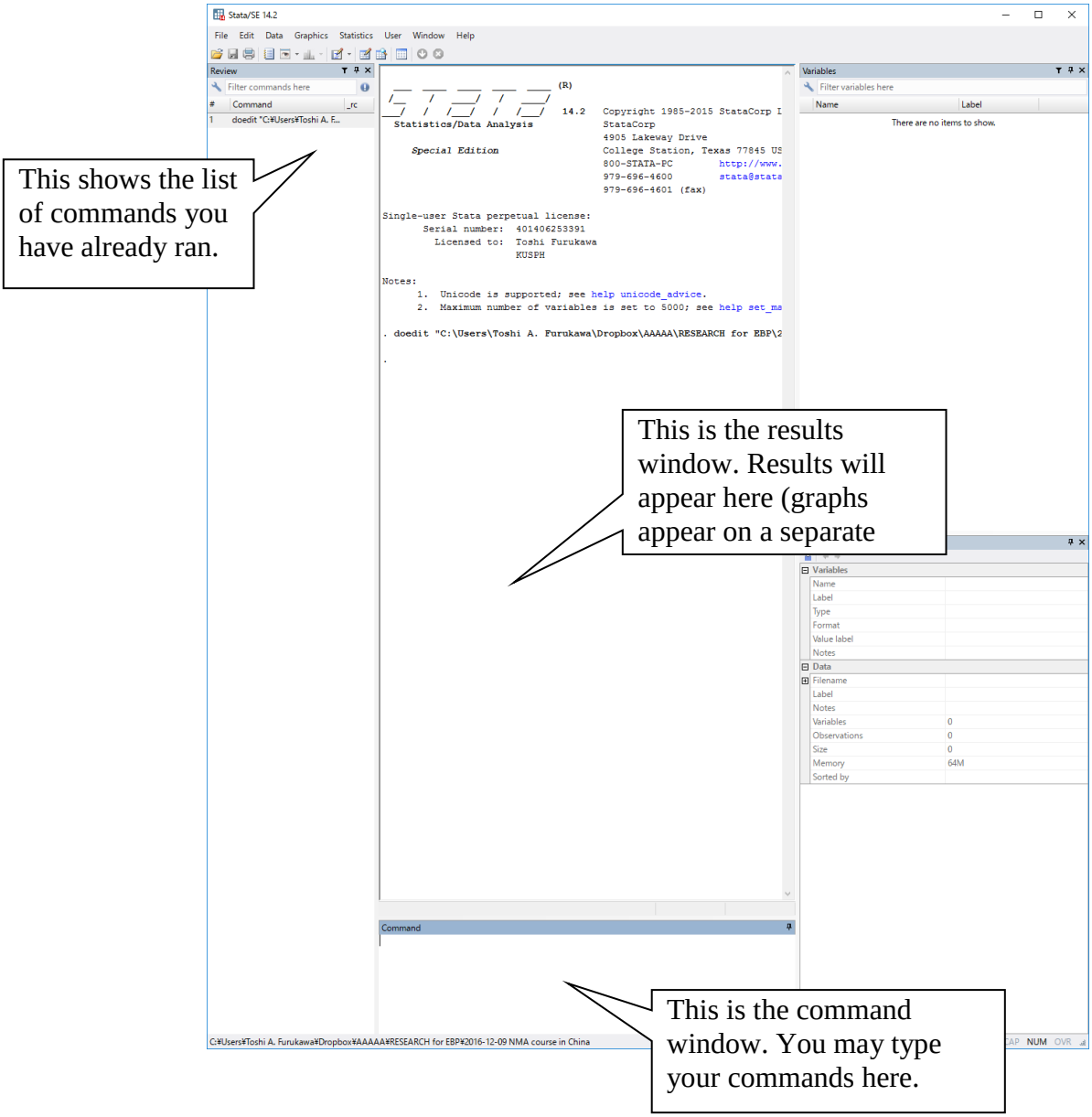
What to do DURING the workshop

To avoid errors due to incompatible fonts (when copy-pasting from .pdf to Stata) or typos (when typing in Stata) all commands can be run from the provided .do file.

