Using R for Epidemiological Analysis

Practical 2: Meta Analysis in R

Introduction

In this practical we will use `R` to perform meta-analyses. We will consider two examples: (i) a binary outcome and (ii) a continuous outcome. For the binary outcome example we will use the `cisapride` dataset that is included within the `meta` package. This examines the risk associated with taking the Cisapride drug versus a placebo in non-ulcer dispepsia patients. For a continuous outcome we will use the `Fleiss93cont` dataset which contains information on the effect of mental health treatment on medical utilisation.

Preliminaries

We need the following package

- `meta` - Package that provides standard methods for meta- analysis.

As in Practical 1, we use the `install.packages()` function to download and install the packages that we need and the `require()` function to load them into the `R` library.

```
install.packages(meta)
```

```
require(meta)
```

Meta-analysis of a Binary Outcome

In this example, we will use the `cisapride` dataset. We will see whether there is a reduced risk associated with taking the Cisapride drug versus a placebo in non-ulcer dispepsia patients. We first need to load in the data, and since this is included within `meta` package we can use the `data()` function.

```
data(cisapride)
```

This function has called the dataset to the `R` workspace. We can look at, and summarise, the data as follows.

```
str(cisapride)
'data.frame':   13 obs. of  5 variables:
 $ study     : Factor w/ 13 levels "Al-Quorain et al. [23]",..: 3 10 6 5 7 11 4 8 2 12 ...
 $ event.cisa: int  15 12 29 42 14 44 14 29 10 17 ...
 $ n.cisa    : int  16 16 34 56 22 54 17 58 14 26 ...
 $ event.plac: int  9 1 18 31 6 17 7 23 3 6 ...
 $ n.plac    : int  16 16 34 56 22 55 15 58 15 27 ...

summary(cisapride)
                      study     event.cisa        n.cisa
 Al-Quorain et al. [23]    :1   Min.   :10.00   Min.   :14.00
 Chung [21]                :1   1st Qu.:14.00   1st Qu.:17.00
```

```
 Creytens [13]              :1   Median :19.00   Median :29.00
 De Nutte et al. [19]       :1   Mean   :23.38   Mean   :32.62
 Deruyttere et al. [16]     :1   3rd Qu.:29.00   3rd Qu.:44.00
 Francois and De Nutte [15]:1   Max.   :44.00   Max.   :58.00
 (Other)                    :7
   event.plac        n.plac
 Min.   : 1.00   Min.   :15.00
 1st Qu.: 6.00   1st Qu.:16.00
 Median :12.00   Median :30.00
 Mean   :13.38   Mean   :32.85
 3rd Qu.:19.00   3rd Qu.:45.00
 Max.   :31.00   Max.   :58.00
```

We can see that this dataset contains the following five variables:

- `study` - Study label
- `event.cisa` - Number of events in cisapride group
- `n.cisa` - Number of observations in cisapride group
- `event.plac` - Number of events in placebo group
- `n.plac` - Number of observations in placebo group

To perform a meta-analysis for a binary outcome you will always need a study identifier and the number of events and observations for the experimental and control groups considered in each study. More information on this dataset can be found by typing `?cisapride` into `R`.

To perform a meta-analysis calculation using relative risks we use the `metabin()` function and assign the output to the variable `mod`:

```
mod <- metabin(event.e  = event.cisa, # Number of events in experimental group
               n.e      = n.cisa,     # Number of observations in experimental group
               event.c  = event.plac, # Number of events in control group
               n.c      = n.plac,     # Number of observations in control group
               data     = cisapride,  # Name of the dataset
               studylab = study,      # Study labels
               sm       = "RR")       # Summary Measure
```

