Meta-analysis (day 1)

Advanced modelling with R

Juan R Gonzalez
juanr.gonzalez@isglobal.org

BRGE - Bioinformatics Research Group in Epidemiology
ISGlobal - Barcelona Institute for Global Health
http://brge.isglobal.org

Preliminaries

```
install.packages(c("rmeta", "meta", "altmeta",
                   "RISmed", "RCurl"))
```

I assume you are familiar with:

- R
- RStudio
- RMarkdown

Outline

- Overview
- Types of effects
- Searching The Literature
- Prepare data
- Pooling Effect Sizes
  - Fixed Effects Model
  - Random Effects Model
  - Forest plot
- Evaluating Heterogeneity
  - Funnel plot
  - Sensitivity Analysis
  - Subgroup Analysis
  - Outlier's detection
- Meta-Regression (Mixed Effects Model)
- Publication Bias
  - Trim-and-fill
  - Fail-Safe N
  - P-curve

What a meta-analysis is?

Quantitative approach for **systematically combining** results of **previous research** to achieve **conclusions** about our scientifc question.

- quantitative: numbers
- systematic : methodological
- combining: putting together
- previous research: evidence based
- conclusions: new knowledge

What a meta-analysis is?

'The analysis of analyses'

Gene V. Glass. Primary, secondary and meta-analysis of research. Educational Researcher, 1976.

What a meta-analysis is?

- More formally... a meta-analysis is the synthesis of:
  - compatible effects (Y_i)
  - (Preferably, but not necessarily, from randomized controlled trials)
- Giving greater weight to studies with:
  - Less variance (V_i), and
  - More precision ($W_i = 1/V_i$)

Types of effects

An **effect** could be almost any aggregate statistic of interest:

- Mean, Mean difference, Mean change
- Risk ratio, Odds ratio, Risk difference
- Incidence rate, Prevalence, Proportion
- Correlation, slope

Estimation of effect size (ES)

- An effect size could be almost any summary statistic (e.g. a mean, a difference in proportions, an adjusted odds ratio, etc.)
- Conventional meta-analytic models assume normality of ESs. Because of the CLT, this will holds for most ESs given large enough samples.
- To normalize ESs, a log-transform is common.

Estimation of effect size (ES)

	Event	Non-Event	Sample Size
Group A	a_i	b_i	n_{iA}
Group B	c_i	d_i	n_{iB}

Figure 1: Contingency table

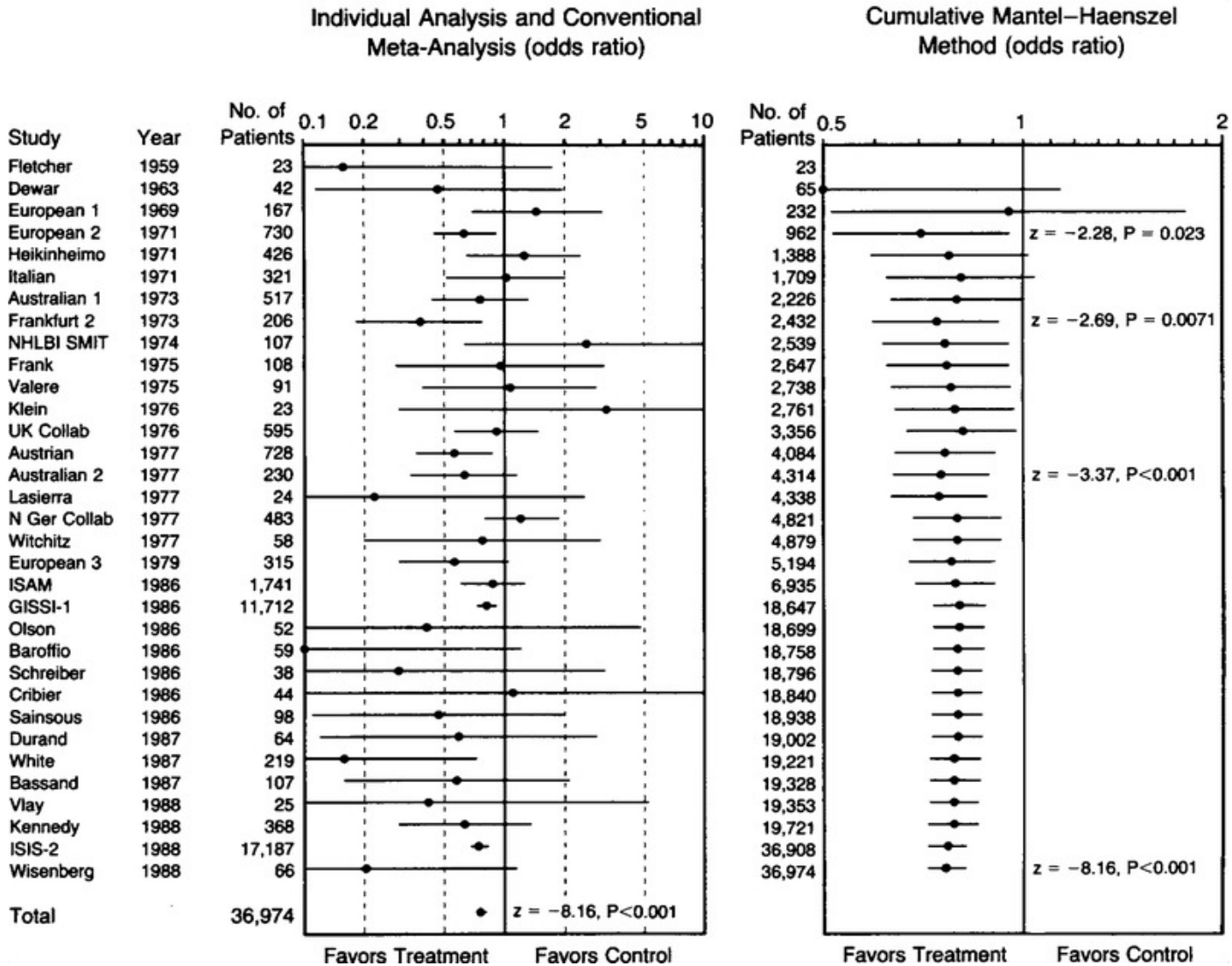
- Effect size estimate (log odds ratio):

$$LOR = \log(\frac{a \times d}{c \times b})$$

- Variance

$$V = \frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}$$

Motivating example



Source: Lau J, et al. N Engl J Med 1992; 327:248-254

We are not yet ready to analyze data

- A meta-analysis starts with a **systematic review**.
- A systematic review is a **scientific summary** of all available evidence on a specific research question.
- An exhaustive search of the literature will require more than R.
- **Note**: If available studies are too few or too different a meta-analysis may **not be appropriate**.

Searching the literature with R

- Not many packages for helping with early stages of a systematic review.
- `RISmed` package can import metadata from NCBI databases into R.
- Using this package, one can search, store, and easily mine metadata on PubMed articles.
- `RISmed` tools are not comprehensive enough to complete a systematic review but may be a helpful aid.

Searching the literature with R (`RISmed`)

- Create `EUtilsSummary` object for specified query.
- Retrieve matching records with `EUtilsGet`.

