Dr Silia Vitoratou

Department: Biostatistics and Health Informatics

Topic materials:
Silia Vitoratou

Contributions:
Zahra Abdulla

Improvements:
Nick Beckley-Hoelscher
Kim Goldsmith
Sabine Landau

Institute of Psychiatry, Psychology and Neuroscience

KING'S College LONDON

Module Title: Introduction to Statistics

Session Title: Equality of proportions (exact tests)

Topic title: Comparing groups II (non-parametric methods)

Learning Outcomes

- Learn when and how to use the **exact** tests for equality of proportions.
- Understand when assumptions have been violated to warrant the use of exact tests.
- Be able to conduct these tests in a statistical software.

Parametric and Non-Parametric Tests

Numerical data	Normality assumed	Normality not assumed
Hypotheses testing	parametric	Non-parametric
one group *versus* a pre-defined value	one sample t-test	Wilcoxon sign rank
one group *versus* another group	two independent samples t-test	Mann-Whitney (Wilcoxon sum rank)
one group (twice or) *versus* another matched group	two paired samples t-test	Wilcoxon sign rank

Previously on 'Introduction to Statistics'.....

Categorical data	Assumptions hold	Assumptions do not hold
Hypotheses testing	Chi-squared tests	Exact tests
one group *versus* a pre-defined value	one sample χ^2-test	Binomial exact test
one group *versus* another group	Pearson's χ^2-test	Fisher's Exact test
one group (twice or) *versus* another matched group	McNemar test	Binomial exact test

Equality of Proportions: Exact Tests

For associations between categorical data, so far we have been considering chi-square (χ^2) tests. However, we can trust the results of the test only if the assumptions hold

Up to 20% cells can have expected count less than 5.
The minimum expected count is larger than 1.

If these assumptions are not satisfied, then test statistic of the chi-square is not reliable

$$\sum \frac{(\text{Observed} - \text{Expected})^2}{\text{Expected}}$$

and we use instead the '**exact**' tests.

One Sample Chi-Square Test - Reminder

When to use

To test if according to the current data, the proportion in the population equals a certain, pre-specified, value.

Hypotheses:

H_0: the proportion in the population equals a certain pre-specified value
H_a: the proportion in the population is different than a certain pre-specified value

Assumptions:

- The observations are randomly and independently drawn
- The number of cells with expected frequencies less than 5, are less than 20%
- The minimum expected frequency is at the very least 1.

SPSS Slide: ‘how to’

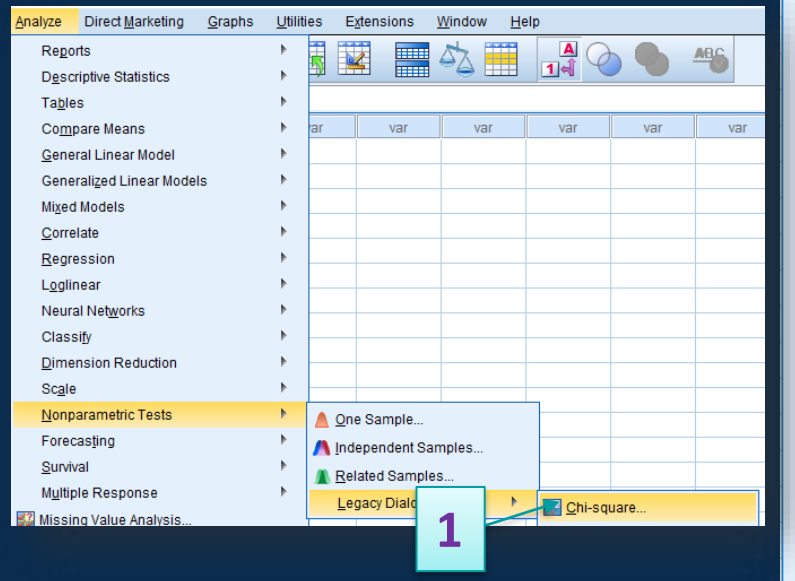
The programme developers had hoped that no more than 1% of the population has used drugs. Is that what our data tell us?

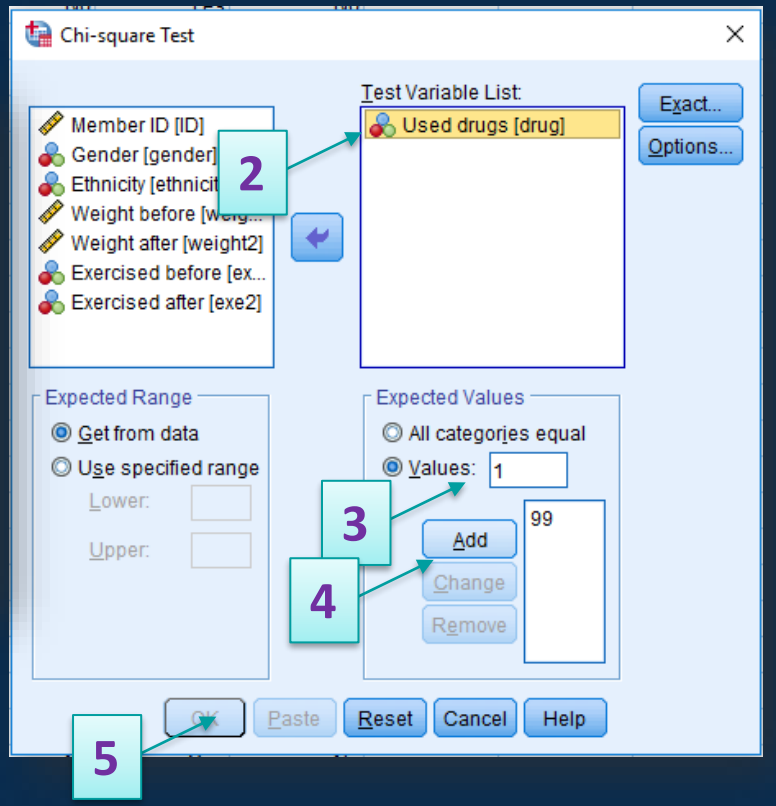
H_0: π=1%

H_a: $\pi \neq$1%

Step 1: Use the appropriate test, here ‘one-sample chi-square t-test’.

