Stata practical 1a – two-stage IPD meta-analysis with a binary outcome

Question 1

ML used to estimate parameters:

Pooled OR = 1.351 95% CI (1.029 to 1.773), τ^2-hat=0.

DL:

Pooled OR = 1.517 95% CI (0.931 to 2.472), τ^2-hat is 0.112. This is much larger than for ML estimation, the pooled OR is larger and the 95% CI for the pooled result is much wider.

REML:

Pooled OR = 1.546 95% CI (0.900 to 2.655), τ^2-hat=0.152. This is larger than DL and ML and thus the pooled OR is larger and the 95% CI is the widest for this estimation method compared to the other two methods.

Question 2

The estimated 95% prediction interval is (0.01 to 304.3) – it is this wide because the estimate of τ is large and the prediction interval uses the *t*-distribution with degrees of freedom equal to k-2 studies. There are only 3 studies in the meta-analysis and therefore the value from the *t*-distribution is very large. Clearly, it is not sensible to derive a prediction interval with few studies.

Stata practical 1b – two-stage IPD meta-analysis with a continuous outcome

Question 1

For some trials the mean SBP is not balanced as closely as other trials. For example, trial 1, 153.0 and 150.9, whereas trial 6, 182.1 and 182.2.

Question 2

ANCOVA: -6.53 95% CI (-8.24 to -4.83), s.e. 0.87;
Final score: -7.21 95% CI (-8.98 to -5.44), s.e. 0.90;
Change score: -5.04 95% CI (-7.06 to -3.03), s.e. 1.03.

The estimates, 95% CI and s.e, are different for all three analyses. The standard error is smallest for ANCOVA as expected. The treatment effect is smallest for the change score analysis. The ANCOVA analysis is more reliable because it adjusts for the baseline blood pressure, which is important particularly when there is baseline imbalance in the groups, and it also accounts for the correlation between baseline and follow-up values. The other methods do not account for both to these issues, and thus their standard errors are always larger.

Question 3

Yes, hypertension treatment does appear effective at reducing systolic blood pressure more than placebo since there is a statistically significant treatment effect estimate of -9.80 95% CI (-11.12 to -8.48).

Question 4

The between-study variance estimate in treatment effect is 3.297. Therefore, yes, there is a lot of heterogeneity in the treatment effect across the trials in the meta-analysis.

Question 5

The two-stage random effects model with REML estimation estimates a 95% prediction interval of (-14.26 to -5.33). In an individual (new) study, the true treatment effect is expected to lie between -14.26 and -5.33, so there is strong evidence to suggest that the hypertension treatment is effective in all populations, even given the large heterogeneity.

Stata practical 1c – two-stage IPD meta-analysis with a survival outcome

Question 1

DL: pooled estimate=0.73, 95% CI (0.54 to 0.99), τ^2-hat=0.312.

ML: pooled estimate=0.73, 95% CI (0.52 to 1.02), τ^2-hat=0.386.

REML: pooled estimate=0.72, 95% CI (0.51 to 1.03), τ^2-hat 0.422.

The estimate of τ changes for the estimation methods – it is higher with ML compared to DL and higher again with REML.

Question 2

Based on the REML analysis, the average effect of treatment suggests that it is effective, however, it is not quite statistically significant and thus further research might be needed.

Question 3

Pooled estimate=0.73, 95% CI (0.49 to 1.08), τ^2-hat=0.312.

The confidence intervals for the pooled estimates are wider with the Hartung-Knapp adjustment, and include a HR value of 1 so there is even less evidence now of a statistically significant effect.

Stata practical 2a – one-stage IPD meta-analysis with a binary outcome

Question 1(a)

$dvt_i \sim bin(1, p_i)$ for patient i=1,…m

$Logit(p_i)=\alpha + \beta*eryt_i$

Question 1(b)

OR 1.25 95% CI (0.96 to 1.63)

Eryt increases odds of dvt by 25% but non-significant result.

