

Practical: Network Meta-Analysis using MetaInsight worksheet

Nicola Cooper, Alex Sutton and Yiqiao Xin

NIHR CRSU - Complex Review Support Unit <http://www.nihrcrsu.org/>

This practical uses the newly developed CRSU app *MetaInsight (Continuous)* which can be accessed from <https://crsu.shinyapps.io/metainsightc/>

We have also developed MetaInsight (Binary) for binary data:

<https://crsu.shinyapps.io/metainsightb/>

MetaInsight 1.1 conducts frequentist NMA. We also recently developed MetaInsight 2.0 Beta which also conducts Bayesian NMA utilising the GEMTC package in R (which itself utilises the Bayesian analysis package JAGS). The last part of the practical will use the Bayesian version of the app, which can be accessed from the links below:

Bayesian NMA for continuous outcome: https://crsu.shinyapps.io/metainsight_continuous2/

Bayesian NMA for binary outcomes: https://crsu.shinyapps.io/metainsight_binary2/

MetaInsight is a web-based App for conducting network meta-analysis for binary and continuous outcomes using fixed or random effects models. It may also be used to facilitate sensitivity analysis via the inclusion and exclusion of studies as this practical will demonstrate.

After completing this practical, you will be able to....

- Upload data into *MetaInsight*
- Perform both frequentist and Bayesian network meta-analysis in *MetaInsight* and interpret the results
- Examine the influence on the results when studies are included and excluded in a sensitivity analysis

In this practical you will be using a real network meta-analysis dataset comparing drug treatments and placebo for reducing obesity (Gray LJ, et al. *A systematic review and mixed treatment comparison of pharmacological interventions for the treatment of obesity. Obes Rev. 2012 Jun;13(6):483-98.*). Interest lies in which of the interventions is most effective for reducing weight. The outcome measure in each study arm is weight loss from study baseline in kilograms. (The whole dataset comes ready loaded into the App but you are going to start by only looking at a subset of the data and learning to load the data into the App).

Firstly, open the *MetaInsight* App in a web browser on your computer:

<https://crsu.shinyapps.io/metainsightc/>

The outcome for the analysis is continuous so we will use the “MetaInsight (continuous)” app.

To load the data into the app, please click ‘Load Data’ page. MetaInsight v1.1 allows data in both long format (1 study arm per row) and wide format (1 whole study per row). On the left hand sidebar of this page, you will see ‘Step 1’ which contains an upload bar for uploading the data and ‘Step 2’ which contains a box for uploading the treatment labels. Instructions on how to upload data and labels are on the main panel (right side) of this page. ‘Long format upload’ tab contains

instructions for uploading long format data and similarly, 'Wide format upload' tab for uploading wide format data. 'View data' tab will display any data that are uploaded. Please note that these will have data in already as the App comes with an example dataset already included for illustration purposes.

The long format data should contain six columns with the following exact headings (case sensitive): 'StudyID', 'Study' (study name), 'T'(treatment code), 'N'(number of participants for each arm), 'Mean'(the mean value of the outcome in each arm, i.e. in this example it would be change in weight from baseline), and 'SD' (the Standard deviation of the outcome in each arm). **Please note that the reference treatment needs to be labelled as 1.**

The second box includes the labels for each of the treatment codes. The first row must be 'Number' tabspace 'Label' as shown in the pre-loaded format, case sensitive.

The pre-loaded example data, and the treatment labels, presented in the correct format, can be downloaded by clicking the 'Download' button on the page.

If you prefer the wide format data, similar to the long format instruction tab, you could find the detailed instructions on the 'Wide format upload' tab. The example data in wide format could also be downloaded.

PART ONE:

Initially we are going to keep things simple by including only comparisons between three treatments: Orlistat, Sibutramine and Placebo.