All the information from our meta-analysis is contained in the object `mod`. This can be summarised using the `summary()` function.

```
summary(mod)
Number of studies combined: k = 13

                         RR           95%-CI    z  p-value
Fixed effect model   1.7570 [1.5439; 1.9995] 8.54 < 0.0001
Random effects model 1.7806 [1.3753; 2.3053] 4.38 < 0.0001

Quantifying heterogeneity:
tau^2 = 0.1439; H = 1.90 [1.43; 2.51]; I^2 = 72.2% [51.3%; 84.1%]

Test of heterogeneity:
     Q d.f.  p-value
 43.09   12 < 0.0001
```

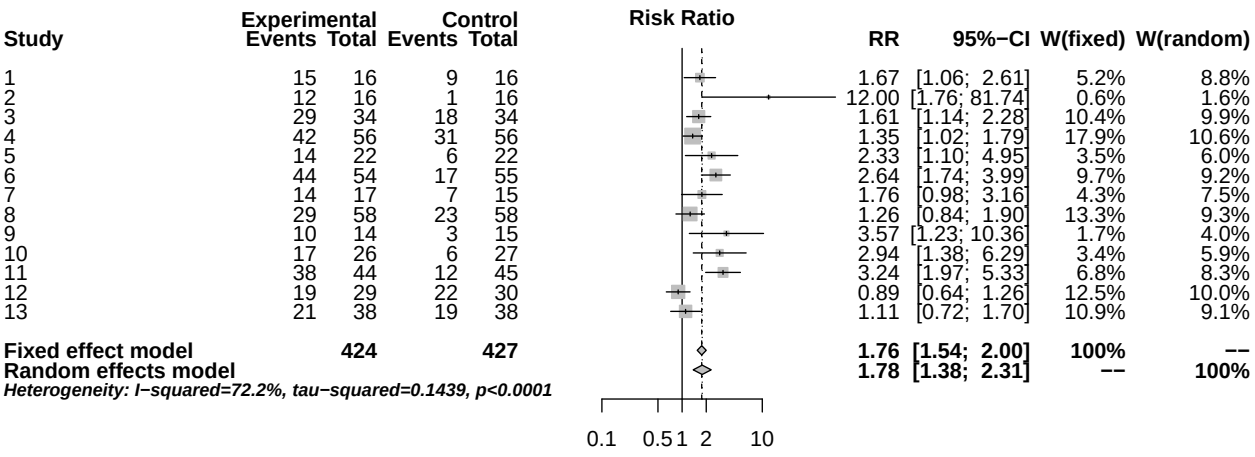
```
Details on meta-analytical method:
- Mantel-Haenszel method
- DerSimonian-Laird estimator for tau^2
```

Questions

1. Is there a difference between the pooled results from the fixed and random effects?
2. Are the results significant?
3. By looking at the I^2 statistic, can you describe the level of heterogeneity?
4. By looking at Cochran's Q statistic, is the heterogeneity significant?