Eryt is not statistically significantly associated with the outcome of DVT since OR 1.25 and 95% CI (0.96 to 1.63). Might suggest further research is needed, as majority of CI is above 1.

Question 2

$Logit(p_{ij})=\alpha_1*study1_i + \alpha_2*study2_i + \alpha_3*study3_i + \beta*eryt_{ij}$ for study j=1,…,N, where $study1_i$ = 1 if in study 1, else 0 etc.

OR 1.35 95% CI (1.03 to 1.77) – evidence of a statistically significant increase in the odds of dvt by 35% for those with eryt, 95% CI suggests increase in odds of DVT between 3% and 77%.

Question 3

The model that ignores clustering has a lower summary OR than the one that properly accounts for clustering; furthermore, the latter suggests a statistically significant increase in odds for those with eryt, whereas the model that ignores the clustering by trial does not suggest that there is statistically significant increase.

The prevalence of dvt (i.e. baseline risk) varies for the three trials – the prevalence is lower for study1 (12.7%) compared to study2 (17.0%) and is highest for study3 (22.4%). The analysis that ignores clustering does not account for this variation in baseline risk, and indeed assumes baseline risk is the same for all studies. This is a strong assumption to make, and leads to a too precise CI and an attenuated effect size. Another way of viewing this is that responses within the same study are correlated, due to being in the same study, and this is only accounted for by allowing for a separate intercept per study.

Question 4

$Logit(p_{ij})= \alpha_1*study1_i + \alpha_2*study2_i + \alpha_3*study3_i + \beta_j*eryt_{ij}$ for study j=1,…N.

$\beta_j=\beta+u_j$, $u_j \sim N(0,\tau^2)$

There is no evidence of between-study heterogeneity since the estimate of the standard deviation is basically zero. Though of course, with 3 studies there is extremely low power to detect genuine heterogeneity.

Stata practical 2b – one-stage IPD meta-analysis with a continuous outcome

Question 1

$Sbpl_{ij} = \alpha_1 * trial1_{ij} + \alpha_2 * trial2_{ij} + \ldots \alpha_{10} * trial10_{ij} + \beta_1 * treat_{ij} + \beta_2 * sbpb_{ij} + \epsilon_{ij}$

$\epsilon_{ij} \sim N(0,\sigma^2)$

Where trial1 = 1 if in trial 1 and 0 otherwise, etc. Note we assumed fixed adjustment effect for baseline (beta2) and the same residual variance in each study (sigma-squared).

Yes, hypertension treatment does appear effective at reducing systolic blood pressure more than placebo since there is a statistically significant treatment effect estimate of -9.38 95% CI (-9.78 to -8.97).

Question 2

$Sbpl_{ij} = \alpha_1 * trial1_{ij} + \alpha_2 * trial2_{ij} + \ldots \alpha_{10} * trial10_{ij} + \beta_{1j} * treat_{ij} + \beta_2 * sbpb_{ij} + \epsilon_{ij}$

$\beta_{1j} = \beta_1 + u_j,\ u_j \sim N(0,\tau^2)$

$\epsilon_{ij} \sim N(0,\sigma^2)$

The estimate of between-study variance in treatment effect is 8.99, 95% CI (2.70 to 29.96). There is a lot of heterogeneity in the treatment effect across the trials in the meta-analysis. The summary treatment effect estimate is -10.32 95% CI (-12.32 to -8.32). Conclusions are largely the same for the summary estimate as in the fixed effect analysis. The average effect of treatment is slightly larger compared to the fixed effect of treatment in Question 1 but the 95% CI for the average treatment effect is wider in the random effects analysis compared to the fixed effect. However, both suggest a strong clinical benefit, on average.

Question 3

Two-stage random effects result with REML estimation:

Summary treatment effect estimate = -10.12 95% CI (-11.98 to -8.26), tau-sq = 7.42.

There is little difference between the one-stage and the two-stage meta-analysis results – the hypertension treatment is still highly effective. There are slight differences in the values due to different modelling assumptions. The two-stage automatically assumes different residual variances and distinct adjustment factors in each trial whereas the one-stage does not.