Analyse -> nonparametric tests -> legacy dialogs -> ‘Chi-square’





Add the variable of interest (used drugs) in to the ‘Variables box’.
Choose to Test ‘values’

Check the coding in the dataset as ‘0’ coding is ‘not used drugs’ ‘add’ in 99 (100%- 1%) first.
Then as ‘1’ coding is ‘used drugs’ ‘add’ in 1 (to represent the 1% in the question).
Click ‘OK’

Output and Interpretation Slide

Step 2: Check the suitability of the data, do the assumptions of the chi-square test hold?

Used drugs

	Observed N	Expected N	Residual
No	290	297.0	-7.0
Yes	10	3.0	7.0
Total	300		

Test Statistics

	Used drugs
Chi-Square	16.498[a]
df	1
Asymp. Sig.	.000

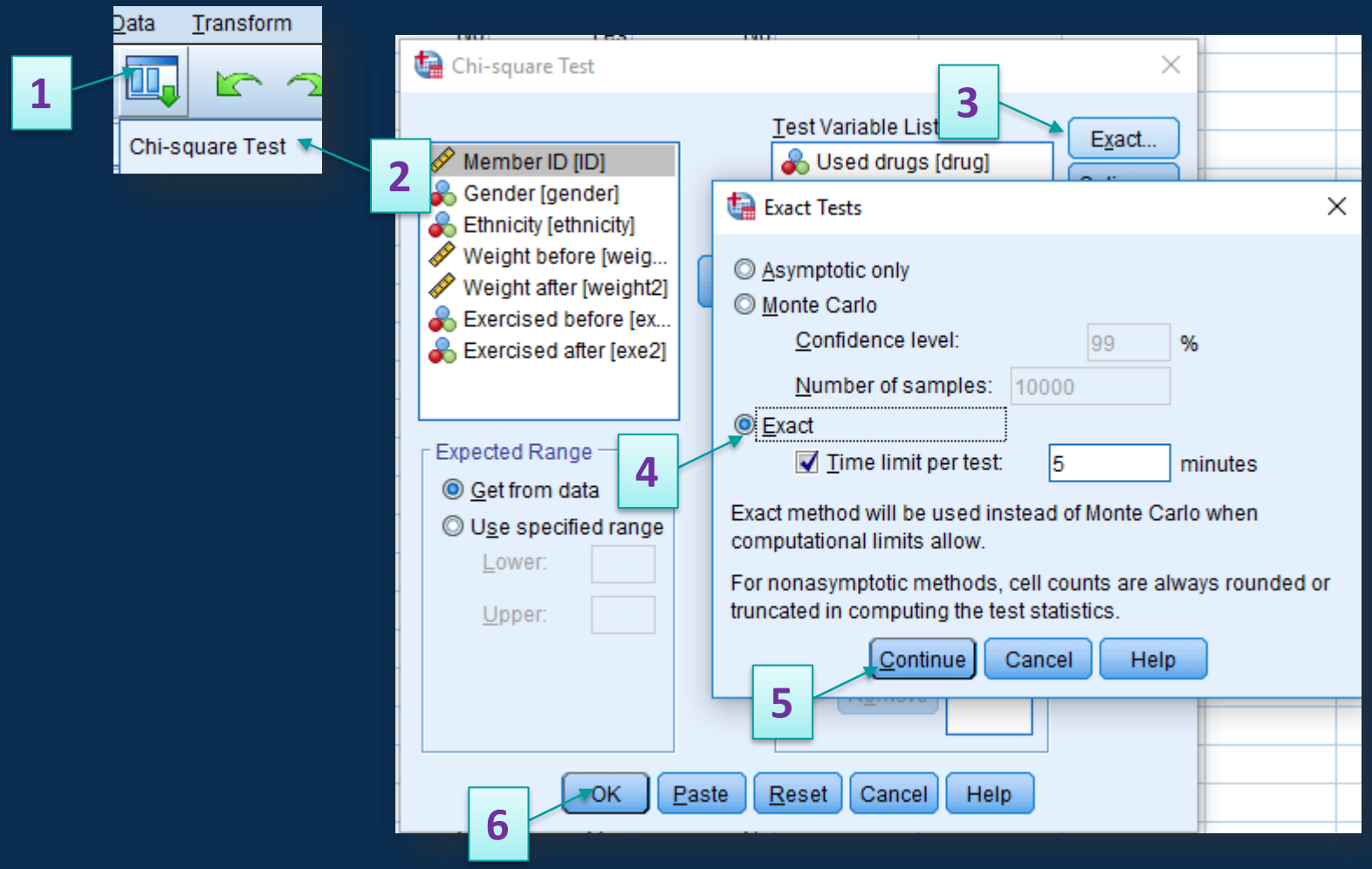
a. 1 cells (50.0%) have expected frequencies less than 5. The minimum expected cell frequency is 3.0.