The data file, named 'Data_3treatments.csv', and the treatment label file, named 'labels_3treatments.txt' can both be downloaded from http://www.nihrcrsu.org/workshops/joint_crsu_cochrane_workshop_2019/

- i) Step 1: Load the data by selecting the downloaded data file in your local drive.
- ii) Step 2: Load the treatment labels from the downloaded 'labels_3treatments' file by copying and pasting it into the box under 'Step 2' on the side bar.
- iii) Check if the data are uploaded correctly in the 'View data' tab and answer the following question:

Question: Most studies only have 2-arms, but three studies have 3-arms. Which are the 3-arm studies?

Answer:

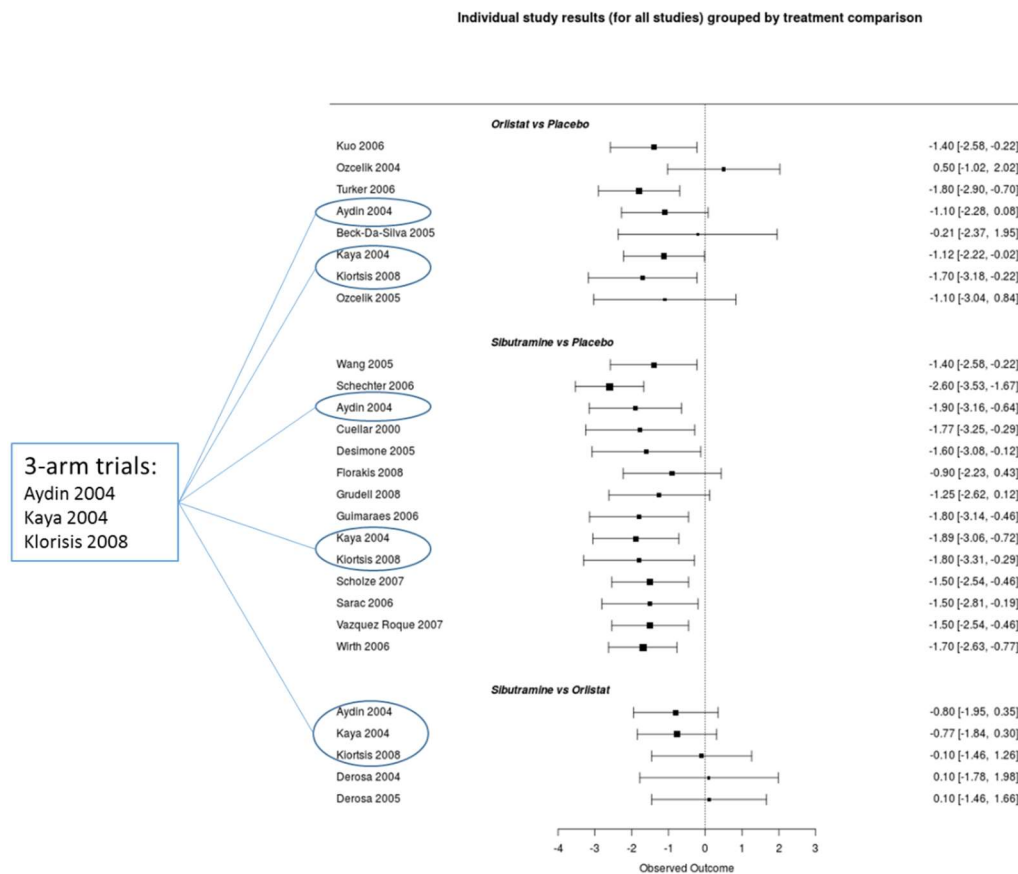
Aydin 2004, Kaya 2004, Kiortsis 2008

- iv) Click on 'Data analysis' at the top of the page and you will see two tabs under this page:
 1. Data summary, which contains:
 - 1a. Study Results, and
 - 1b. Network Plot, and
 2. Frequentist network meta-analysis, which contains:

- 2a. Forest plot, and
- 2b. Comparison of all treatment pairs, and
- 2c. Inconsistency

Next, we will explain what these tabs are for.

- v) '1a. Study Results' contains a plot of the individual study results grouped by treatment comparisons. Observe how the 3-arm studies appear in each of the different treatment comparisons.