Question 4

It is very slow to estimate the parameters. The Kenward-Roger option is also very slow.

Stata practical 2c – one-stage IPD meta-analysis with a survival outcome

Question 1

$h_i(t) = h_0(t)*\exp(\beta*treat_i)$, patient i.

HR=0.79, 95% CI (0.71 to 0.88). Hazard of death is statistically significantly reduced by 21% for those who have received hypertension treatment.

Question 2

$h_{ij}(t) = h_0(t)*\exp(\beta_1*trial1 +\ldots+\beta_{14}*trial14 + \beta_{15}*treat_{ij})$, patient i, trial j.

(trial 15 is the reference group)

Or: $h_{ij}(t) = h_0(t)*\exp(\beta_{0j} + \beta_1*treat_{ij})$, patient i, trial j.

The model assumes that the baseline hazard functions of the trials are proportional to a common baseline hazard function.

Question 3

Trial15 is omitted from the results because it is the reference trial that is included in the baseline hazard function, $h_0(t)$, and so β_{015} is constrained to be 0 (β_{0j} is the proportional effect on the baseline hazard function due to the j^{th} trial for j=1,…,14).

Question 4

HR=0.77, 95% CI (0.69 to 0.86) – yes the results do still suggest that hypertension treatment is effective.

Question 5

Yes, there is heterogeneity in the treatment effect since sd(treat) is 0.61 95% CI (0.39 to 0.95) in the flexible parametric model.

Compared to the previous two-stage meta-analysis using ML, the estimate of τ is very similar ($\hat{\tau}$=0.62 95% CI (0.40 to 0.99) in previous two-stage). Note that the Stata output from ipdmetan is for τ^2 rather than τ.

Question 6

According to the results, hypertension treatment is still effective:

HR=0.713, 95% CI (0.510 to 0.996).

Stata practical 3 - Estimation of effect modifiers (interactions)

Throughout, we are looking to estimate the difference in treatment effect for males compared to females.

Question 1

Meta-regression: regress the study treatment effect estimates ($\hat{\theta}_i$ for study i) against average patient-level covariates:

$$\hat{\theta}_i = \alpha_i + \gamma(proportion\ male) + u_i + \varepsilon_i$$

$$\varepsilon_i \sim N(0, V(\hat{\theta}_i))$$

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

Here, γ is the across-study interaction, which tells us how much the average treatment effect differs in a study with only males compared to a study with only females.

Here, γ-hat is 15.5, 95% CI (8.2 to 22.8), which quantifies the difference in the effect of treatment for male only studies compared to female only studies. For example, if the female only studies have an underlying treatment effect that reduces SBP by 20mmHg, then male only studies have an underlying treatment effect that reduces SBP by 4.5mmHg.

Question 2

In question 1, τ^2 was estimated as 0.833. Here, without a covariate for proportion male, it was estimated as 7.4.

Question 3

$SBPl_{ij} = \alpha_i * trial_i + \beta_1 * SBP_{0ij} + \beta_2 * sex_{ij} + \theta_i * treat_{ij} + \gamma * treat_{ij} * sex_{ij} + \epsilon_{ij}$

Where α_i is the separate baseline (intercept) per trial

$\theta_i = \theta + u_i$, $u_i \sim N(0, \tau^2)$

$\epsilon_{ij} \sim N(0,\sigma^2)$

(This model only fits the amalgamated interaction between treatment and sex).

Stata practical 4a – deriving within-study correlations for multivariate meta-analysis

Question 1

Yes, the treatment does appear to be effective for reducing SBP and DBP.

SBP -6.664, s.e=0.850, 95% CI (-8.332 to -4.996).