Sintax:

- EUtilsSummary([query], [db], [search.limits])
- query: String query as given on PubMed site
- db: String name of NCBI database
- search.limits: Additional arguments to restrict search

Searching the literature with R (RISmed)

```
library("RISmed")
```

The following code performs a PubMed query of all BMJ articles with "rofecoxib" in the title [NSAID that was withdrawn from the market for safety reasons.]

```
rofecoxib <- EUtilsSummary("rofecoxib[ti] +
                           British Medical Journal[jo]",
                           db = "pubmed")
```

You can obtaine the exact serarch (to be added to the paper)

```
QueryTranslation(rofecoxib)
```

```
[1] "rofecoxib[ti] AND (\"Br Med J\"[Journal] OR \"Br Med J (Clin Res Ed)\"[Journal] OR \"BMJ\"[J
```

and the number of papers found

```
QueryCount(rofecoxib)
```

```
[1] 16
```

Searching the literature with R (RISmed)

Now we can extract the metadata for the queried records.

```
metadata <- EUtilsGet(rofecoxib)
metadata # Medline Object
```

```
PubMed query:  rofecoxib[ti] AND ("Br Med J"[Journal] OR "Br Med J (Clin Res Ed)"[Journal] OR "BM
```

```
Records:  16
```

Functions for **Medline** objects

```
ls("package:RISmed")
```

```
 [1] "AbstractText"          "Acronym"              "Affiliation"
 [4] "Agency"                "ArticleId"            "ArticleTitle"
 [7] "Author"                "Cited"                "CollectiveName"
[10] "CopyrightInformation" "Country"              "DayAccepted"
[13] "DayEpublish"           "DayPmc"               "DayPpublish"
[16] "DayPubmed"             "DayReceived"          "ELocationID"
[19] "EUtilsGet"             "EUtilsQuery"          "EUtilsSummary"
[22] "GrantID"               "HourAccepted"         "HourEpublish"
[25] "HourPmc"               "HourPpublish"         "HourPubmed"
[28] "HourReceived"          "ISOAbbreviation"      "ISSN"
[31] "ISSNLinking"           "Issue"                "Language"
[34] "Medline"               "MedlinePgn"           "MedlineTA"
[37] "Mesh"                  "MinuteAccepted"       "MinuteEpublish"
[40] "MinutePmc"             "MinutePpublish"       "MinutePubmed"
[43] "MinuteReceived"        "MonthAccepted"        "MonthEpublish"
[46] "MonthPmc"              "MonthPpublish"        "MonthPubmed"
[49] "MonthReceived"         "NlmUniqueID"          "PMID"
[52] "print"                 "PublicationStatus"    "PublicationType"
[55] "Query"                 "QueryCount"           "QueryId"
[58] "QueryTranslation"      "RefSource"            "RegistryNumber"
[61] "show"                  "summary"              "Title"
[64] "Volume"                "YearAccepted"         "YearEpublish"
[67] "YearPmc"               "YearPpublish"         "YearPubmed"
[70] "YearReceived"
```

Methods for **Medline** objects

```
ArticleTitle(metadata)[1:5]
```

```
[1] "Merck pays $1bn penalty in relation to promotion of rofecoxib."
[2] "Merck to pay $58m in settlement over rofecoxib advertising."
[3] "94% of patients suing Merck over rofecoxib agree to company's offer."
[4] "Merck to pay $5bn in rofecoxib claims."
[5] "Merck appeals rofecoxib verdict."
```

Methods for **Medline** objects

```
Author(metadata)[[1:2]]
```

```
[1] "Janice Hopkins"
```

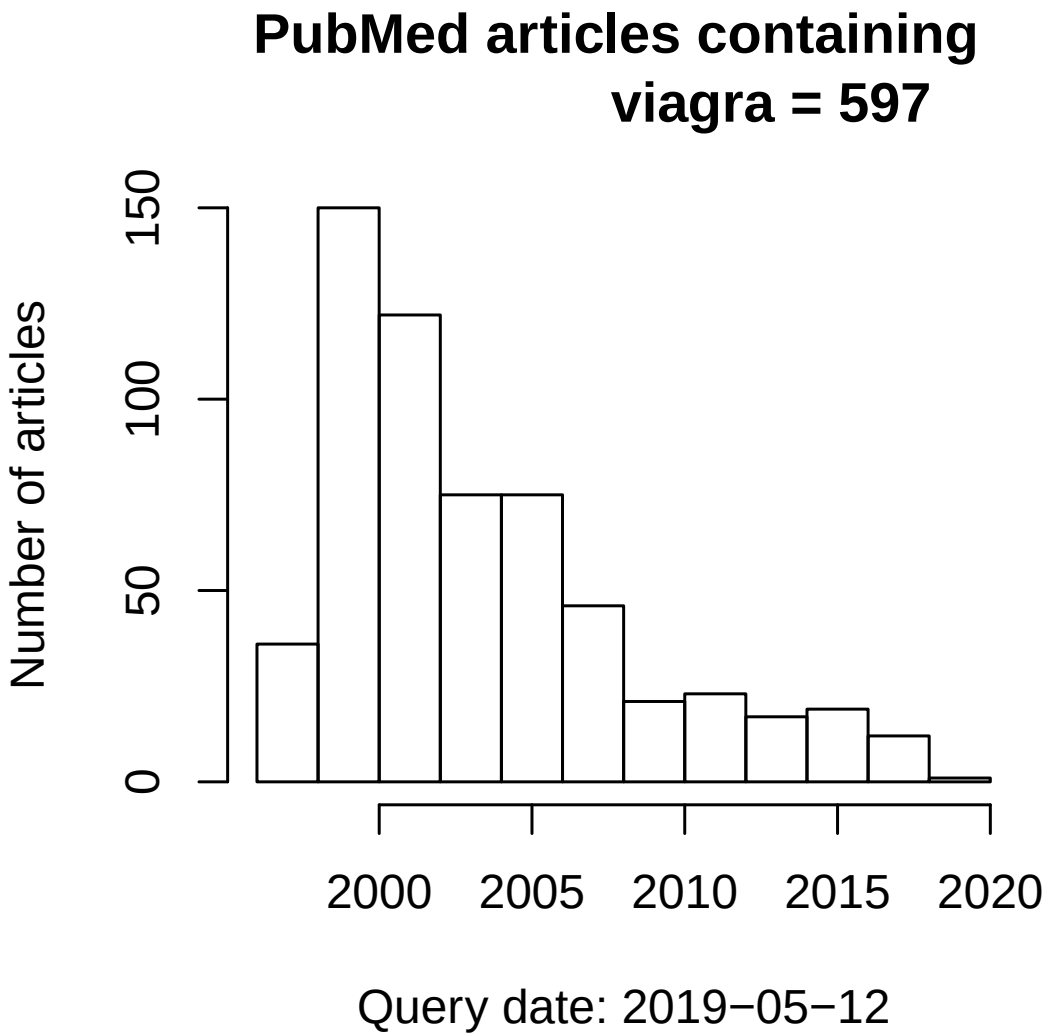
```
YearPubmed(metadata)
```

```
 [1] 2011 2008 2008 2007 2007 2006 2005 2005 2004 2004 2004 2004 2004 2003 2002
[16] 2001
```