Question: What conclusions would you make from observing this plot?

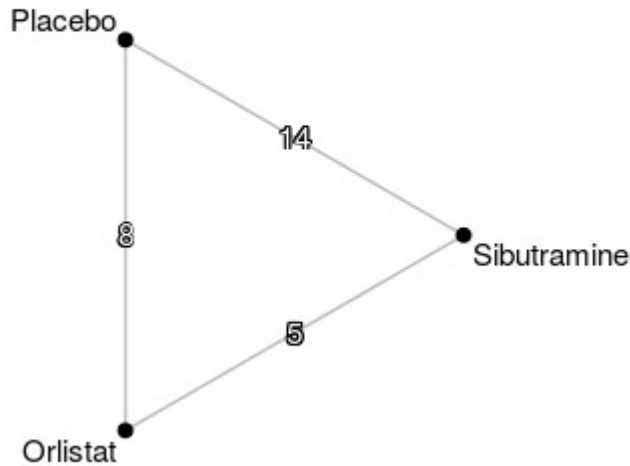
Answer:

All studies consistently favoured Sibutramine to be more effective than Placebo, and the majority of studies favoured Orlistat to be more effective than Placebo with varying degrees of uncertainty. Two out of the five studies favoured Sibutramine when compared to Orlistat and the rest showed no difference in effectiveness between the two active treatments.

- vi) Click on the '1b. Network Plot' tab and you will see a network diagram displaying the 3 treatments included in the analysis together with the numbers along the lines indicating the

number of trials comparing the treatments defined at either end of the line. For example, there are 5 trials comparing Orlistat and Sibutramine, which agrees with the plot of individual study results presented on the '1a. Study Results' tab.

Network plot of all studies



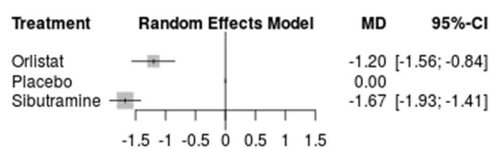
vii) The results of the frequentist network Meta-analysis can be observed by first clicking on the '2. Frequentist network meta-analysis' tab followed by the '2a. Forest plot' tab where the results are presented for each treatment compared to Placebo.

Question: Interpret the results

Answer:

Both Orlistat and Sibutramine are more effective at reducing weight than Placebo at the 5% significance level (i.e. the 95% confidence intervals are below zero).

Results for all studies



viii) The next tab, '2b. Comparison of all treatment pairs', contains three aspects of information:
 a. Treatment ranking: the best treatment is at the top left corner of the diagonal while the worst treatment is at the bottom right corner.

- b. Above the leading diagonal are the relative treatment effects from pairwise estimate (direct evidence only)
- c. Below the leading diagonal are estimates from network meta-analysis (both direct and indirect evidence)

Question: What is the relative effectiveness between Sibutramine and Orlistat?

Answer:

The network estimate suggests that Sibutramine is more effective than Orlistat at reducing weight at the 5% significance level.

Relative treatment effects in ranked order for all studies

Sibutramine	-0.48 [-0.98; 0.02]	-1.70 [-1.96; -1.43]
-0.47 [-0.85; -0.09]	Orlistat	-1.19 [-1.59; -0.79]
-1.67 [-1.93; -1.41]	-1.20 [-1.56; -0.84]	Placebo

- ix) The '2c. Inconsistency' tab contains the results of assessment of inconsistency for all studies. It shows the estimates from network (direct and indirect combined), direct, and indirect evidence. It also shows the difference between the direct and indirect estimate, its 95% confidence intervals, and p value.

Question: Is there any evidence for inconsistency?

Answer:

The estimates from the direct and indirect are consistent – we accept the null hypothesis of consistency.