Only up to 20% of the cells are allowed to have expected frequencies less than 5.

Thus, one assumption is violated!

SPSS Slide: ‘how to’

Use the recall button to go back to the chi-square, and this time click the ‘**exact**’ tab.



Select the ‘Exact’ option
Click ‘Continue’
Click ‘OK’

Output & Interpretation Slide

Used drugs

	Observed N	Expected N	Residual
No	290	297.0	-7.0
Yes	10	3.0	7.0
Total	300		

Test Statistics

	Used drugs
Chi-Square	16.498[a]
df	1
Asymp. Sig.	.000
Exact Sig.	.001
Point Probability	.001

a. 1 cells (50.0%) have expected frequencies less than 5. The minimum expected cell frequency is 3.0.

Based on our sample, the expected proportion of people who have used drugs to lose weight is different (larger) than 1% (exact p-value=0.001).

Pearson's Chi-Square Test

When to use

To test if, according to the current data, the proportions in the population of one variable change based on another variable.

Hypotheses:

H_0: there is no association between the two variables
H_a: there is an association between the two variables

Assumptions:

- The observations are randomly and independently drawn
- The number of cells with expected frequencies less than 5, are less than 20%
- The minimum expected frequency is at the very least 1.
- The observations are not paired

SPSS Slide: ‘how to’

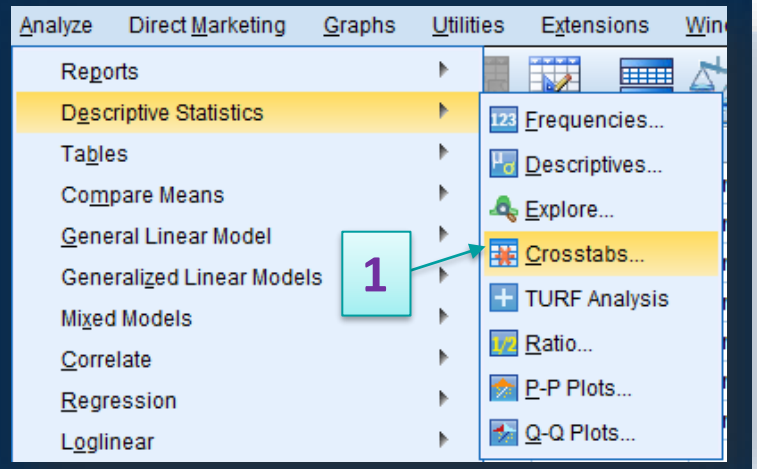
The next question is: do men exercise more than women prior to entering the programme?
Are the proportions of those exercised before the programme, different for men and women?

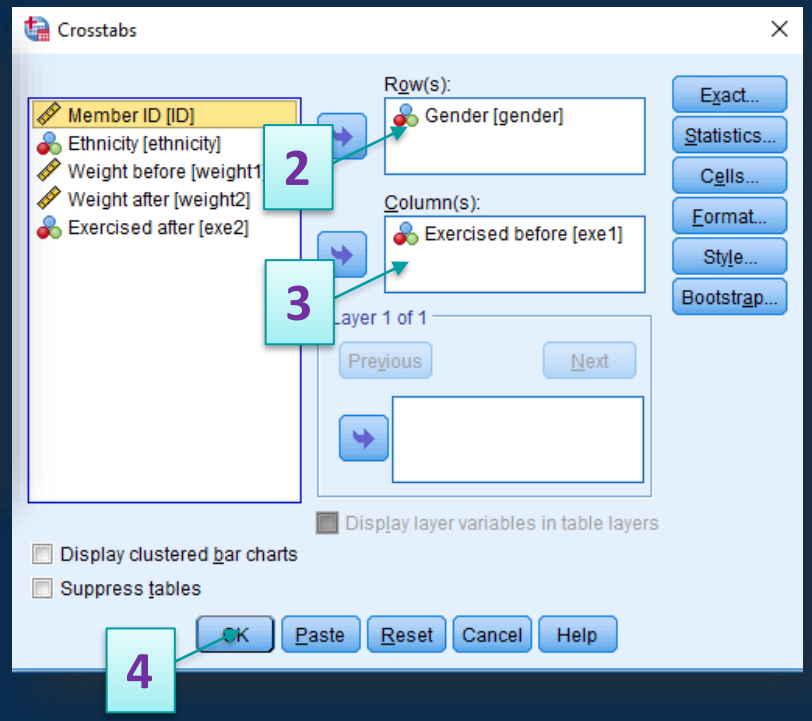
Step 1: Use the appropriate test, here: ‘Pearson’s chi-square test’.