Methods for **Medline** objects

```
pm <- EUtilsSummary("viagra[ti]",
                    db = "pubmed")
metadata.pm <- EUtilsGet(pm)
y <- YearPubmed(metadata.pm)
hist(y, ylab = "Number of articles",
     xlab = paste0("Query date: ",
                   Sys.Date()),
     main = paste0("PubMed articles containing
                    viagra = ", length(y)))
```

Methods for **Medline** objects



Methods for **Medline** objects

Advanced questions: How is the the researcher having more papers?

```
AuthorList <- Author(metadata) # Extract list of authors
LastFirst <- sapply(AuthorList, function(x) paste(x$LastName,
                                                  x$ForeName))
sort(table(unlist(LastFirst)), dec = TRUE)[1:3] # Tabulate & Sor
```

```
Tanne Janice Hopkins       Charatan Fred     Abenhaim Lucien
                   4                   3                   1
```

Exercises

Exercises

- We are interested in performing a meta-analys on physical activity in COPD patients. Search for all papers published in Lancet about this (you only trust this journal)
  - How many papers where published last year?
  - And, how many where accepted? (look for a proper function with `ls(package:RISmed)`)
- Using the "rofecoxib" Medline object,
  - Determine the first year a matching article appeared.
  - What was the title of this article?
  - Do some authors have multiple matching records?
  - If so, which authors?

Preparing data

To conduct Meta-Analyses in R, you need to have your study data prepared. For a standard meta-analysis, the following information is needed for every study:

- The names of the individual studies, so that they can be easily identified later on. Usually, the first author and publication year of a study is used for this (e.g. Ebert et al., 2018).
- The Mean of both the Intervention and the Control group at the same assessment point (*or the number of events* for binary traits).
- The Standard Deviation of both the Intervention and the Control group at the same assessment point (not necessary for binary traits).
- The number of participants (N) in each group of the trial.
- If you want to have a look at differences between various study subgroups later on, you also need a factor variable containing such groups.

Exercise

Preparing data

As per usual, such data is stored in EXCEL spreadsheets. We recommend to store your data there, because this makes it very easy to import data into RStudio.

Author	This signifies the column for the study label (i.e., the first author)
Me	The Mean of the experimental/intervention group
Se	The Standard Deviation of the experimental/intervention group
Mc	The Mean of the control group
Sc	The Standard Deviation of the control group
Ne	The number of participants in the experimental/intervention group
Nc	The number of participats in the control group
Subgroup	This is the label for one of your Subgroup codes. It's not that important how you name it, so you can give it a more informative name (e.g. population). In this column, each study should then be given an subgroup code, which should be exactly the same for each subgroup, including upper/lowercase letters. Of course, you can also include more than one subgroup column with different subgroup codings, but the column name has to be unique

Exercise

1. Create an Excel file with this data and save it as 'example.xls' or 'example.xls'

	Study	Ne	Me	Se	Nc	Mc	Sc
1	Study 1	3	3.9450	1.1157970	3	6.8169	1.7699820
2	Study 2	5	2.2512	0.3275839	5	2.5961	0.6674662
3	Study 3	20	144.1000	25.7000000	20	183.6000	39.0600000
4	Study 4	16	59.9400	14.2800000	16	71.7200	14.3000000
5	Study 5	5	14.6200	2.2940000	5	17.8900	10.1850000
6	Study 6	5	11.1800	5.7490000	5	14.6800	5.0130000
7	Study 7	5	54.0300	1.6200000	5	63.9700	7.6600000
8	Study 8	5	4.9300	1.3500000	5	6.0100	1.0200000
9	Study 9	2	32.9900	5.3750000	2	47.3800	5.0150000
10	Study 10	2	44.8000	7.8790000	2	30.5900	2.1520000

NOTE: file is saved in the GitHub folder `data/example.txt`.
NOTE2: You can open it using Excel!

Preparing data

If you have already calculated the effect sizes for each study on your own, for example using Comprehensive Meta-Analysis or RevMan, there's another way to prepare your data which makes things a little easier. In this case, you only have to include the following columns:

Column	Description
Author	This signifies the column for the study label (i.e., the first author)
TE	The calculated effect size of the study (either Cohen's d or Hedges' g, or some other form of effect size
seTE	The Standard Error (SE) of the calculated effect
Subgroup	This is the label for one of your Subgroup codes. It's not that important how you name it, so you can give it a more informative name (e.g. population). In this column, each study should then be given an subgroup code, which should be exactly the same for each subgroup, including upper/lowercase letters. Of course, you can also include more than one subgroup column with different subgroup codings, but the column name has to be unique

Figure 3:

Goal of meta-analysis

- Pooling the effect sizes to get one overall effect size estimate of the studies.
- There are two approaches: the **Fixed-Effect-Model**, or the **Random-Effects-Model** (Borenstein et al. 2011).
- There is an extensive debate on which model fits best in which context (Fleiss 1993), with no clear consensus in sight.
- Although it has been recommended to only resort to the Random-Effects-Pooling model in clinical psychology and the health sciences (Cuijpers 2016)

Goal of meta-analysis

- Both models only require an effect size, and a dispersion (variance) estimate for each study, of which the inverse is taken.
- This is why the methods are often called generic inverse-variance methods.
- Meta-analyses can be performed with:
  - **continuous variables** (such as effect sizes), as these are the most common ones in psychology and the health science field.
  - **binary outcomes** which might be important if you're focusing on prevention trials or case-control or cohort studies.

Summarizing effects

Normal Assumption

$$Y_i \sim N(\theta, V_i)$$

Summary Effect Size is a **Weighted Average**

$$\hat{\theta} = \sum_i Y_i W_i / \sum_i W_i, \ \mathrm{Var}(\hat{\theta}) = 1 / \sum_i W_i$$

Each Study's Contribution

$$\lambda_i = W_i / \sum_i W_i$$

Figure 4:

Modelling approaches

- **Fixed effects**:
  - Same mean ES, zero between-study variance
- **Random effects**:
  - Different mean ES, between-study variance
- **Mixed effects**:
  - Study-level regression for mean ES

Same mean ES, known variance

$$Y_i = \theta + e_i,$$

$$e_i \sim N(0, V_i).$$

Figure 5:

Random effects

Different mean ES, between-study variance

$$Y_i = \theta + \theta_i + e_i,$$

$$\theta_i \sim N(0, \tau^2),$$

$$e_i \sim N(0, V_i).$$

Figure 6:

Mixed effects (meta-regression)

Study-level regression for mean ES

$$Y_i = \boldsymbol{\beta}' \mathbf{x}_i + \theta_i + e_i,$$

$$\theta_i \sim N(0, \tau^2),$$

$$e_i \sim N(0, V_i).$$

$\mathbf{x}_i$ = Study-level covariates

Figure 7:

Fixed vs. Random effects

- The *FE* model is a description of the studies.
- The *RE* model regards the studies as a sample of a larger universe of studies.
- The *RE* model can be used to infer what would likely happen if a new study were performed, the *FE* model cannot.
- Common practice is to report both fixed and random effects model results.

Random Effects

- Suppose between-study variance (τ^2) is non-zero.
- Methods differ on how they estimate τ^2.
- Many iterative and non-iterative approaches to estimating τ^2 have been proposed.

Random effects

Method	Estimator
DL	DerSimonian-Laird (Most Common)
HE	Hedges
HS	Hunter-Schmidt
SJ	Sidik-Jonkman
ML	Maximum-likelihood
REML	Restricted maximum-likelihood (Default)
EB	Empirical Bayes

No method is universally superior, but Viechtbauer's simulation study (2002) suggests REML has the most recommendable properties.

