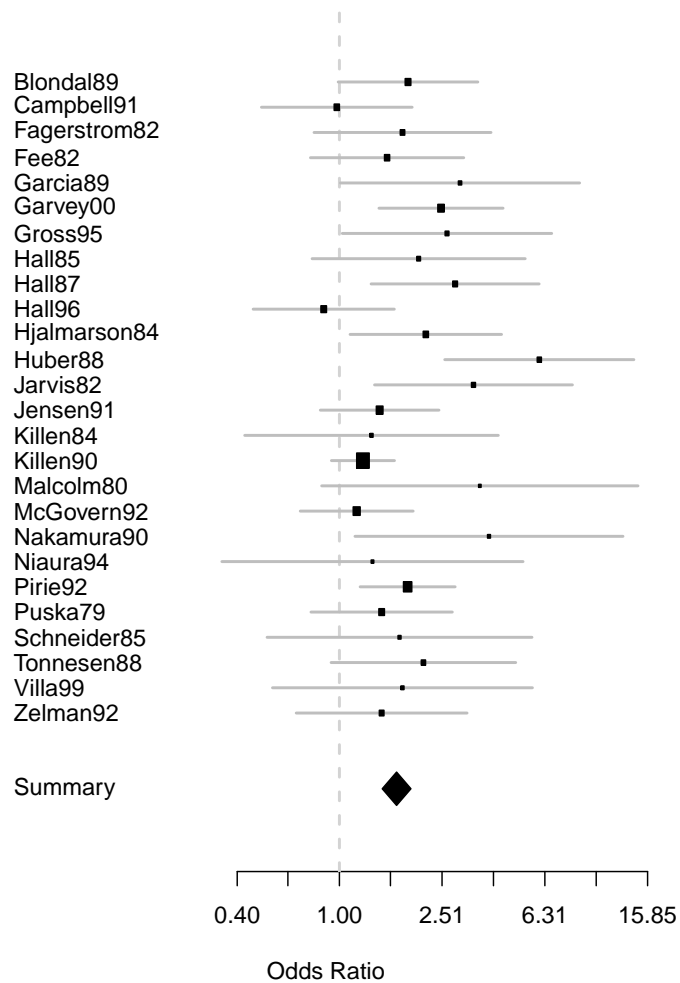

Figure 12.2 Forest plot of observed effect sizes and 95% confidence intervals for the nicotine gum studies.

12.4 Meta-Regression

The examination of heterogeneity of the effect sizes from the studies in a meta-analysis begins with the formal test for its presence, although in most meta-analyses such heterogeneity can almost be assumed to be present. There will be many possible sources of such heterogeneity and estimating how these various factors affect the observed effect sizes in the studies chosen is often of considerable interest and importance, indeed usually more important than the relatively simplistic use of meta-analysis to determine a single summary estimate of overall effect size. We can illustrate the process using the BCG vaccine data. We first find the estimate of the overall effect size from applying the fixed effects and the random effects models described previously:

```
R> data("BCG", package = "HSAUR")
R> BCG_OR <- meta.MH(BCG[["BCGVacc"]], BCG[["NoVacc"]],
+                    BCG[["BCGTB"]], BCG[["NoVaccTB"]],
+                    names = BCG$Study)
R> BCG_DSL <- meta.DSL(BCG[["BCGVacc"]], BCG[["NoVacc"]],
+                    BCG[["BCGTB"]], BCG[["NoVaccTB"]],
+                    names = BCG$Study)
```

The results are inspected using the `summary` method as shown in Figures˜12.3 and 12.4.

```
R> summary(BCG_OR)

Fixed effects ( Mantel-Haenszel ) meta-analysis
Call: meta.MH(ntrt = BCG[["BCGVacc"]], nctrl = BCG[["NoVacc"]], ptrt = BCG[["BCGTB"]]
    pctrl = BCG[["NoVaccTB"]], names = BCG$Study)
------------------------------------
    OR (lower  95% upper)
1  0.39    0.12      1.26
2  0.19    0.08      0.46
3  0.25    0.07      0.91
4  0.23    0.18      0.31
5  0.80    0.51      1.26
6  0.38    0.32      0.47
7  0.20    0.08      0.50
8  1.01    0.89      1.15
9  0.62    0.39      1.00
10 0.25    0.14      0.42
11 0.71    0.57      0.89
12 1.56    0.37      6.55
13 0.98    0.58      1.66
------------------------------------
Mantel-Haenszel OR =0.62 95% CI ( 0.57,0.68 )
Test for heterogeneity: X^2( 12 ) = 163.94 ( p-value 0 )
```

