MI260: Introduction to model-based meta-analysis

Hands-on Exercises: Measures of heterogeneity and random effects meta-analysis

Introduction

In this exercise, we'll work through calculating measures of heterogeneity and fitting a random-effects meta-analysis of the atorvastatin 80 mg LDL data.. We'll begin by doing these calculations 'by hand', and then use the `metafor` package for fitting the models for us. I'll then ask you to go through similar steps for the pancreatic cancer mortality incidence data.

Files Provided

Datasets:

- AtorvastatinStudies.csv – LDL data for placebo-controlled studies with atorvastatin arms
- CochranePancreaticData12months.csv– mortality incidence data from 15 studies comparing gemcitabine to a gemcitabine+chemotherapy regimen

R code:

- RandomEffectsExercise-Part1.r – code for reading data, calculating measures of heterogeneity and doing a random effects analysis of the atorvastatin LDL data
- RandomEffectsExercise-Part2.r – similar code the pancreatic cancer 12 month mortality incidence data

Part 1:
LDL data from 8 placebo-controlled atorvastatin studies

1. Using the difference in mean percent change from baseline LDL between 80 mg atorvastatin and placebo, calculate the Q statistic and test the null hypothesis of no between-study variability in the mean placebo-adjusted response. Would you conclude that there is significant between-study heterogeneity based on this test?

```
k <- nrow(ator80)-1
Q <- sum( (ator80$seDiff^-2)*(ator80$diff-feEst)^2 )
pvalue.Q <- 1-pchisq(Q,k)
```

2. Calculate the I^2 and H statistics. How would you interpret these values? What might you do to explore this heterogeneity further?

```
H <- sqrt(Q/k)

I2 <- 100*(Q-k)/Q
```

3. Given your estimate of Q, derive the DerSimonean and Laird estimate of τ^2 and the random effects estimate and 95% confidence interval for the mean difference from placebo. How do these compare to the estimates based on the fixed effects model?

```
tau2.dl <- max( 0, (Q - k) / ( sum(w) - sum(w^2)/sum(w) ) )
wre <- 1/(ator80$seDiff^2 + tau2.dl)
reEst <- sum(wre * d$diff) / sum(wre)
varReEst <- 1/sum(wre)
reCI <- reEst + c(-1.96, 0, 1.96)*sqrt(varReEst)

# Compare the relative weights between fixed and random effects
# models
print(w / sum(w))
print(wre / sum(wre))
```

4. Fit the random effects meta-analysis model using the rma function. Get both the DerSimonean and Laird and the REML estimates of τ^2. Do the estimates differ? How do these influence your estimates of the mean naproxen difference from placebo and 95% confidence intervals? How do these compare to the fixed effects analysis? Do you prefer the fixed or random effects model?

```
rema.dl <- rma(yi=diff, sei=seDiff^2, data=ator80, method="DL",
slab=trial)
confint(rema.dl)
coef(rema.dl)

rema.reml <- rma(yi=diff, vi=seDiff^2, data=ator80,
method="REML", slab=trial)
confint(rema.reml)
```

5. Make a forest plot of the mean difference from placebo based on the rma object. Add the estimate from the fixed effect model to the figure for comparison.

```
forest(rema.reml, order=rev(order(ator80$diff)),
      ylim=c(-2.5,nrow(ator80)+3), refline=NA,
      mlab="Random Effects Estimate",
      main="Difference in %CFB LDL for 80 mg atorvastatin")
text(-120.5, 7.5, "Study", cex=0.95, adj=0, font=2)
text(-30, 7.5, "Mean Diff. + 95%CI",cex=0.95, adj=0, font=2)
addpoly(feEst, vi=1/sum(w), rows=-2,
            mlab="Fixed Effects Estimate",
```

Part 2:

Month 6 Mortality Incidence in pancreatic cancer from Cochrane Collaboration

We'll repeat the same sort of analysis as above but now using pancreatic cancer Month 12 Mortality incidence data.

1. Calculate the Q statistic and test the null hypothesis of no between-study variability in the log odds ratio. Would you conclude that there is significant between-study heterogeneity of effect based on this test?

2. Calculate the I^2 and H statistics for the log odds ratio comparing G+C to G. How would you interpret these values?

3. Given your estimate of Q, derive the DerSimonean and Laird estimate of τ^2 and the random effects estimate and 95% confidence interval for the log odds. How do these compare to the estimates based on the fixed effects model?

4. Fit the random effects meta-analysis model using the `rma` function. Get both the DerSimonean and Laird and the REML estimates of τ^2. Do the estimates differ? How do these influence your estimates of the probability of paresthesia and 95% confidence intervals? How do these compare to the fixed effects analysis? Do you prefer the fixed or random effects model?

5. Make a forest plot of the log odds ratio based on the `rma` object. Include the meta-analysis estimates from the two random effects models and the fixed effects model.