DBP -2.989, s.e=0.522, 95% CI (-4.012 to -1.965)

Question 2

Fit a model using seemingly unrelated regression command sureg.

$$SBP_{ij} = \alpha_{1i} + \beta_{1i} * SBP_{0ij} + \theta_{1i} * trt_{ij} + \varepsilon_{ij1}$$

$$DBP_{ij} = \alpha_{2i} + \beta_{2i} * DBP_{0ij} + \theta_{2i} * trt_{ij} + \varepsilon_{ij2}$$

$$\varepsilon_{ij1} \sim N(0, \sigma_{i1}^2)$$

$$\varepsilon_{ij2} \sim N(0, \sigma_{i2}^2)$$

$$cov(\varepsilon_{ij1}, \varepsilon_{ij2}) = \sigma_{i12}$$

SBP -6.628, se 0.849, 95% CI (-8.293 to -4.964).

DBP -2.975, se 0.521, 95% CI (-3.997 to -1.953)

Correlation estimate: 0.7756.

The estimates are very similar to before, though slightly different with marginally smaller standard errors, due to the utilisation of correlation.

Question 3

The patient level correlation is the correlation in the residuals for SBP and DBP which is 0.776, and this tells us that there is a strong positive correlation between final SBP and DBP values in patients in trial 1.

Question 4

Within-study covariance between SBP and DBP treatment effect estimates is 0.3433.

Within-study correlation between SBP and DBP treatment effect estimates is $\frac{cov(SBP,DBP)}{\sqrt{var(SBP)*var(DBP)}} = \frac{0.3433}{\sqrt{0.7214*0.2717}} = 0.7755.$

Question 5

The estimate of correlation from the bootstrap samples is 0.7691, which is very similar albeit slightly smaller.

Question 6

The estimate of correlation for the partially and fully adjusted hazard ratios is 0.9915. Thus, there is a near perfect correlation between them, which is perhaps not surprising.

Stata practical 4b – getting to know 'mvmeta'

Question 1

Yes, the results from *metaan y1 se1, reml* and *mvmeta y v, nounc vars(y1)* are identical. This is because both used REML, and the *nounc* option was specified within *mvmeta* (i.e. thus neither method is inflating the confidence intervals to account for the uncertainty of the estimated tau-squared). If the between-study standard deviation from *mvmeta* is squared, this gives the same value as the estimate of τ^2 in the *metaan* results.

Question 2

Accounting for the additional uncertainty slightly increases the standard error but this does not change the conclusions.

Question 3

The analysis dropped 17 studies because there is no data for the fully adjusted result in these 17 studies, and we were only fitting a univariate analysis.

Question 4

The fitted model for the bivariate meta-analysis model:

$$\begin{pmatrix} Y_{i1} \\ Y_{i2} \end{pmatrix} \sim N\left\{ \begin{pmatrix} \theta_{i1} \\ \theta_{i2} \end{pmatrix}, \boldsymbol{S}_i \right\}$$

$$\boldsymbol{S}_i = \begin{pmatrix} s_{i1}^2 & \rho_{W_i} s_{i1} s_{i2} \\ \rho_{W_i} s_{i1} s_{i2} & s_{i2}^2 \end{pmatrix}$$

$$\begin{pmatrix} \theta_{i1} \\ \theta_{i2} \end{pmatrix} \sim N\left\{ \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}, \boldsymbol{D} \right\}$$

$$\mathbf{D} = \begin{pmatrix} \tau_1^2 & \rho_B \tau_1 \tau_2 \\ \rho_B \tau_1 \tau_2 & \tau_2^2 \end{pmatrix}$$

The β's represent the true pooled (mean) effects. The estimates of these in the results are β_1=0.271, 95% CI (0.219 to 0.324), β_2=0.347, 95% CI (0.288 to 0.405). (Note: These

Stata practical 4c – advanced multivariate and network meta-analysis

Question 1

Yes, all four outcomes show significant results for the treatment effect in this multivariate meta-analysis model.

Question 2

The borrowing of strength (BoS) is 0.9, 0.5, 4.2, and 20.8 for outcomes 1, 2, 5, and 6, respectively.