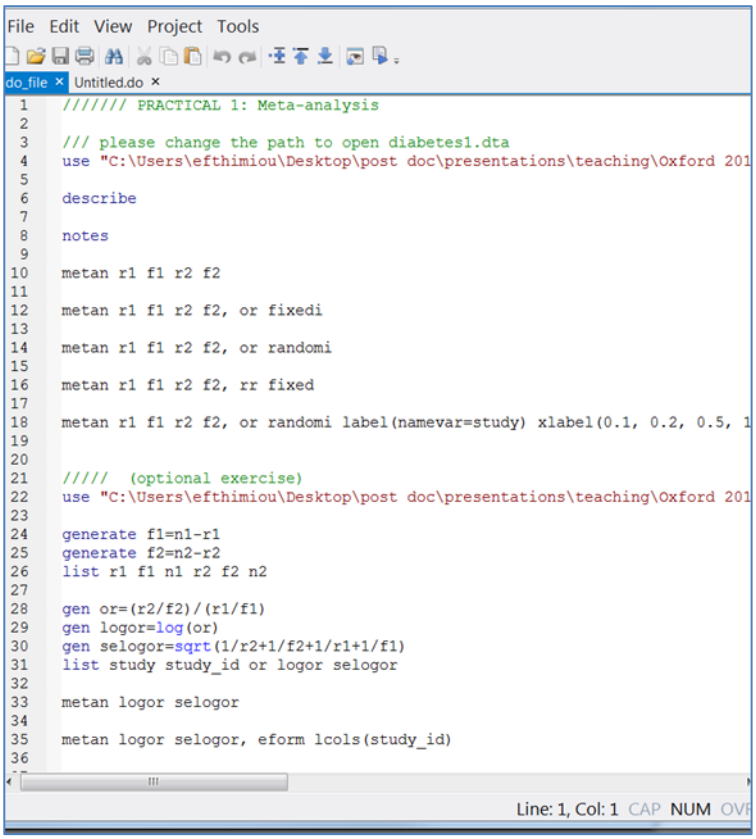
1 Setting up STATA for today's session

Let's first open the file `do_file.do` (by double-clinking it).

You will see two windows. The first is the main Stata window:



The second is a so-called .do file, which contains the commands with some additional comments (text in green, after the /////).



Please note that commands (in black) usually has the following structure:

```
. XXX YYY, ZZZ
```

where XXX is the command name, YYY is a list of variables, and ZZZ is a list of options. Some commands do not have YYY and/or ZZZ.

You can select one or more commands in the .do file, and then you can use either Ctrl-D or click on the icon located at the right-end of the top bar to run the commands.

If you want to know more about a specific command in Stata, you can always run:

```
. help XXX
```

to call for a window explain all options available and with examples.

Practical 1: Meta-analysis

The aim of this practical is to introduce the **metan** command.

1 Diuretics for pre-eclampsia (meta-analysis)

Dataset **diuretics.dta** contains data from nine trials of diuretics for pre-eclampsia (dichotomous outcome). Open the dataset in Stata (either by double clicking the file in Windows to open it in a new Stata session, or by first starting Stata and then opening the file using the menus or the **use** command).

Use the **describe** and **notes** commands to obtain details of the dataset.

We can perform a meta-analysis using the default methods of the **metan** command as follows:

```
metan r1 f1 r2 f2
```

The default method is a fixed-effect analysis on the risk ratio scale using the Mantel-Haenszel method.

Now type:

```
metan r1 f1 r2 f2, or fixedi
```

The **or** argument species an analysis on the odds ratio scale and **fixedi** requests a fixed-effect analysis using the standard inverse-variance method.

The random-effects meta-analysis is produced as follows:

```
metan r1 f1 r2 f2, or randomi
```

The following table lists the options available for methods of meta-analysis of dichotomous outcome data. The option **nograph** suppresses the forest plot, so makes the analysis quicker to run.

or	odds ratio
rr	risk ratio
rd	risk difference
fixed	Mantel-Haenszel fixed-effect
random	Mantel-Haenszel random-effects
fixedi	Inverse-variance fixed effect
randomi	Inverse-variance random-effects

The default method is equivalent to

```
metan r1 f1 r2 f2, rr fixed
```

Try a few different methods.

Do the methods affect the conclusions you would draw?

Use **help metan** to look at the options available for the command. You can add information on your forest plot, e.g. like:

```
metan r1 f1 r2 f2, or randomi label(namevar=study) xlabel(0.1, 0.2, 0.5, 1, 2, 4) effect(Odds ratio) favours(favours diuretic # favours control) texts(180)
```

2 Meta-analysis for calcium channel blockers versus beta blockers: more about metan (OPTIONAL)

Dataset **diabetes1.dta** contains data from five trials comparing two antihypertensive treatments, calcium channel blockers (CCB) and beta-blockers (BB), measuring the onset of diabetes as an outcome. Use the **describe** and **notes** commands to obtain details of the dataset.