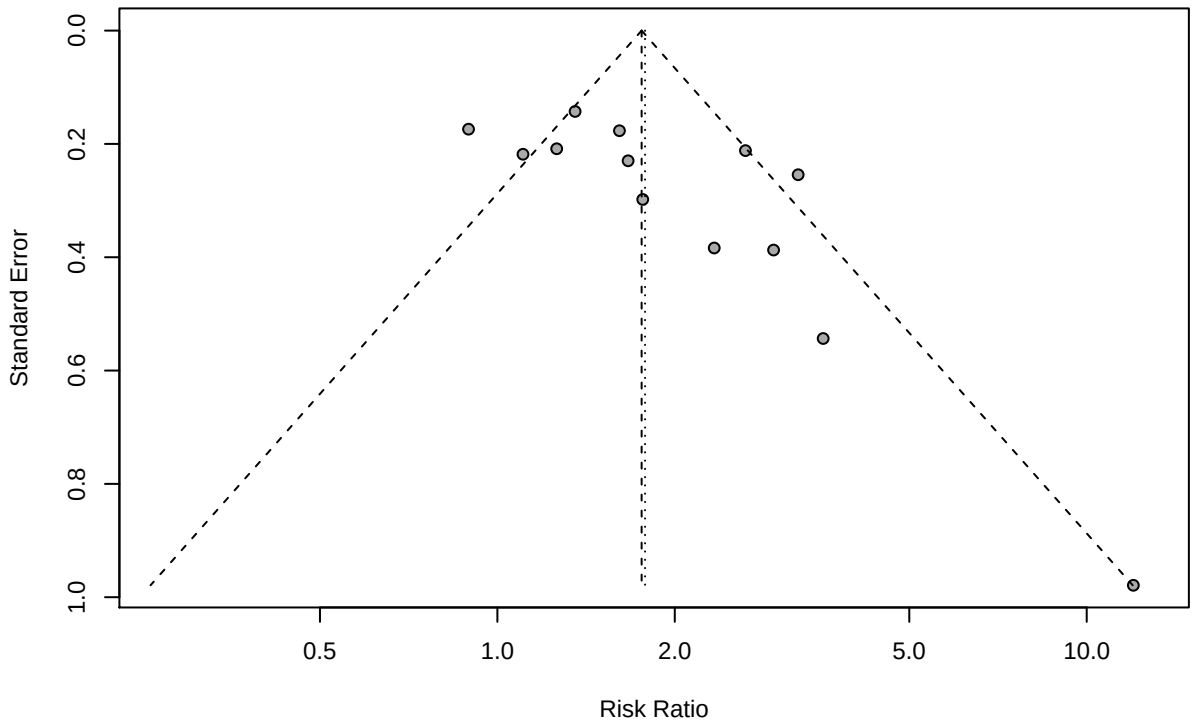
In addition, we can create a visual summary of the meta-analysis. We can create a forest plot using object `mod` created above by using it as an input to `forest()` function.

```
forest(mod)
```



We can create a funnel plot by using the `funnel()` function in a similar way.

```
funnel(mod)
```



Questions

1. Do any results look strange?
2. Is there evidence of publication bias?

If you have time, you may want to repeat this analysis for another dataset to understand how to implement a binary meta-analysis in `R`. You should use the `Fleiss93` dataset which can be called to your `R` workspace using the command `data(Fleiss93)`. The meta-analysis on this data will examine the possible effects of aspirin in preventing death after myocardial infarction. Specific information about this dataset can be seen by using the command `?Fleiss93`. Study the data, and carefully adapt the above code to implement the meta-analysis.

Meta-analysis of a Continuous Outcome

In this example, we will use the `Fleiss93cont` dataset. We will examine the effect of of mental health treatment on medical utilisation. We first need to load in the data, which is also included within `meta` package, therefore we can use the `data()` function.

```
data(Fleiss93cont)
```

This function has called the dataset to the `R` workspace. We can look at, and summarise, the data as follows.

```
str(Fleiss93cont)
'data.frame':   5 obs. of  8 variables:
 $ study : Factor w/ 5 levels "Davis","Florell",..: 1 2 3 4 5
 $ year  : int  1973 1971 1975 1975 1977
 $ n.e   : int  13 30 35 20 8
 $ mean.e: num  5 4.9 22.5 12.5 6.5
 $ sd.e  : num  4.7 1.71 3.44 1.47 0.76
```

```
 $ n.c   : int  13 50 35 20 8
 $ mean.c: num  6.5 6.1 24.9 12.3 7.38
 $ sd.c  : num  3.8 2.3 10.65 1.66 1.41

summary(Fleiss93cont)
     study        year           n.e           mean.e           sd.e
 Davis  :1   Min.   :1971   Min.   : 8.0   Min.   : 4.90   Min.   :0.760
 Florell:1   1st Qu.:1973   1st Qu.:13.0   1st Qu.: 5.00   1st Qu.:1.470
 Gruen  :1   Median :1975   Median :20.0   Median : 6.50   Median :1.710
 Hart   :1   Mean   :1974   Mean   :21.2   Mean   :10.28   Mean   :2.416
 Wilson :1   3rd Qu.:1975   3rd Qu.:30.0   3rd Qu.:12.50   3rd Qu.:3.440
             Max.   :1977   Max.   :35.0   Max.   :22.50   Max.   :4.700
      n.c           mean.c           sd.c
 Min.   : 8.0   Min.   : 6.10   Min.   : 1.410
 1st Qu.:13.0   1st Qu.: 6.50   1st Qu.: 1.660
 Median :20.0   Median : 7.38   Median : 2.300
 Mean   :25.2   Mean   :11.44   Mean   : 3.964
 3rd Qu.:35.0   3rd Qu.:12.30   3rd Qu.: 3.800
 Max.   :50.0   Max.   :24.90   Max.   :10.650
```