These values are lower than the fibrinogen example because there is no missing data in this example, and usually BoS is small in such situations.

Question 3

Yes, the treatment reduces PP more than control on average across the studies. PP: -5.59, 95% CI (-6.90 to -4.28).

Question 4

If there is a missing value for the covariance for even one trial in the dataset, *mvmeta* will not run. We could impute a value but what value do we choose?

Question 5

The treatment effect estimate for B versus A is -0.161, 95% CI (-0.252 to -0.070).

The treatment effect estimate for C versus A is 0.002, 95% CI (-0.061 to 0.066).

Question 6

All the between-study variances and covariances were fixed at the same value. This was done to be able to estimate the model parameters and to make sure that all comparisons are possible, not just those of the treatments compared to A. Using an unstructured between-study variance-covariance matrix is likely to be impossible in most situations with 3 or more treatments, due to not all pairs being available in all studies. So, although not ideal, the simplification is used mainly for pragmatic reasons to allow modelling to continue. Between-study correlations are forced to be 0.5 ensures all contrasts have the same τ^2 also.

Question 7

This estimates the treatment effect for B versus C.

Keele University - Important Information: Your IT Account

These details are for your personal use only. Do not disclose this information to anyone else and destroy this page once you have memorized the details.

If you receive this information in error please contact PrimaryCare IT

By registering at Keele, you are agreeing to abide by the IT Conditions of Use.

http://www.keele.ac.uk/itconditions

Name	Temporary 2 Day Account
Account/username - **Password -**	zzg61 v3jnk5e8

Your user name and password should be used to access all of these resources:

- Connecting devices to the internet
- Network Drive Access

To Connect to WIFI do the the following:

- Connect to the "Visitor" network.
- Open a web browser, and go to www.google.co.uk
- Log in with the username and password provided.
- Internet access should now be working.

If there are any issues please contact us on the below details.

PrimaryCare IT
Email: primarycare.it@keele.ac.uk Tel: 01782 734708 / 734893 Room: 0.50 (Admin Office)

COURSE TIMETABLE

DAY 1:

9:00 – 9:30: Registration & tea/coffee

9:30 - 9:45: Welcome & Agenda for the day (includes MAD MINUTES for faculty)

9:45 - 10:15: Lecture 1: Rationale for meta-analysis & IPD (includes objectives for the course)

10:15 - 11:15: Lecture 2: Two-stage IPD meta-analysis

11:15 - 11:40: Tea & Coffee

11:40 – 11:55: "MAD MINUTE": Part A

11:55 – 13:00: Stata Practical 1: Two-stage IPD meta-analysis

13:00 – 13:45: Lunch

13:45 - 14:45: Lecture 3: One-stage IPD meta-analysis

14:45 – 15:00: "MAD MINUTE": Part B

15:00 - 16:00: Stata practical 2: One-stage IPD meta-analysis

16:00 – 16:30: Tea and coffee

16:30 – 16:45: "MAD MINUTE": Part C

16:45 – 17:45: Lecture 4: Special topics (Power, one-stage weights, prognostic factors, DTA)

19:00: EVENING MEAL AT LOCAL PUB (SNEYD ARMS, KEELE)

DAY 2:

09:00 – 10:00: Lecture 5: Estimation of effect modifiers (interactions)

10:00 - 10:45: Stata practical 3: Interactions

10:45 – 11:05: Tea and Coffee

11:05 - 11:30: Stata practical 3: Interactions (continued)

11:30 – 12:30: Lecture 6 (Part 1): Multivariate meta-analysis using IPD

12:30 – 13:15: Lunch

13:15 – 14:00: Lecture 6 (Part 2): Multivariate meta-analysis using IPD

14:00 – 15:15: Stata practical 4: Multivariate and network meta-analysis

15:15 – 15:30: Tea and coffee

15:30 – 16:45: GUEST LECTURE: Dr Thomas Debray (Utrecht) "Multiple imputation in an IPD meta-analysis"

16:45: CLOSE & FEEDBACK