Analyse -> Descriptive Statistics-> Crosstabs

$H_0: \pi_m = \pi_f$

$H_a: \pi_m \neq \pi_f$

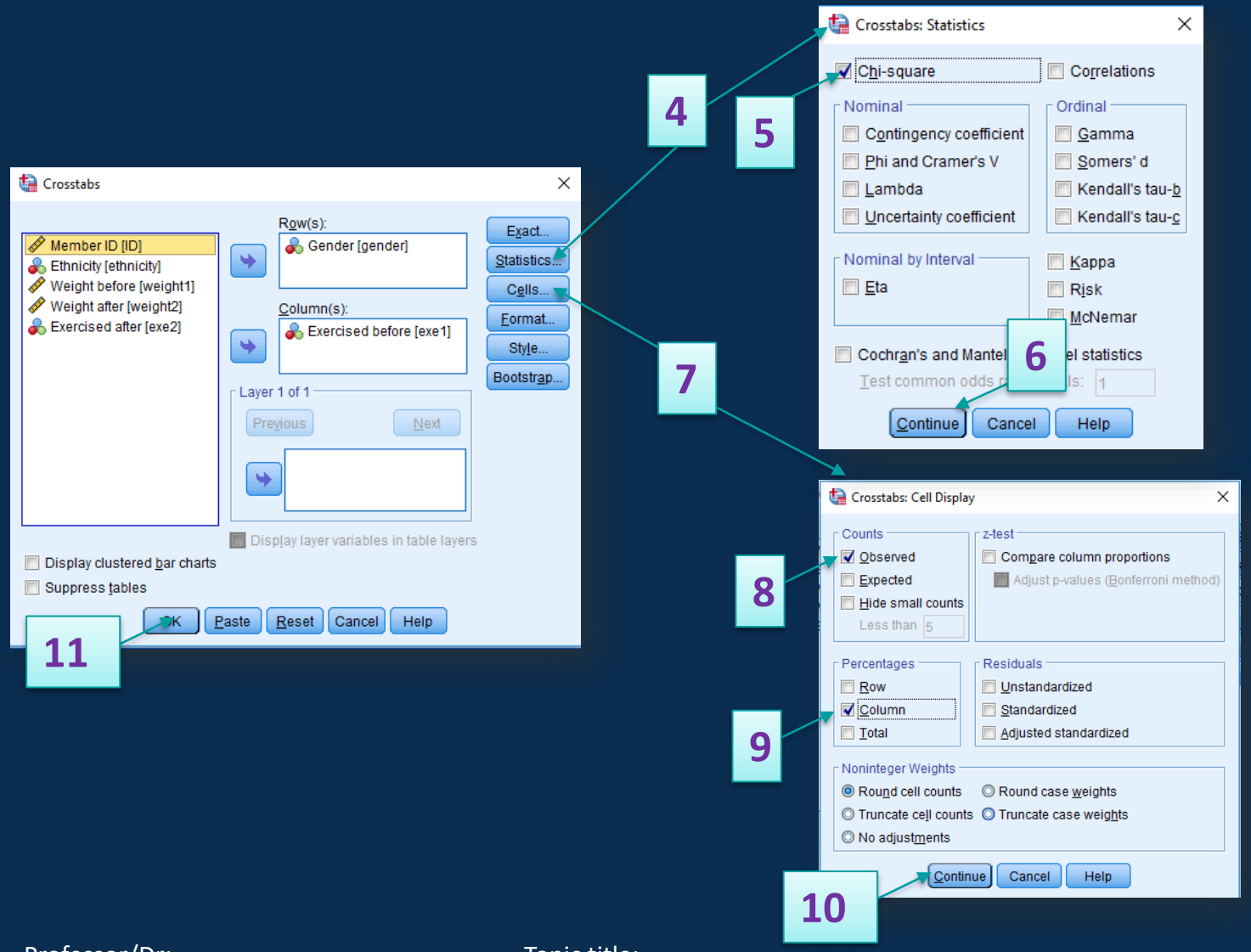




Add the variable of interest (gender) in to the ‘Rows box’.
Add the second variable interest (Exe before) in the ‘columns box’

SPSS Slide: ‘how to’

Step 1: Use the appropriate test, here: ‘Pearson’s chi-square test’.



Output and Interpretation Slide

Unfortunately, due to a technical problem, most of the data we had in 'exercise before' were accidentally deleted. We only have info for 29 people.

Gender * Exercised before Crosstabulation

			Exercised before		
			No	Yes	Total
Gender	Female	Count	17	3	20
		% within Exercised before	100.0%	25.0%	69.0%
	Male	Count	0	9	9
		% within Exercised before	0.0%	75.0%	31.0%
Total		Count	17	12	29
		% within Exercised before	100.0%	100.0%	100.0%

Step 2: Check the suitability of the data, do the assumptions of the chi-square test hold?

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	18.488[a]	1	.000		
Continuity Correction[b]	15.149	1	.000		
Likelihood Ratio	22.428	1	.000		
Fisher's Exact Test				.000	.000
Linear-by-Linear Association	17.850	1	.000		
N of Valid Cases	29				

a. 1 cells (25.0%) have expected count less than 5. The minimum expected count is 3.72.

b. Computed only for a 2x2 table

- Only up to 20% of the cells are allowed to have expected frequencies less than 5.
- **Thus, one assumption is violated!**

Output and Interpretation Slide

Gender * Exercised before Crosstabulation

			Exercised before		
			No	Yes	Total
Gender	Female	Count	17	3	20
		% within Exercised before	100.0%	25.0%	69.0%
	Male	Count	0	9	9
		% within Exercised before	0.0%	75.0%	31.0%
Total		Count	17	12	29
		% within Exercised before	100.0%	100.0%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	18.488[a]	1	.000		
Continuity Correction[b]	15.149	1	.000		
Likelihood Ratio	22.428	1	.000		
Fisher's Exact Test				.000	.000
Linear-by-Linear Association	17.850	1	.000		
N of Valid Cases	29				

a. 1 cells (25.0%) have expected count less than 5. The minimum expected count is 3.72.

b. Computed only for a 2x2 table

Among females, the proportion of those who exercised before the programme was lower than those who did not exercise before the programme (25% versus 100%, respectively). This difference was statistically significant according to Fisher's exact test (exact $p<0.001$).

Therefore, we conclude that men tend to exercise (before the programme) more often than women, in the population. The variables 'gender' and 'exe1' are related.