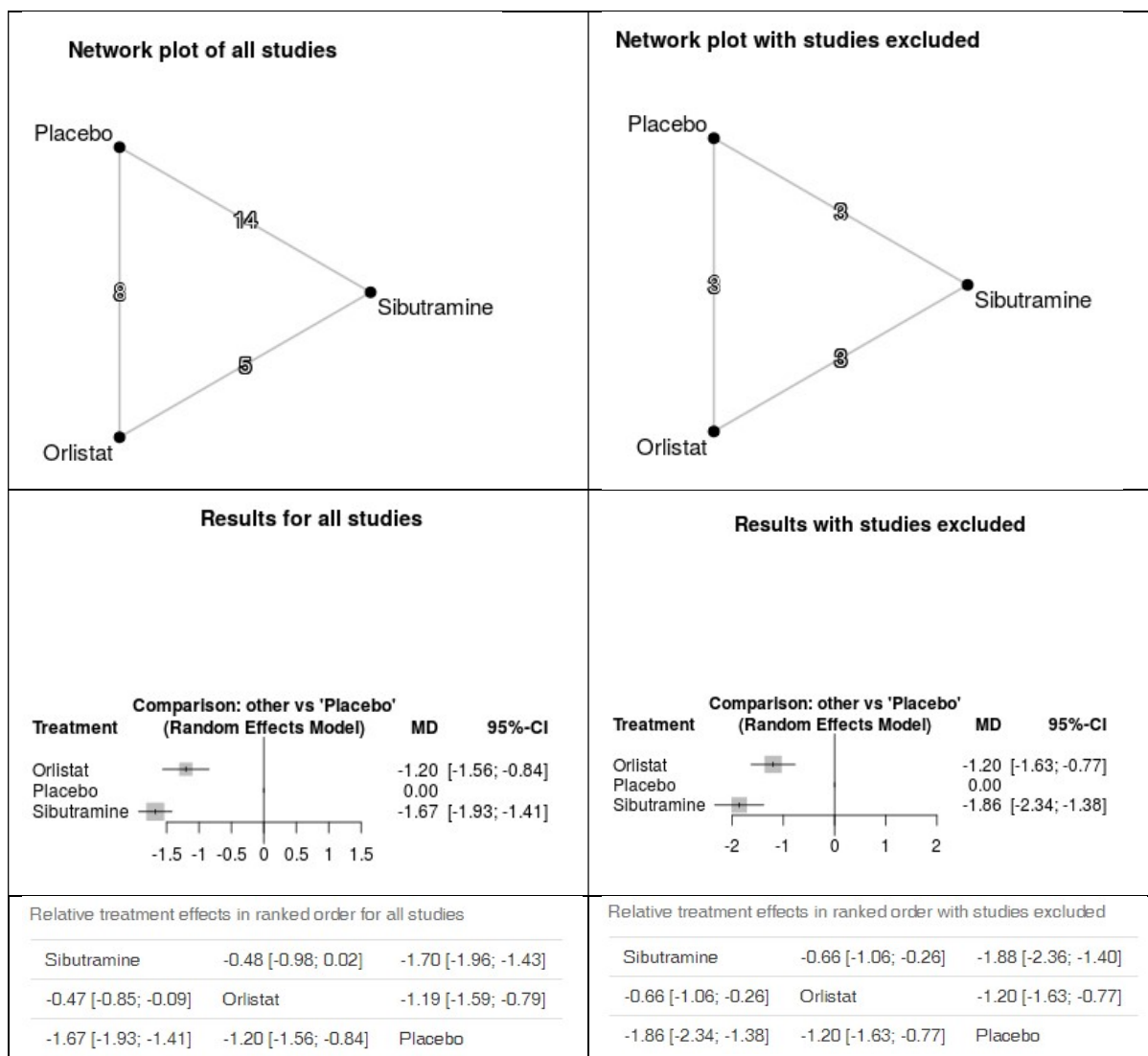
- x) It is easy to exclude studies in *MetaInsight* and to examine the influence this has on the overall results of the network meta-analysis. On the left-hand-side of the screen is a list of all the studies loaded into the App with radial buttons to select those studies you may want to exclude in a sensitivity analysis.