We will start by using the **generate** command to calculate the log odds ratio and its standard error for each study. First, derive the number of individuals who did *not* develop diabetes in the two treatment arms:

```
generate f1=n1-r1
generate f2=n2-r2
list r1 f1 n1 r2 f2 n2
```

Note, **r1 f1 n1** are data for BB (treatment 1); **r2 f2 n2** for CCB (treatment 2).

Now derive the odds ratio, then the log odds ratio and its corresponding standard error for each study. We will take the ratio of diabetes odds on CCB over diabetes odds on BB (odds of diabetes for treatment 2 over odds of diabetes for treatment 1). Formulas are available in the appendix of the slides "Introduction to meta-analysis, indirect comparisons and mixed treatment comparisons".

```
gen or=(r2/f2)/(r1/f1)
gen logor=log(or)
gen selogor=sqrt(1/r2+1/f2+1/r1+1/f1)
list study study_id or logor selogor
```

We will now use the **metan** command to perform a fixed- and then a random-effects meta-analysis, using inverse-variance weighting. The basic output for fixed-effects is produced by typing:

```
metan logor selogor
```

The **metan** command has many options to improve the output. To display the output on the odds ratio scale, and to see the study IDs, type:

```
metan logor selogor, eform lcols(study_id)
```

To produce a random-effects analysis:

```
metan logor selogor, eform lcols(study_id) random
```

Complete the following table of results from the random-effects meta-analysis, where we use notation μ (logOR) and τ^2 from the lectures.

Table 1

Parameter	***Estimate***	***95% confidence interval***
μ CCB vs BB		
OR CCB vs BB		
τ^2 CCB vs BB		N/A

The analysis can be performed directly from the binary data as follows:

```
metan r2 f2 r1 f1, lcols(study_id) or randomi
```

The **or** option specifies use of odds ratios (alternatives are **rr** and **rd** for risk ratios and risk differences, respectively), and the **randomi** option specifies a random-effects analysis using a standard inverse-variance weighted average. To suppress the forest plot, add the option **nograph**. Check that the results agree with the analysis above.
The **metan** command stores the estimated treatment effects and the standard errors of the log effect for each study in the automatically created variables **_ES** (effect size) and **_selogES**. For example, if we use the **metan** command with odds ratios then **_ES** is the odds ratio (identical to the variable or calculated at the beginning of the practical) and the value of **_selogES** is the standard error of the log odds ratio (identical to variable **selogor**, also calculated earlier).

Note: Some Stata meta-analysis commands require the user to provide the log odds ratio and its standard error; the **metan** command is useful for deriving these variables, as it is quicker than computing them directly.

```
metan r2 f2 r1 f1, lcols(study_id) or randomi nograph
list study or _ES selogor _selogES
```

In addition, **metan** stores the results of the meta-analysis. We can see these using

```
return list
```

Use the following commands to complete the following table (**g** is shorthand for **generate**):

```
g mu = log(r(ES))
display mu
display r(selogES)
display r(ES)
display r(ci_low)
display r(ci_upp)
display r(tau2)
```

Table 2

Parameter	***Estimate***	***Uncertainty***
μ CCB vs BB		Standard error:
OR CCB vs BB		95% Confidence interval:
τ^2 CCB vs BB		N/A

Note: An alternative way to obtain results on the log odds ratio scale is to use the option **`log`** in the **`metan`** command.

Practical 2: Fitting the network meta-analysis model

In this practical we will perform network meta-analysis with multivariate meta-analysis (with **`network`**).

3 Network meta-analysis using multivariate meta-analysis

We will use the **`network`** package that calls the **`mvmeta`** command, so that we can properly take account of the multi-arm studies. The package does everything 'automatically', but make sure you understand each step so that you know what is going on.

We will use a different data set that has been formatted more appropriately for **`network`**. Open **`diabetes2.dta`**.
The treatments codes are:
1 = placebo, 2 = beta blockers (BB), 3 = diuretics, 4 = calcium channel blockers (CCB), 5 = ACE inhibitors, 6 = angiotensin-receptor blockers (ARB).

Look at part of the data, e.g.

```
list study t r n
```

What does each row in the data set represent?

We now need to create the treatment effect estimates and a variance-covariance matrix for the observations. Recall that the multivariate approach requires us to have data for the reference treatment in every study.

```
network setup r n, stud(study) trt(t) ref(1) or numcodes
```

Look again at the data

```
edit
```

How many rows are there per study?

What does each variable **`_y_2`**, **`_y_3`**,…, **`_y_6`** represent?