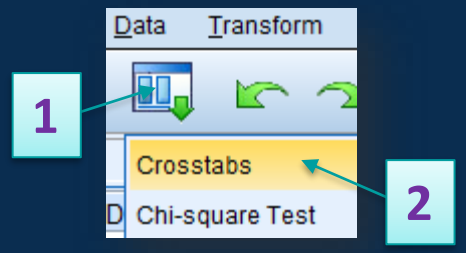
SPSS Slide: ‘how to’

Neither Pearson’s chi-square, nor Fisher’s exact test are restricted to 2x2 tables. We may have 2 groups but more than two categories in the categorical variable.

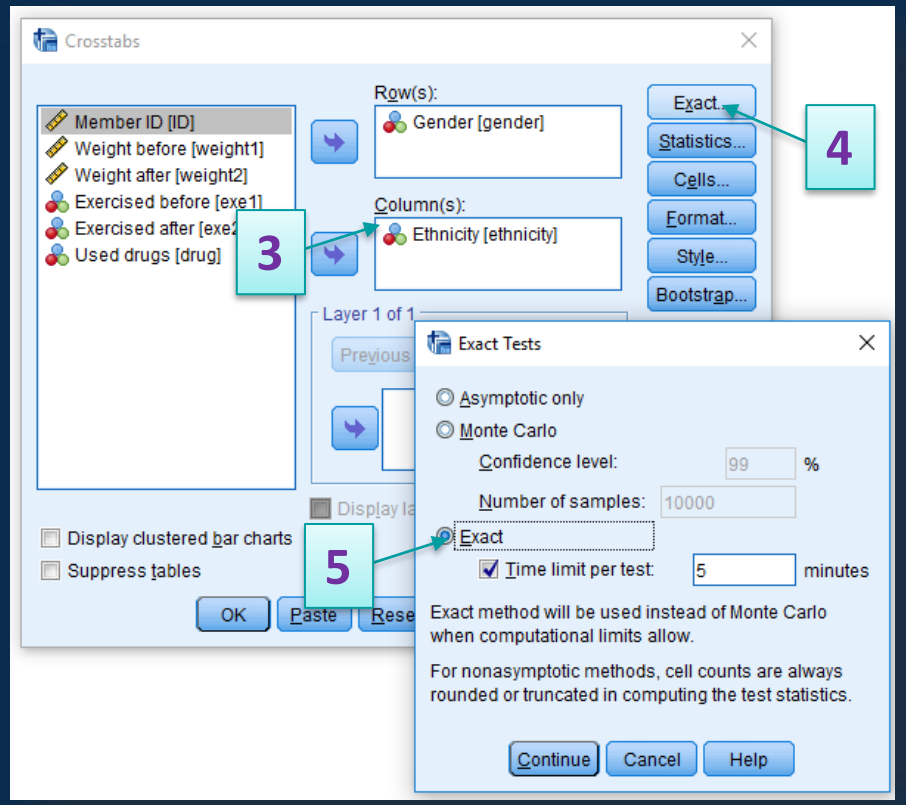
The next question is: is ethnicity (more than 2 categories) associated with gender?

Step 1: Use the appropriate test, here: ‘Pearson’s chi-square test’.

Use the recall button and simply replace ‘exe1’ with ‘ethnicity’



While Fisher’s exact test appears automatically in 2x2 tables, but that is NOT the case in 2xP tables. We need to go the ‘**exact**’ tab (as we did for the one sample previously).



Interpretation Slide

Step 2: Check the suitability of the data, do the assumptions of the chi-square test hold?

Gender * Ethnicity Crosstabulation

			Ethnicity				
			White	Black	Asian	Other	Total
Gender	Female	Count	53	44	46	0	143
		% within Ethnicity	42.7%	49.4%	58.2%	0.0%	47.7%
	Male	Count	71	45	33	8	157
		% within Ethnicity	57.3%	50.6%	41.8%	100.0%	52.3%
Total		Count	124	89	79	8	300
		% within Ethnicity	100.0%	100.0%	100.0%	100.0%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	12.136[a]	3	.007	.006		
Likelihood Ratio	15.219	3	.002	.002		
Fisher's Exact Test	12.663			.004		
Linear-by-Linear Association	.801[b]	1	.371	.394	.204	.035
N of Valid Cases	300					

a. 2 cells (25.0%) have expected count less than 5. The minimum expected count is 3.81.

b. The standardized statistic is -.895.

According to Fisher's exact test (exact p=0.004), there was an association between gender and ethnicity in our sample. For example, for women the highest percentage was Asian ethnicity (58%).

McNemar Test

When to use

To test if, according to the current data, the proportions in the population of a variable change based on another matched variable.