Question: Keeping only the 3-arm trials (i.e. those studies comparing the 3 treatments of interest directly), observe how the results presented on the following tabs change: 'Network Plot', 'Forest Plot', and 'Comparison of all treatment pairs'.

Answer:

Despite removing 18 of the 21 studies from the analysis, the results are similar leading to similar conclusions regarding the most effective treatment.

Including all studies	Including only 3-arm trials
-----------------------	-----------------------------



PART TWO:

In Part One, we kept the analysis simple by including only comparisons between the main three treatments of interest – Orlistat, Sibutramine and Placebo. We are now going to include:

- Further arms for two of the studies included in the simplified analysis that administered different treatments
 - Further studies that compare additional treatments that connect to the network
- i) Use the full data, named 'Data_full.csv', and the treatment label file, named 'labels_full.txt', contain the complete data and the complete list of treatment labels, respectively.
 - ii) Upload the new dataset into *Metalnsight* (as outlined in Part One) and copy and paste the complete list of treatment labels.
 - iii) Note the new comparisons are included in the '1a. Study Results' tab and observe the new network displayed in the '1b. Network Plot' tab. In particular, it is important to confirm that all the new treatments are connected to the network.

Question: Interpret the results on the '2a. Forest Plot' tab. Is a different treatment now estimated to be the most effective?

Answer:

All active treatments are estimated to be more effective than Placebo. Rimonbant is estimated to be the most effective treatment despite the relatively small amount of evidence (1 trial, see the '1b. Network Plot' tab) and correspondingly uncertain treatment effect estimate. Caution is required when interpreting ranks ('2b. Comparison of all treatment pairs' tab) as it does not take uncertainty into account.

PART THREE:

In Part one & two, we used the frequentist version of MetaInsight (V1.1) to conduct NMA and interpret the results. In this part, we will use the version that includes Bayesian analysis (V2.0) to conduct NMA with a Bayesian approach. Within the Bayesian app, NMA is conducted with the R package GEMTC that itself calls the Bayesian simulation analysis software JAGS (<http://mcmc-jags.sourceforge.net/>).

We will use the same full dataset as in part two for the Bayesian app practical, which is also the default dataset preloaded in the app. The Bayesian MetaInsight (continuous) can be accessed from: https://crsu.shinyapps.io/metainsight_continuous2/

Under the 'Data analysis' tab, click on '3. Bayesian network meta-analysis'. Under this tab, you can see the following tabs with their functions summarized in the table below:

3a. Forest plot	Forest plot of the Bayesian NMA results (each treatment compared to the reference treatment), and model fit statistics
3b. Comparison of all treatment pairs	Estimates of comparison of all treatment pairs from the NMA
3c. Ranking table	Ranking table and median rank chart for each treatment to be the best
3d. Nodesplit model	Inconsistency tests with a "node-splitting" model (assesses whether the direct and indirect evidence is inconsistent)
3e. Bayesian result details	Direct result output from R including mean, SD and quantiles from the iterations for each treatment effect, and Gelman convergence assessment plot for checking model convergence.
3f. Deviance report	Three plots are provided for checking model fit of individual data points: a. residual deviance from NMA model and UME inconsistency model, b. per-arm residual deviance for all studies, c. leverage plot. Deviance data output are displayed at the end of the page. Studies with outlying results can be identified using these tools
3g. Model details	3g.1 – model codes 3g.2 – initial values: initial values used in each of the four chains for the analysis. 3g.3 – Download simulations: simulated result data are available to download (e.g. to import into other packages).

Now we will illustrate these functions below.

1. Let's start by looking at the '3a. Forest plot' page.

The Bayesian model takes longer to run than the frequentist model. To avoid the page constantly updating, we include a button so whenever the selection of options on the side bar are finalised, the button should be clicked to tell the app to start running the analysis. For this example, before we run the analysis, we need to make sure the radio button of 'for treatment rankings, smaller outcome values are' set as 'desirable', since the lower BMI estimates the intervention is more effective.