R packages for meta-analysis

- `meta` (Author: Guido Schwarzer)
- `rmeta` (Author: Thomas Lumley)
- `metafor` (Author: Wolfgang Viechtbauer)

Illustrative example (continuous variable)

Effect of amlopidine on work capacity

- **Overview**: 8 randomized clinical trials of the calcium channel blocker amlopidine vs placedo
- **Treatment goal**: improve work capacity
- **Primary endpoint**: ratio of exercise time
- **Data**: mean ratio in both groups and variances

Illustrative example (continuous variable)

```
library(meta)
data("amlodipine", package="meta")
res.cont <- metacont(n.amlo, mean.amlo, sqrt(var.amlo),
                     n.plac, mean.plac, sqrt(var.plac),
                     data=amlodipine, studlab=study)
```

Illustrative example (binary variable)

Corticosteroid to reduce premature labor deaths

- **Overview**: 7 randomized clinical trials of corticosteriod therapy vs. placebo
- **Treatment goal**: preventing neonatal deaths in preterm labor
- **Primary endpoint**: death
- **data**: the trial center, the number of deaths in the treatment group, the total number of patients in the treatment group, the number of deaths in the control group and the total number of patients in the control group

```
data("cochrane", package="rmeta")
cochrane
```

```
          name ev.trt n.trt ev.ctrl n.ctrl
1     Auckland     36   532      60    538
2        Block      1    69       5     61
3        Doran      4    81      11     63
4        Gamsu     14   131      20    137
5     Morrison      3    67       7     59
6 Papageorgiou      1    71       7     75
7      Tauesch      8    56      10     71
```

Fixed and Random Effects with R

```
library(rmeta)
cochrane
```

```
          name ev.trt n.trt ev.ctrl n.ctrl
1     Auckland     36   532      60    538
2        Block      1    69       5     61
3        Doran      4    81      11     63
4        Gamsu     14   131      20    137
5     Morrison      3    67       7     59
6 Papageorgiou      1    71       7     75
7      Tauesch      8    56      10     71
```

```
model.FE <- meta.MH(n.trt, n.ctrl, ev.trt, ev.ctrl,
                    names=name, data=cochrane)
model.RE <- meta.DSL(n.trt, n.ctrl, ev.trt, ev.ctrl,
                     names=name, data=cochrane)
```

Fixed and Random Effects with R

```
model.FE
```

```
Fixed effects ( Mantel-Haenszel ) Meta-Analysis
Call: meta.MH(ntrt = n.trt, nctrl = n.ctrl, ptrt = ev.trt, pctrl = ev.ctrl,
    names = name, data = cochrane)
Mantel-Haenszel OR =0.53    95% CI ( 0.39, 0.73 )
Test for heterogeneity: X^2( 6 ) = 6.9 ( p-value 0.3303 )
```

```
model.RE
```

```
Random effects ( DerSimonian-Laird ) meta-analysis
Call: meta.DSL(ntrt = n.trt, nctrl = n.ctrl, ptrt = ev.trt, pctrl = ev.ctrl,
    names = name, data = cochrane)
Summary OR= 0.53    95% CI ( 0.37, 0.78 )
Estimated random effects variance: 0.03
```

Fixed and Random Effects with R

```
library(meta)
res.bin.FE <- metabin(ev.trt, n.trt, ev.ctrl, n.ctrl,
                      data=cochrane, studlab = name,
                      comb.random = FALSE)
res.bin.RE <- metabin(ev.trt, n.trt, ev.ctrl, n.ctrl,
                      data=cochrane, studlab = name,
                      comb.fixed = FALSE)
```

Fixed and Random Effects with R

```
summary(res.bin.FE)
```

```
Number of studies combined: k = 7

                        RR           95%-CI     z  p-value
Fixed effect model 0.5646 [0.4254; 0.7493] -3.96 < 0.0001

Quantifying heterogeneity:
tau^2 = 0.0376; H = 1.09 [1.00; 1.58]; I^2 = 16.0% [0.0%; 59.9%]

Test of heterogeneity:
    Q d.f. p-value
 7.15    6  0.3076

Details on meta-analytical method:
- Mantel-Haenszel method
```

```
summary(res.bin.RE)
```

```
Number of studies combined: k = 7

                         RR           95%-CI     z p-value
Random effects model 0.5710 [0.4008; 0.8135] -3.10  0.0019

Quantifying heterogeneity:
tau^2 = 0.0376; H = 1.09 [1.00; 1.58]; I^2 = 16.0% [0.0%; 59.9%]

Test of heterogeneity:
    Q d.f. p-value
```

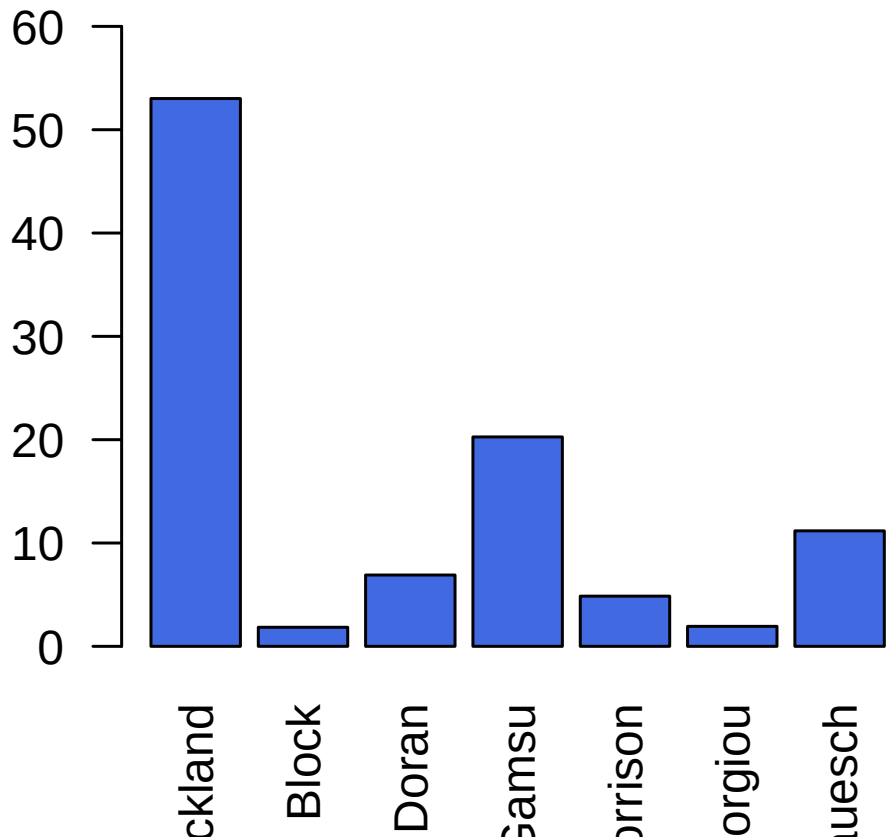
Fixed and Random Effects with R

```
res.bin <- metabin(ev.trt, n.trt, ev.ctrl, n.ctrl,
                   data=cochrane, studlab = name)
res.bin
```

```
                RR           95%-CI %W(fixed) %W(random)
Auckland    0.6068 [0.4086; 0.9011]      49.7       41.6
Block       0.1768 [0.0212; 1.4719]       4.4        2.7
Doran       0.2828 [0.0945; 0.8463]      10.3        9.3
Gamsu       0.7321 [0.3862; 1.3878]      16.3       22.6
Morrison    0.3774 [0.1022; 1.3938]       6.2        6.8
Papageorgiou 0.1509 [0.0190; 1.1959]      5.7        2.8
Tauesch     1.0143 [0.4287; 2.3997]       7.4       14.1

Number of studies combined: k = 7

                        RR           95%-CI     z  p-value
Fixed effect model   0.5646 [0.4254; 0.7493] -3.96 < 0.0001
Random effects model 0.5710 [0.4008; 0.8135] -3.10   0.0019

Quantifying heterogeneity:
tau^2 = 0.0376; H = 1.09 [1.00; 1.58]; I^2 = 16.0% [0.0%; 59.9%]

Test of heterogeneity:
    Q d.f. p-value
 7.15    6  0.3076

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

Study contribution