Figure 12.3   R output of the `summary` method for `BCG_OR`.

```
R> summary(BCG_DSL)
```

```
Random effects ( DerSimonian-Laird ) meta-analysis
Call: meta.DSL(ntrt = BCG[["BCGVacc"]], nctrl = BCG[["NoVacc"]], ptrt = BCG[["BCGTB"
    pctrl = BCG[["NoVaccTB"]], names = BCG$Study)
------------------------------------
     OR (lower  95% upper)
1  0.39    0.12      1.26
2  0.19    0.08      0.46
3  0.25    0.07      0.91
4  0.23    0.18      0.31
5  0.80    0.51      1.26
6  0.38    0.32      0.47
7  0.20    0.08      0.50
8  1.01    0.89      1.15
9  0.62    0.39      1.00
10 0.25    0.14      0.42
11 0.71    0.57      0.89
12 1.56    0.37      6.55
13 0.98    0.58      1.66
------------------------------------
SummaryOR= 0.47  95% CI ( 0.32,0.69 )
Test for heterogeneity: X^2( 12 ) = 163.16 ( p-value 0 )
Estimated random effects variance: 0.37
```

Figure 12.4 R output of the `summary` method for `BCG_DSL`.

To assess how the two covariates, latitude and year, relate to the observed effect sizes we shall use multiple linear regression but will weight each observation by $W_i = (\hat{\sigma}^2 + V_i^2)^{-1}, i = 1, \ldots, 13$, where $\hat{\sigma}^2$ is the estimated between-study variance and V_i^2 is the estimated variance from the ith study. The required R code to fit the linear model via weighted least squares is:

```
R> studyweights <- 1 / (BCG_DSL$tau2 + BCG_DSL$selogs^2)
R> y <- BCG_DSL$logs
R> BCG_mod <- lm(y ~ Latitude + Year, data = BCG,
+                weights = studyweights)
```

and the results of the `summary` method are shown in Figure˜12.5. There is some evidence that latitude is associated with observed effect size, the log-odds ratio becoming increasingly negative as latitude increases, as we can see from a scatterplot of the two variables with the added weighted regression fit seen in Figure˜12.6.

12.5 Publication Bias

We can construct a funnel plot for the nicotine gum data using the R code depicted with Figure˜12.8. There does not appear to be any strong evidence of publication bias here.

```
R> summary(BCG_mod)
```

```
Call:
lm(formula = y ~ Latitude + Year, data = BCG, weights = studyweights)

Residuals:
    Min      1Q   Median      3Q     Max
-1.66012 -0.36910 -0.02937  0.31565  1.26040

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -16.199115  37.605403  -0.431   0.6758
Latitude     -0.025808   0.013680  -1.887   0.0886 .
Year          0.008279   0.018972   0.436   0.6718
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7992 on 10 degrees of freedom
Multiple R-squared: 0.4387,        Adjusted R-squared: 0.3265
F-statistic: 3.909 on 2 and 10 DF,  p-value: 0.05569
```

Figure 12.5   R output of the `summary` method for `BCG_mod`.

```
R> plot(y ~ Latitude, data = BCG, ylab = "Estimated log-OR")
R> abline(lm(y ~ Latitude, data = BCG, weights = studyweights))
```
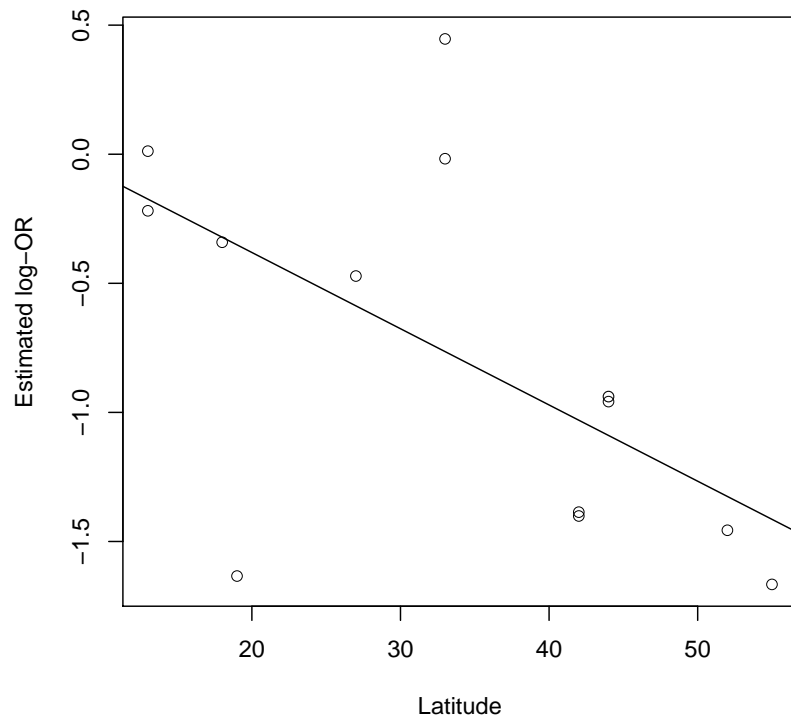