Now click the 'Click here to run the main analysis for all studies' button. You will see a progress bar shown at the right bottom corner indicating the model is running. Depending on the complexity of the model and the speed of your computer, the results usually take about 10 – 20 seconds to run, but can be up to 1 min, so – please be patient and please do not re-click the button if the model is still running.

For this example, the results should be shown in about 6 seconds.

Question: compare the results on this page with the frequentist result on '2. Frequentist network meta-analysis – 2a. forest plot' page. What do you find?

Answer:

The mean estimate and the 95%CI are very similar when comparing the frequentist and Bayesian results. The between-study standard deviations are similar as well. This means, for this dataset, the frequentist and Bayesian approach agree with each other to a high degree.

2. Click on '3b. Comparison of all treatment pairs' and compare the results with the frequentist result. Please note that the all comparison table of Bayesian NMA is not ordered based on the ranking of the treatment effect. So you may need to take some time to match the pairs in both tables to be able to compare.

Also, in contrast to the 'comparison of all treatment pairs' tab in the frequentist NMA results, this table only contains the estimates from the network meta-analysis, i.e. does not contain estimates from pairwise meta-analysis (i.e. direct evidence comparison).

Question: how do the treatment effect estimates compare between frequentist and Bayesian NMA results?

Answer:

The other treatment effect estimates from the two approaches are also similar.

3. Click the '3c. Ranking table', and now you will see the probability for each treatment to rank 1, 2, ... and 6. On the right, you will also see a rank chart based on the median rank of each treatment.

(N.B.: this is an output that is not available from the frequentist analysis)

Question: please interpret the ranking.

Answer:

Rimonabant has the highest probability to be the most effective treatment (95% of the simulations ranked sibutramine the first place), followed by Metformin and Orli_Sibut, then Sibutramine, Orlistat, and placebo at the last.

Please note that different from the frequentist app where the model could be re-run at any stage, the Bayesian model can only run by clicking the button on the '3a. forest plot' page. If you want to change any radio button selections at other pages (e.g. 3c. ranking table), you will need to always go back to the '3a. forest plot' page to click the button to re-run. This applies to the sensitivity analysis as well. This means if you change the exclusion selection, you would need to go back to the '3a. forest plot' page to click the button in order to re-run the sensitivity analysis.

Compared with Placebo	Mean Difference (95% CrI)
Metformin	-2.04 (-2.92, -1.06)
Orli_Sibut	-2.04 (-2.90, -1.09)
Orlistat	-1.10 (-1.48, -0.669)
Rimonabant	-3.77 (-5.53, -1.93)
Sibutramine	-1.85 (-1.83, -1.32)

4. Click the '3d. Nodesplit model' page. Here on this page, you can run the nodesplit models to check the consistency between direct and indirect evidence. It takes a long time to run (and the app may run out of memory and drop offline halfway when many people are running them at the same time). So let's only run a simpler model in this practical by excluding some studies. We would like to examine the consistency between the direct evidence and indirect evidence for the more recent studies (2005 onwards). On the list of 'select studies to exclude', please select all the studies that are published before 2004, including 2004 (as shown on the figure on the right).

Click the button 'Click here to run the nodesplitting analysis with studies excluded'. Then you should get the result below (please note that the result may slightly vary because of the Bayesian analysis):

Select studies to exclude:

- Kuo 2006
- Ozcelik 2004
- Turker 2006
- Wang 2005
- Schechter 2006
- Aydin 2004
- Beck-Da-Silva 2005
- Borges 2007
- Cuellar 2000
- Desimone 2005
- Derosa 2004
- Derosa 2005
- Florakis 2008
- Grudell 2008
- Guimaraes 2006
- Kaya 2004
- Kiortsis 2008
- Ozcelik 2005
- Scholze 2007
- Sarac 2006
- Sari 2004
- Sathuapalan 2008
- Vazquez Roque 2007
- Wirth 2006

NB: If a whole treatment is removed from the analysis the NMA will return an error message. To overcome this, please remove the treatment from the data.