```
vari <- (res.bin$seTE)^2
contrib <- 1/vari/sum(1/vari) * 100
barplot(contrib, names = res.bin$studlab, ylim = c(0, 60),
        las = 2, col = "royalblue")
```



Exercise

Exercise

- Using the `example` dataset:
  - Estimate the pooled mean difference using Fixed and Random effects model
  - Which is the study having more contribution to such estimates?

How Much Can Estimates Of τ^2 Differ?

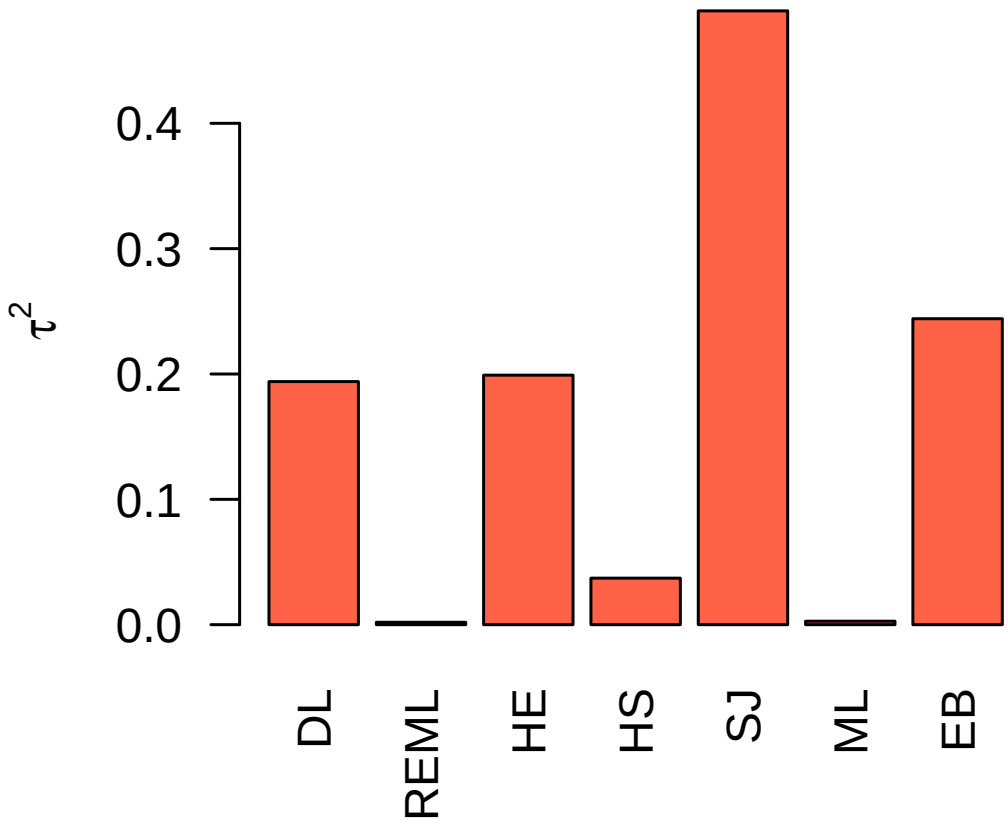
- DerSimonian-Laird (method.tau="DL")
- Restricted maximum-likelihood estimator (method.tau="REML")
- Maximum-likelihood estimator (method.tau="ML")
- Hunter-Schmidt estimator (method.tau="HS")
- Sidik-Jonkman estimator (method.tau="SJ")
- Hedges estimator (method.tau="HE")
- Empirical Bayes estimator (method.tau="EB").

How Much Can Estimates Of τ^2 Differ?

```
estimators <- c("DL", "REML", "HE", "HS", "SJ", "ML", "EB")
taus <- sapply(estimators, function(method) {
  metabin(ev.trt, n.trt, ev.ctrl, n.ctrl,
          data=cochrane, method.tau=method)$tau
})
```

```
barplot(taus, las=2, col="tomato",
        ylab=expression(tau^2))
```

How Much Can Estimates Of τ^2 Differ?



Testing for Heterogeneity: Q test

$$Q = \sum_i W_i (Y_i - \hat{\theta})^2$$

- Q is the weighted deviations about the summary effect size.
- Larger values of Q reflect greater between-study heterogeneity.
- When $\tau^2 = 0$, $Q \sim \chi^2(K-1)$, which leads to a chi-squared test for heterogeneity.

Figure 8:

```
res.bin$Q
```

```
[1] 7.14583
```

```
res.bin$pval.Q
```

```
[1] 0.3075713
```

Comments on Q test

- The chi-squared approximation is valid when study sample sizes are large.
- Type I error is generally accurate if normal distribution assumption and sample sizes are not too small.
- Q-test has low power (<0.80) when the number of studies and/or sample sizes is small.

-*NOTE*: If there are few trials in the meta-analysis (as is usually the case), the Q-test is likely underpowered for detecting true heterogeneity.

Other indices of Heterogeneity

- τ^2: variation among the effects observed in different studies (between-study variance)
- Higgin's I^2: Percentage of "unexplained" variance
- H^2, H-index: Unexplained heterogeneity
- Intra-class correlation (ICC): consistence among studies (0-1)

Thresholds for I^2

Judging the severity of measured heterogeneity is subjective, however Higgins suggests these rules of thumb:

- 0% to 30% Low
- 30% to 60% Moderate
- 50% to 90% Substantial
- 75% to 100% Considerable

H-index

Is the ratio of Q to the Q-test's degrees of freedom,

$$H^2 = \frac{Q}{df},$$

$$1/H^2 = 1 - \frac{I^2}{100}.$$

H index is the $\sqrt{H^2}$.

$H > 1$ suggests there is unexplained heterogeneity.

Figure 9:

Example

Confidence intervals are provided

```
res.bin
```

```
                RR           95%-CI %W(fixed) %W(random)
Auckland    0.6068 [0.4086; 0.9011]      49.7       41.6
Block       0.1768 [0.0212; 1.4719]       4.4        2.7
Doran       0.2828 [0.0945; 0.8463]      10.3        9.3
Gamsu       0.7321 [0.3862; 1.3878]      16.3       22.6
Morrison    0.3774 [0.1022; 1.3938]       6.2        6.8
Papageorgiou 0.1509 [0.0190; 1.1959]      5.7        2.8
Tauesch     1.0143 [0.4287; 2.3997]       7.4       14.1