Hypotheses:

H_0: there is no association between the two (paired) variables
H_a: there is an association between the two (paired) variables

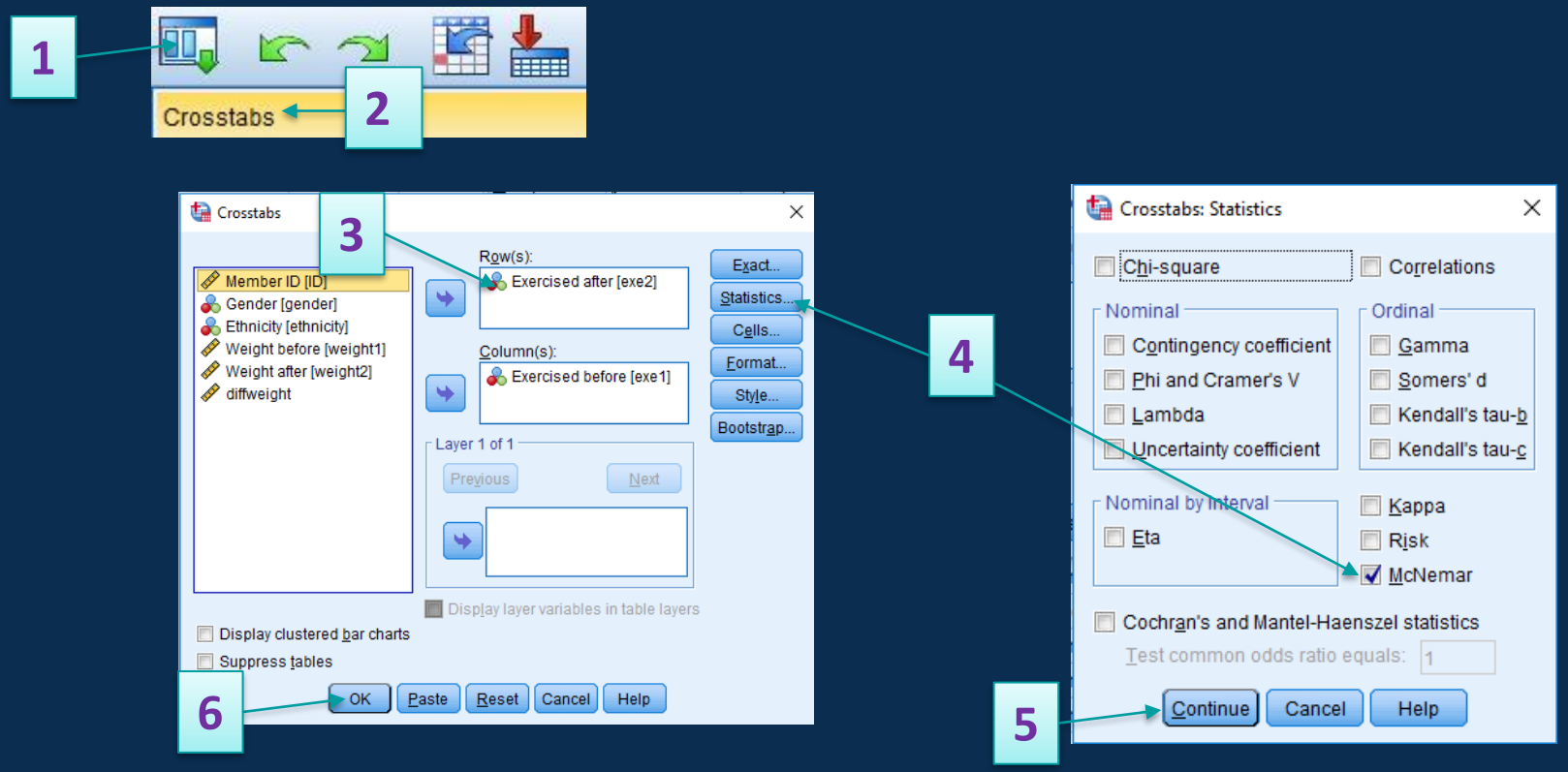
Assumptions:

- The observations are randomly and independently drawn
- There are at least 25 observations in the discordant cells
- The data are paired

SPSS Slide: ‘how to’

The next question is: do people exercise more after the programme than before?
Are the proportions of those exercised before the programme, different of those exercised after the programme?

Step 1: Use the appropriate test, here: ‘McNemar chi-square test’.



Output and Interpretation Slide

Step 1: Use the appropriate test, here: 'McNemar test'.

Exercised after * Exercised before Crosstabulation

Count

		Exercised before		
		No	Yes	Total
Exercised after	No	9	7	16
	Yes	8	5	13
Total		17	12	29

Chi-Square Tests

	Value	Exact Sig. (2-sided)
McNemar Test		1.000[a]
N of Valid Cases	29	

a. Binomial distribution used.

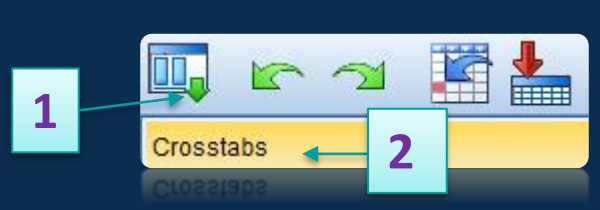
Step 2: Check the suitability of the data:

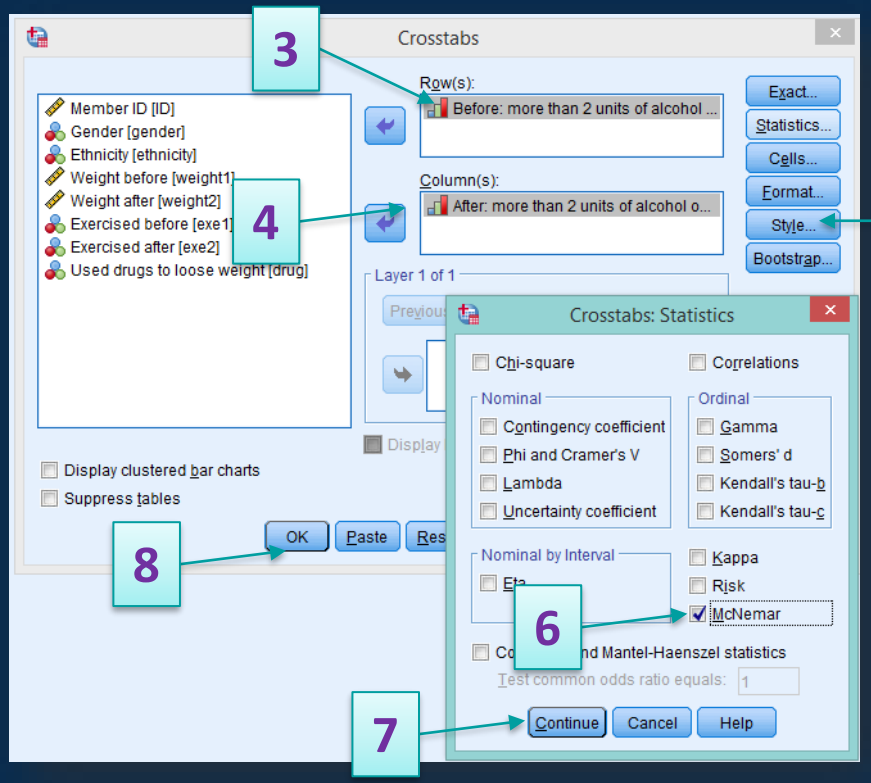
For McNemar's test, we needed at least 25 discordant observations, and paired data. We only have 15 observations. SPSS instead of printing the asymptotic p-value prints the exact test. However, SPSS always prints the exact p-value (binomial), even if the assumptions hold! (SPSS only does this).

Therefore, in SPSS only, the McNemar p-value is always valid, as long as we have paired categorical data.

SPSS Slide: ‘how to’

We can compute the McNemar test for PxP tables as well.





Analyse-> Descriptive Statistics->Crosstabs
Place before variable in ‘row’ box and after variable in ‘column’ box.
In ‘Statistics’ choose McNemar
Click ‘continue’
Click ‘ok’

Output and Interpretation Slide

Step 1: Use the appropriate test, here: 'McNemar test'.

Before: more than 2 units of alcohol on a weekend * After: more than 2 units of alcohol on a weekend Crosstabulation

			After: more than 2 units of alcohol on a weekend			
			Never	Sometimes	Always	Total
Before: more than 2 units of alcohol on a weekend	Never	Count	65	3	0	68
		% within Before: more than 2 units of alcohol on a weekend	95.6%	4.4%	0.0%	100.0%
	Sometimes	Count	9	148	0	157
		% within Before: more than 2 units of alcohol on a weekend	5.7%	94.3%	0.0%	100.0%
	Always	Count	0	13	62	75
		% within Before: more than 2 units of alcohol on a weekend	0.0%	17.3%	82.7%	100.0%
Total		Count	74	164	62	300
		% within Before: more than 2 units of alcohol on a weekend	24.7%	54.7%	20.7%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
McNemar-Bowker Test	16.000	2	.000
N of Valid Cases	300		

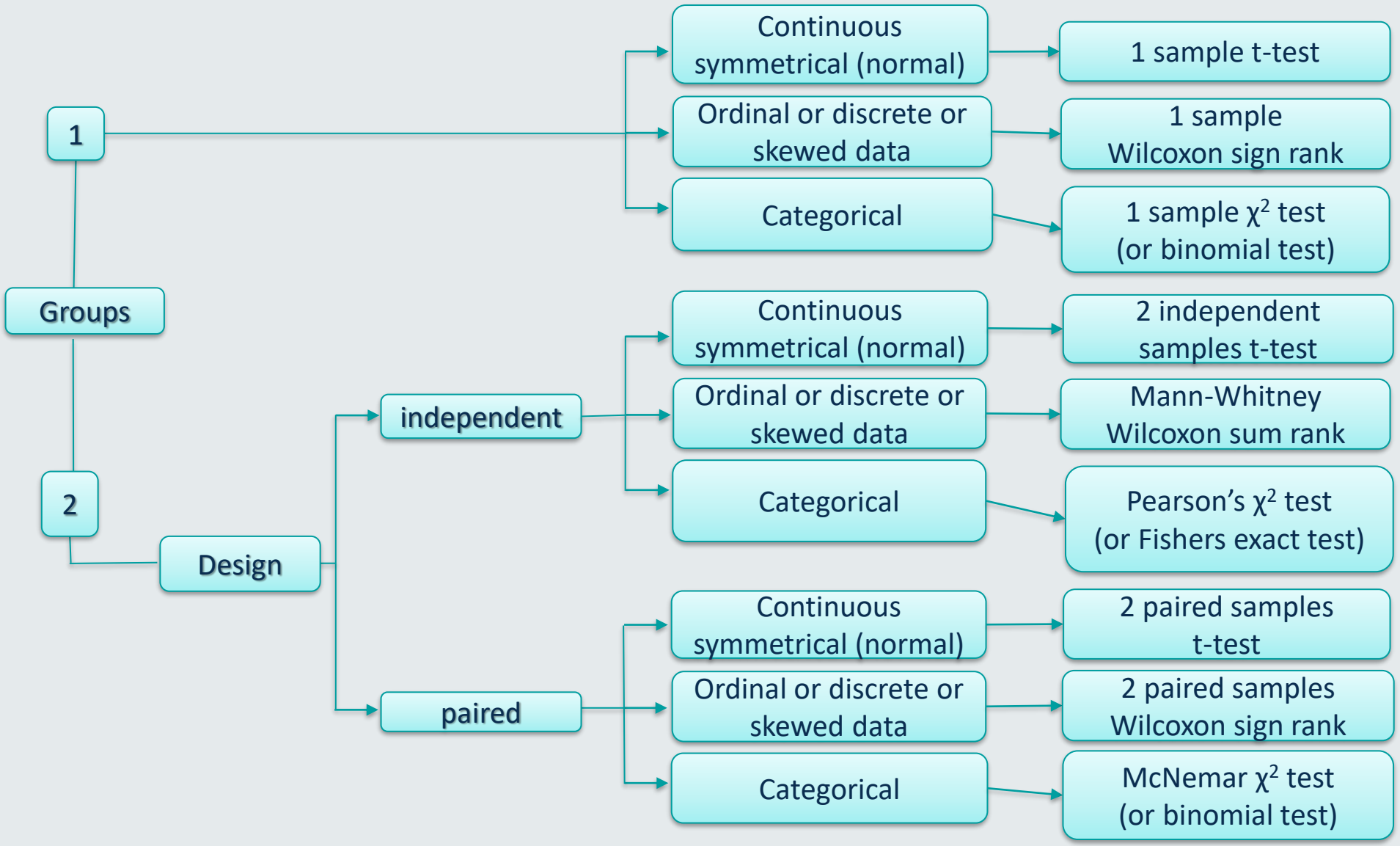
SPSS automatically prints the McNemar – Bowker test, which is the extension of the original test for pxp tables.