Inconsistency test with nodesplitting model with studies excluded

[Click here to run the nodesplitting analysis with studies excluded](#)

comparison	p.value	CrI
d.Orlistat.Placebo	0.766750	NA
-> direct	NA	1.5 (0.89, 2.1)
-> indirect	NA	1.8 (0.22, 3.2)
-> network	NA	1.5 (1.0, 2.1)
d.Orlistat.Sibutramine	0.784200	NA
-> direct	NA	-0.013 (-0.98, 0.96)
-> indirect	NA	-0.18 (-0.94, 0.58)
-> network	NA	-0.11 (-0.67, 0.46)
d.Placebo.Sibutramine	0.729575	NA
-> direct	NA	-1.7 (-2.0, -1.3)
-> indirect	NA	-1.4 (-3.0, 0.30)
-> network	NA	-1.7 (-1.9, -1.3)

[Download](#)

Question: please interpret the nodesplitting model results

Answer:

There are three comparisons for which both direct and indirect evidence exist (as defined by closed loops in the network diagram). The estimates from direct and indirect evidence show little evidence of inconsistency (as indicated by large p-values).

- Click the '3e. Bayesian result details' page. This page contains the direct R output of running the GEMTC package on the data. The current version of the app uses default setting for the number of simulated iterations for burn-in (5000) and for the model results (20000). We hope to allow more flexibility in the future so that the users can adjust the burn-in period and model results period based on the Gelman convergence assessment plot (although sensible defaults are chosen and work well for this practical).

Question: please compare the results details and the forest plot on '3a. forest plot' page and identify the estimates that are used in the forest plot.

Answer:

The 50% is the point estimate and 2.5% and 97.5% constitute the 95% Credible interval.

- Click the '3f. deviance report' tab. This page presents three diagnostic plots of the model fit in terms of the individual data points.

Question: please try to interpret the plots based on the descriptions underneath each plot.

Answer: The top figure plots the residual deviance from NMA model (horizontal axis) and the unrelated mean effect (ume) inconsistency models (vertical axis) along with the line of

equality. The points on the equality line mean there is no improvement in model fit when using the inconsistency model, suggesting that there is no evidence of inconsistency. There are three points with larger residual deviance (>2), of which two of them are on the line indicating that the NMA model and une inconsistency model provide similar fit, and the other point (at the very top right corner) is below the line indicating the UME model provides better fit. In the deviance data table at the bottom of the page, dev.ab is the residual deviance for each trial arm from the NMA model. Looking at the highest values we can see that the point in the very top right corner is from the study Borges 2007, and the other two points (with deviance >2) on the line are from Ozcelik 2004.

The 2nd plot is the contribution for each study arm to the residual deviance. Ideally each point should contribute around one, with bigger values indicating poorer fit and thus higher residuals. So for the example there are a few points with higher values, i.e. both arms of Ozcelik 2004, and the 2nd arm of Borges 2007.

So it may be sensible to explore these checking no mistakes with data entry, exploring whether the studies differ from the others in any important way, whether the studies were poor quality. It may be sensible to do a sensitivity analysis excluding them (note taking whole studies out not just individual arms) to assess the impact they have.

The contour plot takes this a step further by simultaneously looking at residual deviance and leverage (which is a measure of influence) simultaneously. The text suggests leverage values outside the contour of 3 (which is a points contribution to the DIC) are considered poorly fitting. In the example two points are (just) outside 3. These are a little harder to identify than the previous plot as it is the combination of leverage and deviance which matter. The points leverage values are not extreme but it is the (square root) of the residual deviance which are - i.e. the 2 highest values - so it is mostly likely one arm of Ozcelik and one arm of Borges.

We are hoping in the next version of the app, the feature of 'hover over' interaction will be added to make it easier to identify the points.

7. Click the '3g. Model details'. This page contains the information behind all the summary output, including model codes and initial values and you can download the simulated output data. These are primarily included to add transparency to the app and provide code that ensures any analysis is reproducible (good practice for data analysis).

(End of the practical)