Network command allows you to switch between the format of the data:

```
network convert pairs
```

Look again at the data. How many rows are there per study? Use the sort command if needed (**`sort study`**)

There is also a third form:

```
network convert standard
```

Now use the map option to get a network graph:

```
network map
```

The graph can be exported from Stata in other programs for editing (eg. PowerPoint)
By typing help network you can also explore additional options for this command.

If you bring the data in the pairs format, you can also use the networkplot command, which allows some additional options

```
network convert pairs
networkplot _t1 _t2, lab(PL BB CCB CCB ACE ARB)
```

To run the network meta-analysis:

```
network convert augmented

network meta c
```

By default, the command uses the common assumption of equal heterogeneities for all comparisons.

Which treatments would appear to be best, and worst (versus placebo)? Note that r in the data are incidence of diabetes and that the *Coef*. reported in STATA is the logOR of each treatment versus placebo.

Compute ('by hand') the odds ratio comparing ACE (treatment 5) with BB (treatment 2).

We can use the **`lincom`** command to compare specific treatments. To compare ACE (treatment 5) with BB (treatment 2):

```
lincom [_y_5]_cons -[ _y_2]_cons, eform
```

To derive all summary odds ratios for all treatment comparisons based on the consistency equations, run the following loop.

```
foreach trt1 in 2 3 4 5 6{
 foreach trt2 in 2 3 4 5 6{
```

```
if "`trt1'"=="`trt2'" continue
if "`trt2'">"`trt1'" lincom [_y_`trt2']_cons-[_y_`trt1']_cons,eform
  }
 }
```

As a reminder:
1 = placebo, 2 = beta blockers (BB), 3 = diuretics, 4 = calcium channel blockers (CCB), 5 = ACE inhibitors, 6 = angiotensin-receptor blockers (ARB).

Extract these results:

Table 2

Parameter	***Estimate***	***Uncertainty***
$OR_{CCB\ vs\ BB}$		CI:
$OR_{ACE\ vs\ BB}$		CI:
$OR_{ACE\ vs\ CCB}$		CI:
τ^2		-

In this analysis we chose treatment 1 to be the reference. As discussed in the lectures, this choice is arbitrary. Check that redoing the analyses using another treatment does not substantially change results. There might be some small differences due to augmentation technique. In general, for estimation purposes it is good to choose well connected treatments.

4 Forest plot of all pairwise summary effects

We can use the **intervalplot** command to draw a forest plot of all estimated pairwise summary effects. The option **null()** specifies the line of no effect, while the option **separate** separates the different comparisons according to the comparator treatment. The options **range()** and **xlab()** handle the appearance of the horizontal axis and the option **margin()** the blank margins around the plot:

```
intervalplot, mvmeta lab(Placebo BB Diuretics CCB ACE ARB) eform
     null(1) sep  range(0.4 2) xlab(0.4 0.7 1.5 2) marg(5 40 5 5)
```

5 Example with continuous data (OPTIONAL)

Open the data set **glaucoma.dta**. There are 24 studies and eight treatments. The treatment codes are:
1 = placebo, **2 = travoprost**, **3 = timolol**, **4 = betaxolol**, **5 = latanoprost**, **6 = dorzolamide**, **7 = brinzolamide**, **8 = brimonidine**, **9 = bimatoprost**.

```
use "glaucoma.dta", clear
```

Look at part of the data, e.g.

```
list study t mean sd n
```

Prepare the data in the appropriate format for conducting network meta-analysis using the multivariate approach (consider the treatment 1 as reference).

```
network setup mean sd n,stud(study) trt(t) smd ref(1) numcodes
```

Look again at the data

```
edit
```

How many rows are there per study?

What does each variable **_y_2**, **_y_3**,..., **_y_6** represent?

What does each variable **_S_2_2**, **_S_3_3**,..., **_S_6_6** represent?

What does each variable **_S_2_3**, **_S_2_4**,..., **_S_5_6** represent?