Number of studies combined: k = 7

                         RR           95%-CI     z  p-value
Fixed effect model   0.5646 [0.4254; 0.7493] -3.96 < 0.0001
Random effects model 0.5710 [0.4008; 0.8135] -3.10   0.0019

Quantifying heterogeneity:
tau^2 = 0.0376; H = 1.09 [1.00; 1.58]; I^2 = 16.0% [0.0%; 59.9%]

Test of heterogeneity:
    Q d.f. p-value
 7.15    6  0.3076

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

Prediction intervals

- Previous measures have some drawbacks (mainly depend on studies sample size)
- **Prediction intervals (PIs)** are a good way to overcome this limitation (IntHout et al. 2016)
- PIs give us a range for which we can expect the effects of future studies to fall based on our present evidence in the meta-analysis.
- If our prediction interval, for example, lies completely on the positive side favoring the intervention, we can be quite confident to say that despite varying effects
- The intervention might be at least in some way beneficial in all contexts we studied in the future. If the confidence interval includes zero, we can be less sure about this, although it should be noted that broad prediction intervals are quite common, especially in medicine and psychology.

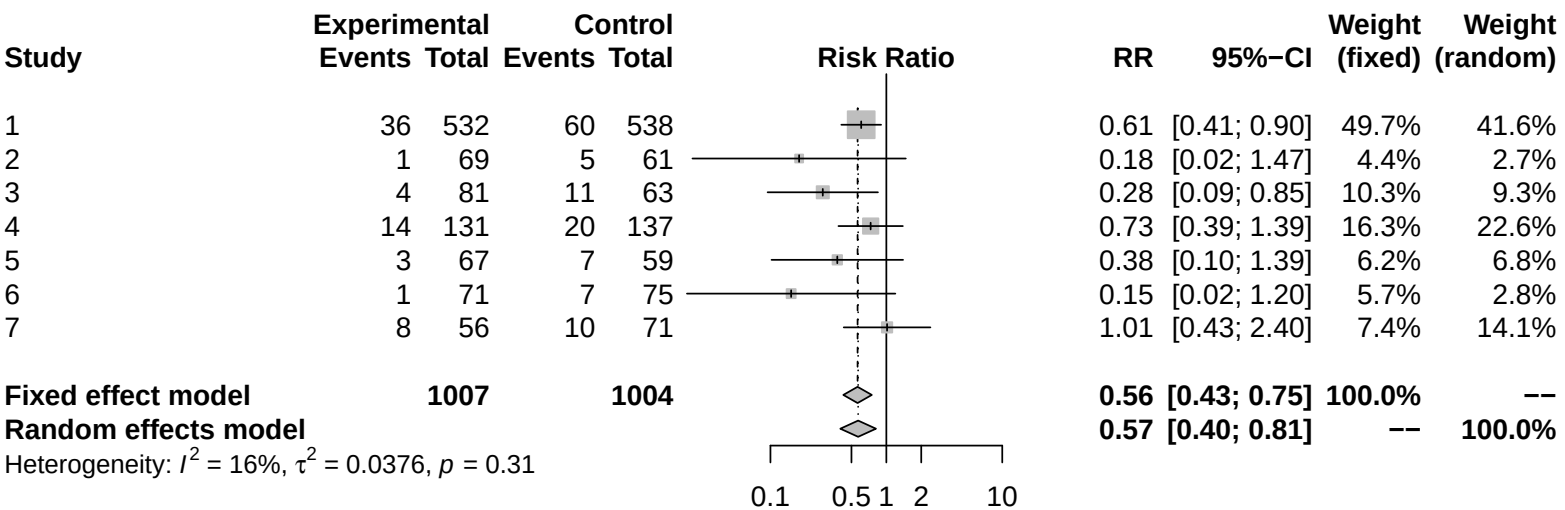
Visualization of Heterogeneity

Forest plot: *Seeing the forest through the trees. . .*

- Is a plot of effect sizes and their precisions
- Is the most common way to report the results of a meta-analysis
- Can help identify patterns across effects
- Can help spot large variation in effects or possible outliers

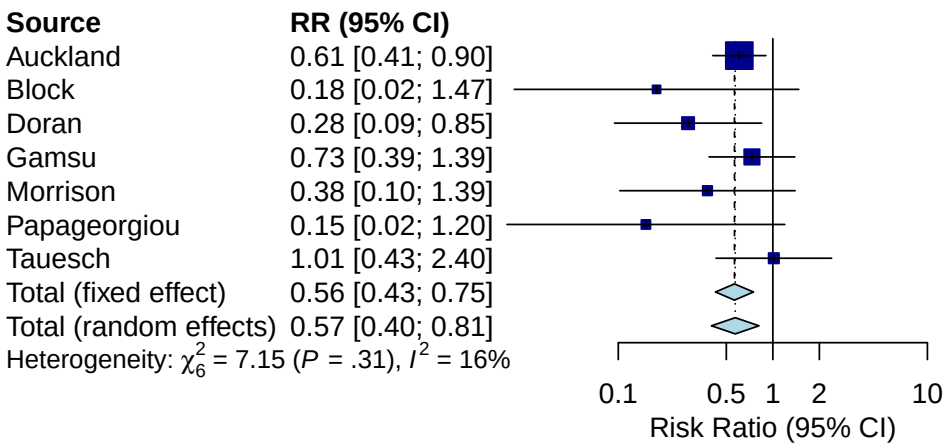
Visualization of Heterogeneity

```
forest(res.bin, studlab=cochrane$name)
```



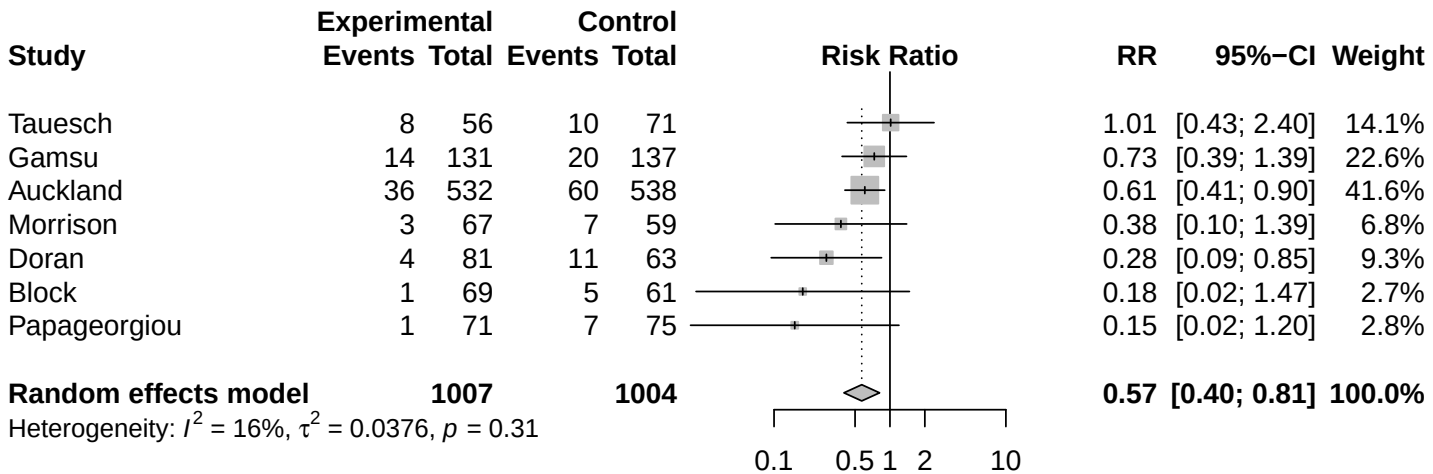
Visualization of Heterogeneity

```
forest(res.bin, layout = "JAMA")
```