We can see that this dataset has the following variables:

- `study` - Study label
- `year` - Year of publication
- `n.e` - Number of observations in psychotherapy group
- `mean.e` - Estimated mean in psychotherapy group
- `sd.e` - Standard deviation in psychotherapy group
- `n.c` - Number of observations in control group
- `mean.c` - Estimated mean in control group
- `sd.c` - Standard deviation in control group

More information on this dataset can be found by typing `?Fleiss93cont` into `R`. To perform a meta-analysis in the continous case you need a study identifier, in addition to the number of events, mean effect and standard deviation of the effect for both the experimental and control groups.

To perform a meta-analysis calculation using relative risks in the continuous case we use the `metacont()` function and assign the output to the variable `mod`:

```
mod <- metacont(n.e     = n.e,              # Number of observations in experimental group
                mean.e  = mean.e,           # Estimated mean effect in experimental group
                sd.e    = sd.e,             # Standard deviation of effect in experimental group
                n.c     = n.c,              # Number of observations in control group
                mean.c  = mean.c,           # Estimated mean effect in control group
                sd.c    = sd.c,             # Standard deviation of effect in control group
                data    = Fleiss93cont,     # Name of the dataset
                studlab = paste(study,year), # Study labels
                sm      = "ROM")            # Summary Measure
```

All the information from our meta-analysis is contained in the object `mod`, and can be summarised by using the `summary()` function.

```
summary(mod)
Number of studies combined: k = 5
```

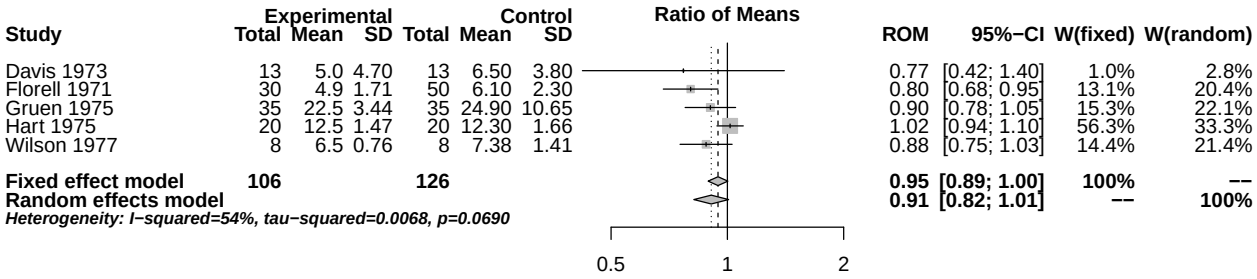
```
                         ROM           95%-CI     z  p-value
Fixed effect model   0.9457 [0.8916; 1.0030] -1.86   0.0631
Random effects model 0.9082 [0.8187; 1.0076] -1.82   0.0692

Quantifying heterogeneity:
tau^2 = 0.0068; H = 1.48 [1.00; 2.43]; I^2 = 54.0% [0.0%; 83.1%]

Test of heterogeneity:
    Q d.f.  p-value
 8.70    4   0.0690

Details on meta-analytical method:
- Inverse variance method
- DerSimonian-Laird estimator for tau^2
```

Questions

1. Is there a difference between the pooled results for the fixed and random effects?
2. Are the results significant?
3. By looking at the I^2 statistic, can you describe the level of heterogeneity?
4. By looking at Cochran's Q statistic, is the heterogeneity significant?