Network command allows you to switch between the format of the data:

network convert pairs

Look again at the data. How many rows are there per study? Use the sort command if needed (**sort study**)

Now use the map option to get a network graph:

network map

To run the network meta-analysis:

network convert augmented

network meta c

By default, the command uses the common assumption of equal heterogeneities for all comparisons. That is, we assumed a variance-covariance matrix of the following form for the five basic parameters.

$$\mathbf{T}^2 = \begin{pmatrix} \tau^2 & \frac{\tau^2}{2} & \frac{\tau^2}{2} & \frac{\tau^2}{2} & \frac{\tau^2}{2} \\ \frac{\tau^2}{2} & \tau^2 & \frac{\tau^2}{2} & \frac{\tau^2}{2} & \frac{\tau^2}{2} \\ \frac{\tau^2}{2} & \frac{\tau^2}{2} & \tau^2 & \frac{\tau^2}{2} & \frac{\tau^2}{2} \\ \frac{\tau^2}{2} & \frac{\tau^2}{2} & \frac{\tau^2}{2} & \tau^2 & \frac{\tau^2}{2} \\ \frac{\tau^2}{2} & \frac{\tau^2}{2} & \frac{\tau^2}{2} & \frac{\tau^2}{2} & \tau^2 \end{pmatrix}$$

Which treatments would appear to be best, and worst (versus placebo)?

Compute ('by hand') the SMD comparing treatment 3 with treatment 2.

To derive all summary SMD for all treatment comparisons based on the consistency equations, run the following loop.

```
foreach trt1 in 2 3 4 5 6 7 8 9 {
 foreach trt2 in 2 3 4 5 6 7 8 9 {

if "`trt1'"=="`trt2'" continue
if "`trt2'">"`trt1'" lincom [_y_`trt2']_cons-[_y_`trt1']_cons,eform
```

```
  }
}
```

Extract these results:

Table 2

Parameter	***Estimate***	***Uncertainty***
$OR_{3 \text{ vs } 2}$		CI:
$OR_{4 \text{ vs } 2}$		CI:
$OR_{4 \text{ vs } 3}$		CI:
τ^2		-

Practical 3: Assessing inconsistency in network meta-analysis

6 Loop-specific approach for the evaluation of inconsistency

Open the data set **`diabetes2.dta`**.

The treatment codes are:
1 = placebo, **2 = beta blockers (BB)**, **3 = diuretics**, **4 = calcium channel blockers (CCB)**, **5 = ACE inhibitors, 6 = angiotensin-receptor blockers (ARB)**.

Use the **`network`** package to prepare the data in a format where each row represents a pairwise comparison from a study

```
network setup r n, stud(study) trt(t) ref(1) or

network convert pairs
```

Use the **`ifplot`** command to draw a forest plot of all inconsistency factors. The option **`eform`** plots the ratio of odds ratios (*ROR*) between direct and indirect estimates, which is estimated as *ROR=exp(IF)*. We can use the option **`plotoptions()`** to handle the appearance of the plot using standard options of the **`metan`** command:

```
ifplot _y _stderr _t1 _t2 study, eform lab(P BB D CCB ACE ARB)
plotopt(texts(120))
```

ROR stands for the ratio of odds ratios between direct and indirect evidence.

How many closed loops are included in the network?

How many loops display significant inconsistency in the network?

Are there other loops that might be potential sources of inconsistency in the network?

The default setting for **`ifplot`** is to assume a common heterogeneity variance for all comparisons in a loop. Run the command again allowing this to differ for each comparison in a loop (comparison-specific heterogeneity estimates):

```
ifplot _y _stderr _t1 _t2 study, eform lab(P BB D CCB ACE ARB) tau2(comparison)
```

Do the results change?

Run the command using the common between-study variance estimated by the network meta-analysis model for all loops from the previous practical (i.e. set $\tau^2 = 0.11682021^2$):

```
di  .1168^2

ifplot _y _stderr _t1 _t2 study, eform lab(P BB D CCB ACE ARB) tau2(0.014)
```

Do the results change?

7 Assessing inconsistency using the node-splitting approach

We can assess inconsistency using the node-splitting approach. In order to split comparison A-B type:

```
network convert standard

network sidesplit A B
```

Is there a difference between direct and indirect evidence for this comparison?

You can use network to perform node-splitting to all comparisons in one go:

```
network sidesplit all
```

How many treatment comparisons have some indication for inconsistency?

How do results compare with the loop-specific approach?

8 Example with continuous data (OPTIONAL)

If you have time, open the data set **`glaucoma.dta`**. There are 24 studies and eight treatments plus placebo. The treatment codes are: **1 = placebo**, **2 = travoprost**, **3 = timolol**, **4 = betaxolol**, **5 = latanoprost**, **6 = dorzolamide**, **7 = brinzolamide**, **8 = brimonidine**, **9 = bimatoprost**.
Prepare the data in the appropriate format for conducting network meta-analysis using the multivariate approach

```
network setup mean sd n,stud(study) trt(t) smd ref(1)
```

and try to assess inconsistency