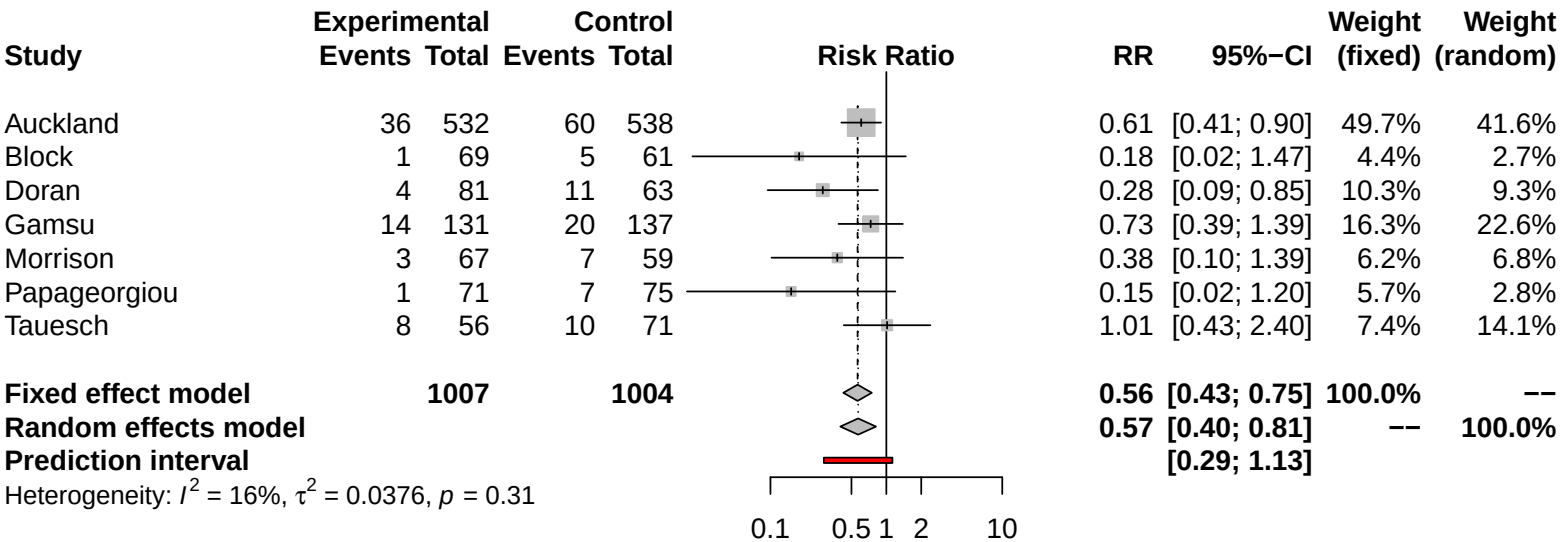
Visualization of Heterogeneity

```
forest(res.bin, sortvar=-TE, comb.fixed=FALSE)
```



Viualization of Prediction Interval

```
forest(res.bin, prediction = TRUE)
```



Exercise

Exercise

- Using the `example` dataset:
  - Create a forest plot. By looking at the figure, do you think there i heterogeneity?
  - Estimate the τ^2 parameter using DerSimonian-Laird, RMLE and Empirical Bayes estimators. Which estimator provides the lower value of between-study variance (e.g. less hetereogeneity)?
  - Which is the conclusion about heterogeneity using Higgins estimator? In that case, which is the recommended method to pool the effects accross studies?
  - Which is the predicted interval of the pooled effect? What is the conclusion?

Meta-analysis of observational studies

- Case control and cohort studies can provide only effect estimates (e.g Odds and Hazard Ratios, repectively)
- We can also have effect estimates and p-values
- We obtain beta and standard error estimates
- Meta-analysis can be performed using this aggregated data

Meta-analysis of observational studies

OR and 95%CI

```
studies <- paste0("Study", 1:5)
or <- c(1.23, 1.36, 1.08, 1.24, 1.81)
ciInf <- c(1.12, 1.06, 0.97, 1.20, 1.55)
beta <- log(or)
beta
```

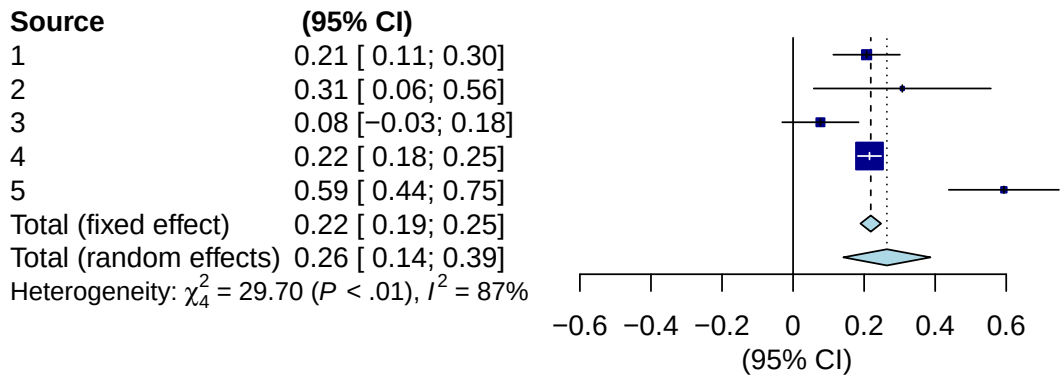
```
[1] 0.20701417 0.30748470 0.07696104 0.21511138 0.59332685
```

```
se <- (beta - log(ciInf)) / 1.96
se
```

```
[1] 0.04779872 0.12715091 0.05480625 0.01672950 0.07911832
```

Meta-analysis of observational studies

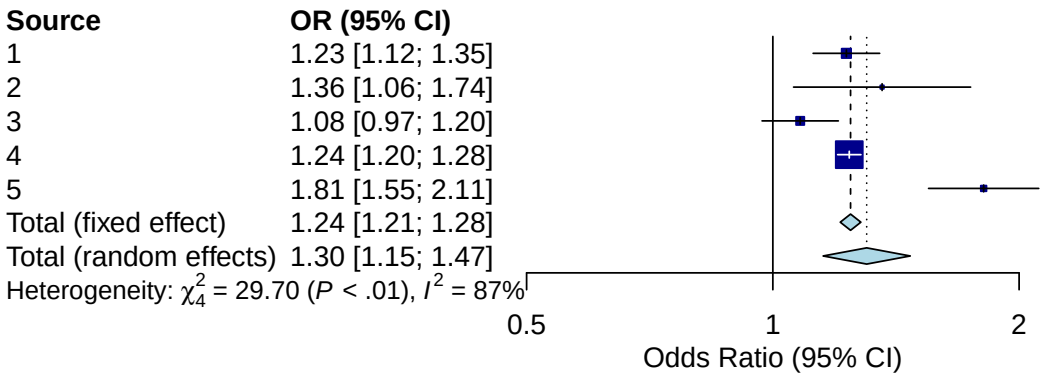
```
res.obs <- metagen(beta, se)
forest(res.obs, layout = "JAMA")
```