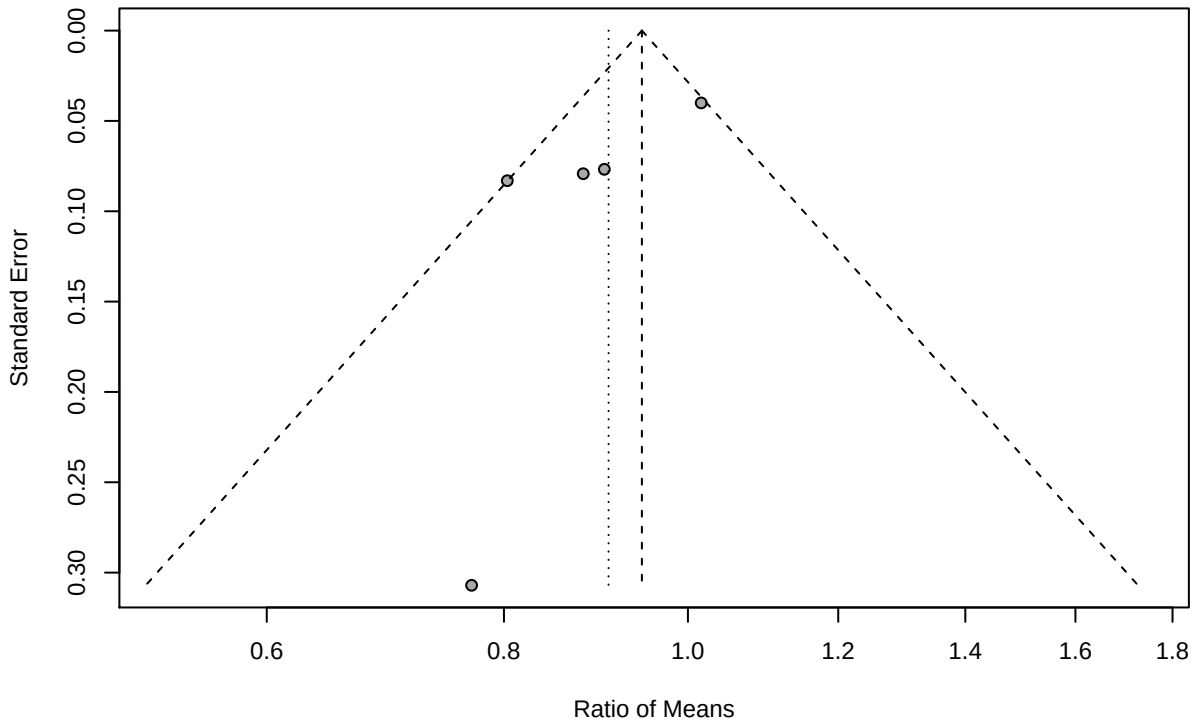
In addition, we can create a visual summary of the meta-analysis. We can create a forest plot using object `mod` created above and using it as input to the `forest()` function.

```
forest(mod)
```



We can also create a funnel plot by using the `funnel()` function in a similar way.

```
funnel(mod)
```



Questions

1. Do any results look strange?
2. Is there evidence of publication bias?

If you have time, you may want to repeat this analysis for another dataset to understand how to implement a continuous meta-analysis in `R`. You should use the Woody Plants dataset which can be called to your `R` workspace using the command `data(woodyplants)`. The meta-analysis on this data will try to understand the effect of elevated CO_2 on total biomass of woody plants. Specific information about this dataset can be seen by using the command `?woodyplants`. Study the data, and carefully adapt the above code to implement the meta-analysis.