There is an association between the amount of alcohol consumed before and the amount of alcohol consumed after the programme (McNemar Browker χ^2=16.0, df=2, p<0.001).

Cheat Sheet

	Numerical data		Categorical data	
Hypotheses testing	parametric	non-parametric	chi-squared tests	exact tests
one group *versus* a pre-defined value	one sample t-test	Wilcoxon sign rank	one sample χ^2-test	One sample Binomial exact test
one group *versus* another group	two independent samples t-test	Mann-Whitney (Wilcoxon sum rank)	Pearson's χ^2-test	Fisher's Exact test
one group (twice or) *versus* another matched group	two paired samples t-test	Wilcoxon sign rank	McNemar test	Binomial exact test

Cheat Sheet in a Flow Chart…



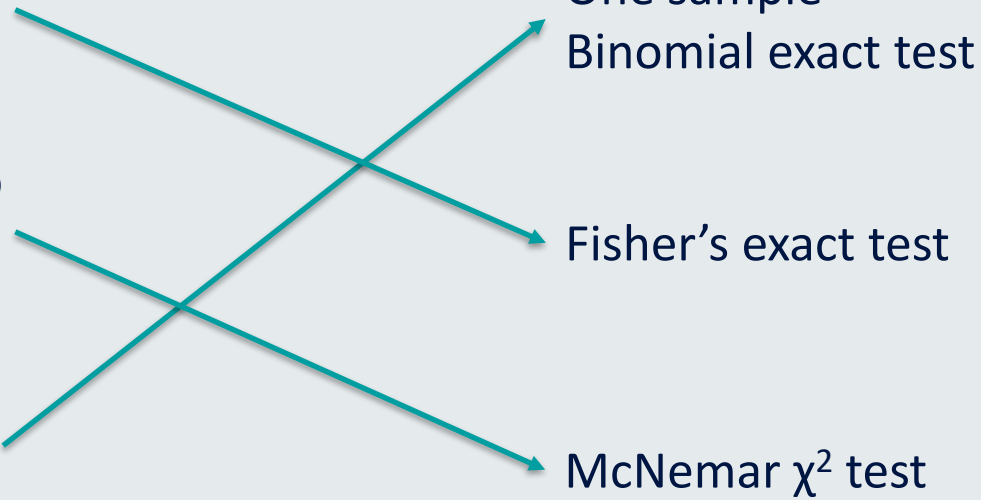
Knowledge Test

The assumptions have been violated. Match the scenario with the correct test.

Tom wants to test if boys' proportions in ADHD high/low classification groups are different than those of girls.

Tom wants to test if mothers' reported ADHD high/low classification for children are different than those reported by fathers.

Tom wants to test if children's high classification ADHD proportion is higher than 50%.



One sample Binomial exact test

Fisher's exact test

McNemar χ^2 test

Reflection

Reflecting on your own field of study.

Write down three examples from your research that would require the use of each of the three exact tests.

Reference List

Agresti and Finlay (2009) Statistical Methods for the Social Sciences, 4th Edn, Pearson Hall, Upper Saddle River, NJ.

- Comparison of Two Groups, Ch 7, pages 183-209
- Analyzing Association between Categorical Variables, Ch 8, pages 221-239

Field (2005) Discovering Statistics using SPSS, 2nd Edn, Sage, London.

- Comparing Two Means, Ch 7
- Categorical Data, Ch 16

Thank you

Please contact your module leader or the course lecturer of your programme, or visit the module's forum for any questions you may have.

If you have comments on the materials (spotted typos or missing points) please contact Dr Vitoratou:

Silia Vitoratou, PhD
Psychometrics & Measurement Lab,
Department of Biostatistics and Health Informatics
IoPPN, King's College London, SE5 8AF, London, UK
silia.vitoratou@kcl.ac.uk

For any other comments or remarks on the module structure, please contact one of the three module leaders of the Biostatistics and Health Informatics department:

Zahra Abdula: **zahra.abdulla@kcl.ac.uk**

Raquel Iniesta: **raquel.iniesta@kcl.ac.uk**

Silia Vitoratou: **silia.vitoratou@kcl.ac.uk**