Meta-analysis of observational studies

Show original scale

```
res.obs <- metagen(beta, se, sm="OR")
forest(res.obs, layout = "JAMA")
```



Exercises

Exercises

- This table shows the information obtained in the literature about exiting studies on conscious sedation (CS) versus general anesthesia (GA) during endovascular acute ischemic stroke treatment

Study name	Statistics for each study			
	Odds ratio	Lower limit	Upper limit	p-Value
Sugg et al, AJNR, 2010	0.12	0.01	1.03	0.05
Davis et al, Anesthesiology, 2012	0.11	0.04	0.30	0.00
Abou-Chebl et al, Stroke, 2010	0.52	0.40	0.68	0.00
Abou-Chebl et al, Stroke, 2014	0.49	0.29	0.83	0.01
Li et al, J Neurosurg Anesthesiol, 2014	0.74	0.22	2.51	0.63
Jumaa et al, Stroke, 2010	0.35	0.16	0.78	0.01
Nichols et al, JNIS, 2010	0.19	0.06	0.56	0.00
Hassan et al, Neurocrit Care, 2012	0.25	0.11	0.59	0.00

Figure 10:

Exercises

The outcome reported corresponds to the effect of CV vs GS on having a good functional outcome after the intervention.

- Carry out a meta-analysis study to determine whether CS outperforms GS.
- Is there heterogeneity accross studies
- Create a forest plot using JAMA format and RevMan5 which is the recommended output for Cochrane reviews.

Sesitivity analysis

- A single outlying trial could be the source of substantial heterogeneity
- To identify suspicious cases, a leave-one-out method can be used whereby we rerun the meta-analysis, iteratively removing studies.
- In the `meta` package this is accomplished with the `metainf` function.

Sensitivity analysis

```
lev1out <- metainf(res.bin, pooled = "random")
lev1out
```

```
Influential analysis (Random effects model)

                           RR           95%-CI p-value   tau^2    I^2
Omitting Auckland      0.5078 [0.2923; 0.8821]  0.0162  0.1384  30.3%
Omitting Block         0.5927 [0.4194; 0.8376]  0.0030  0.0291  14.6%
Omitting Doran         0.6200 [0.4487; 0.8568]  0.0038  0.0109   5.4%
Omitting Gamsu         0.5098 [0.3207; 0.8102]  0.0044  0.0816  24.1%
Omitting Morrison      0.5801 [0.3891; 0.8648]  0.0075  0.0610  25.1%
Omitting Papageorgiou  0.6008 [0.4366; 0.8269]  0.0018  0.0136   7.4%
Omitting Tauesch       0.5348 [0.3791; 0.7545]  0.0004  0.0168   7.4%

Pooled estimate        0.5710 [0.4008; 0.8135]  0.0019  0.0376  16.0%

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

Outlier detection

- Suppose that a meta-analysis collects n studies. The observed effect size in study i is y_i and its within-study variance is s_i^2. Also, the inverse-variance weight is $w_i = 1/s_i^2$.
- Hedges and Olkin (1985) describe the outlier detection procedure for FE. Using the studies except study i, the pooled estimate of overall effect size is $\bar{\mu}_{(-i)} = \sum_{j \neq i} w_j y_j / \sum_{j \neq i} w_j$. The residual of study i is $e_i = y_i - \bar{\mu}_{(-i)}$. The variance of e_i is $v_i = s_i^2 + (\sum_{j \neq i} w_j)^{-1}$, so the standardized residual of study i is $\epsilon_i = e_i / \sqrt{v_i}$.
- Viechtbauer and Cheung (2010) describes the outlier detection procedure for RE. Using the studies except study i, let the method-of-moments estimate of between-study variance be $\hat{\tau}^2_{(-i)}$. The pooled estimate of overall effect size is $\bar{\mu}_{(-i)} = \sum_{j \neq i} \tilde{w}_{(-i)j} y_j / \sum_{j \neq i} \tilde{w}_{(-i)j}$, where $\tilde{w}_{(-i)j} = 1/(s_j^2 + \hat{\tau}^2_{(-i)})$. The residual of study i is $e_i = y_i - \bar{\mu}_{(-i)}$, and its variance is $v_i = s_i^2 + \hat{\tau}^2_{(-i)} + (\sum_{j \neq i} \tilde{w}_{(-i)j})^{-1}$. Then, the standardized residual of study i is $\epsilon_i = e_i / \sqrt{v_i}$.

Outlier detection

The function `metaoutlier` from package `altmeta` returns a list which contains standardized residuals and identified outliers. A study is considered as an outlier if its standardized residual is greater than 3 in absolute magnitude. Let's illustrate the detection of outliers on `amlodipine` example.

```
library(altmeta)
out <- metaoutliers(res.cont$TE,       # observed effect sizes
                    res.cont$seTE^2)   # within-studies variances
```

```
This function uses random-effects meta-analysis because Ir2 >= 30%.
```

```
out$outliers
```

```
[1] "All the standardized residuals are smaller than 3"
```

Exercise

Exercise

1. Using the `example` dataset:
   - Does the removal of any trial change the main conclusion about the effect?
   - Is there any study that can be considered as an outlier, and hence, being removed from the analysis?
2. The data `Fleiss93` contains information about a meta-analysis on aspirin in preventing death after myocardial infarction

- Load the data into R by `data(Fleiss93, package="meta")`
- Which trial contributes the most to the OR meta-analysis
- Do any of the trials reduce I^2 to $< 30\%$?
- Does the removal of any trial change the main conclusion about the efficacy of the aspirin?
- Is there any outlier?

Session info

```
sessionInfo()
```

```
R version 3.5.0 (2018-04-23)
Platform: x86_64-w64-mingw32/x64 (64-bit)
Running under: Windows 10 x64 (build 17134)

Matrix products: default

locale:
[1] LC_COLLATE=Spanish_Spain.1252  LC_CTYPE=Spanish_Spain.1252
[3] LC_MONETARY=Spanish_Spain.1252 LC_NUMERIC=C
[5] LC_TIME=Spanish_Spain.1252

attached base packages:
[1] stats     graphics  grDevices utils     datasets  methods   base

other attached packages:
[1] altmeta_2.2  rmeta_3.0    meta_4.9-4   RISmed_2.1.7

loaded via a namespace (and not attached):
 [1] Rcpp_1.0.1       lattice_0.20-35  codetools_0.2-15 digest_0.6.15
 [5] grid_3.5.0       nlme_3.1-137     metafor_2.0-0    magrittr_1.5
 [9] evaluate_0.13    stringi_1.2.2    Matrix_1.2-14    rmarkdown_1.12
[13] tools_3.5.0      stringr_1.3.1    xfun_0.5         yaml_2.2.0
[17] compiler_3.5.0   htmltools_0.3.6  knitr_1.22
```