Figure 12.6   Plot of observed effect size for the `BCG` vaccine data against latitude, with a weighted least squares regression fit shown in addition.
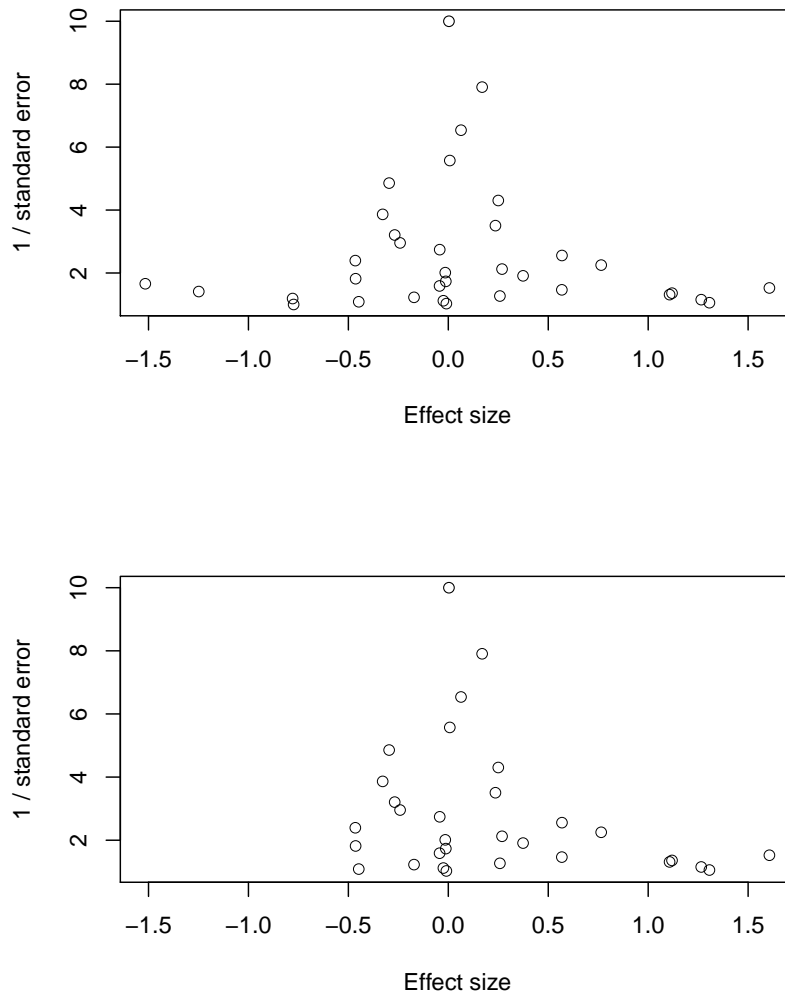
Figure 12.7 Example funnel plots from simulated data. The asymmetry in the lower plot is a hint that a publication bias might be a problem.

```
R> funnelplot(smokingDSL$logs, smokingDSL$selogs,
+              summ = smokingDSL$logDSL, xlim = c(-1.7, 1.7))
R> abline(v = 0, lty = 2)
```
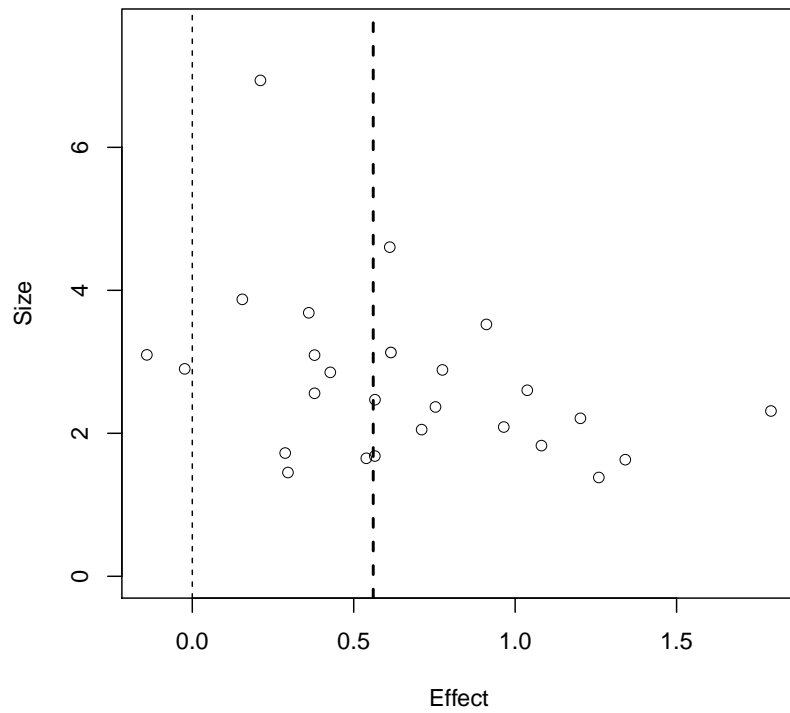

Figure 12.8   Funnel plot for nicotine gum data.