Introduction to Ophthalmic Statistics

Gabriela Czanner PhD CStat
Department of Biostatistics
Department of Eye and Vision Science
University of Liverpool
and
Clinical Eye Research Center
Royal Liverpool University Hospital

Email: czanner@liverpool.ac.uk
Web: http://pcwww.liv.ac.uk/~czanner/

21 September 2016 (3.45 and 4.15)

MERSEY POSTGRADUATE TRAINING PROGRAMME

This is part of the workshop series:

Basic Statistics for Eye Researchers and Clinicians

Goal

- To give an overview of concepts
- … of statistical data analysis methods …
- … that occur in ophthalmology.

Notes

- We choose 10 main concepts only
- There are many more things to know
- We show how the concepts *connects* with each other

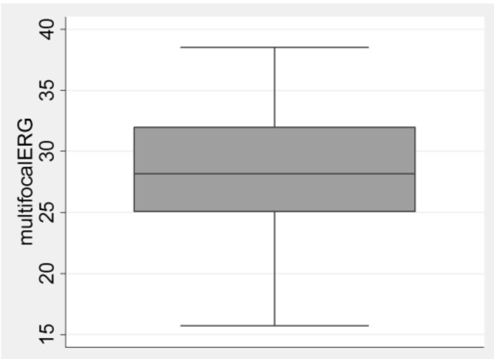
Outline

- The 10 main things to know about *statistical methods*

  …from simple to more complex notes
- Discussion
- References

1. There are two categories of statistical methods

Fig. Boxplot of mfERG central density in early DR (n=20)



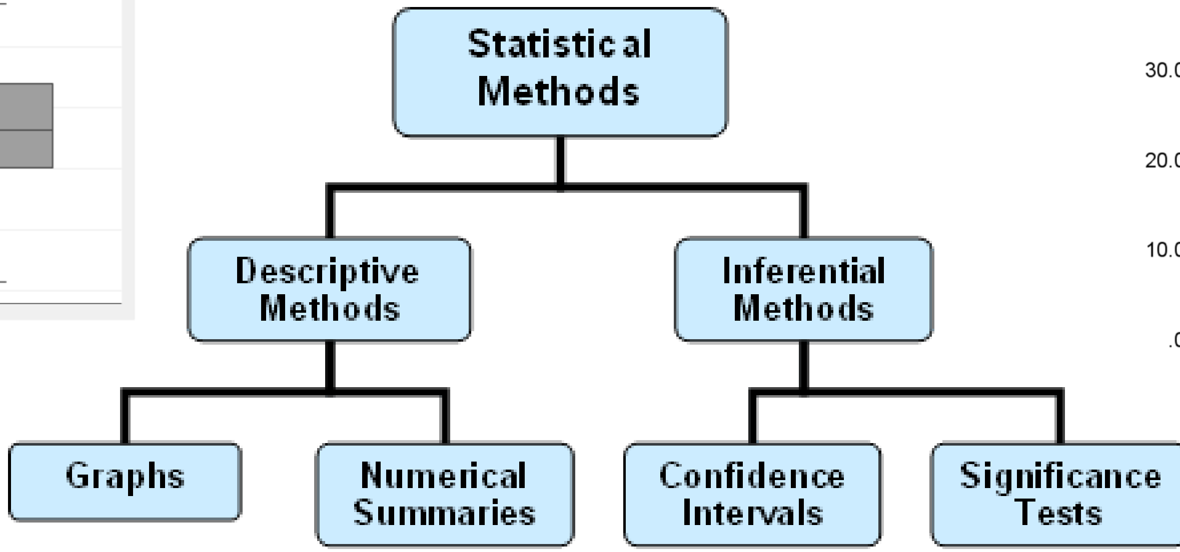
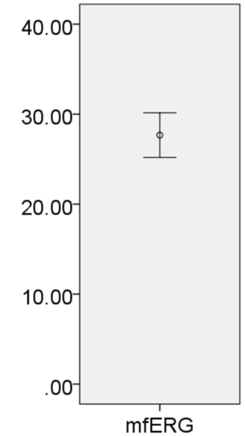


Fig. Confidence interval for mean mfERG in early DR (n=20)



Descriptive statistics is the discipline of quantitatively describing the data.

Inference is the process of deducing properties of an underlying distribution by analysis of data.

Thing to know: If our study is to test a hypothesis, we still need to employ both descriptive and inferential statistical methods.

2. What we measure on eye and patient is a *variable*

What is a variable?

- It is something whose value can ***vary*** across subjects and within subjects e.g. when same subject measured repeatedly over time or on same visit
- E.g. BCVA is a variable
- Some variables can be named as ***primary or secondary outcomes***

How is a variable measured?

- E.g. How to measure visual acuity? Snellen Chart and Early Treatment Diabetic Retinopathy study Chart (ETDRS). No appropriate conversion exists.
- E.g. How to measure blood pressure? Can we measure it as dichotomised, e.g. as high and low?

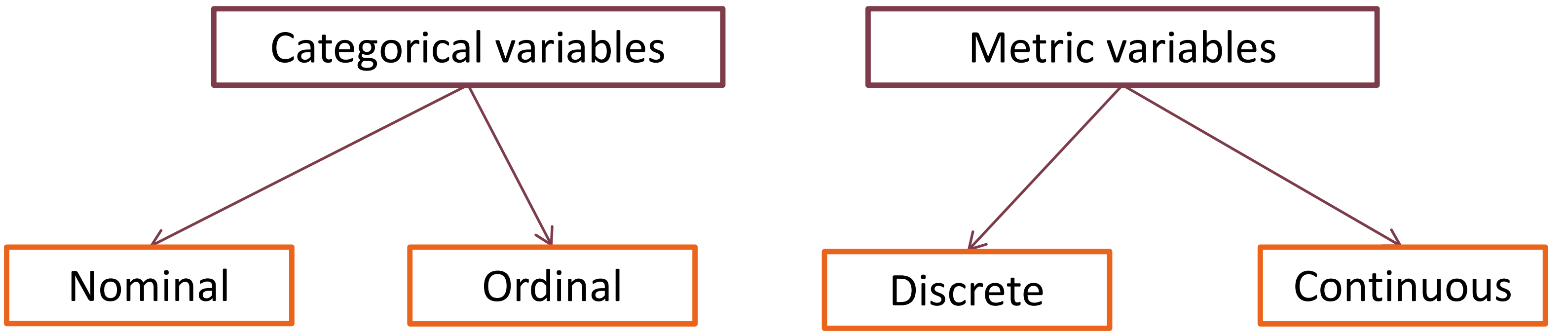
Data are values you get when you measure a variable

- E.g. Age = 40.2, 40.5, 31.1, 51.2, 31.2 years
- Gender: Male and Female
- Quality of Fluorescein Angiography image: Excellent, Good, Fair, Poor

Thing to know: Try to avoid dichotomisation.

Cumberland PM, Czanner G, Bunce C, Dore C, Freemantle N, Garcia-Finana M. Ophthalmic statistics note: the perils of dichotomising continuous variables. *Br J Ophthalmol* 2014; 98:841-843.

Types of variables



Nominal	Ordinal	Discrete	Continuous
Values in arbitrary categories	**Values in ordered categories**	**Integer values on proper numeric line or scale**	**Continuous values on proper numeric line or scale**
E.g. Gender: Male/Female	E.g. Quality of image: Excellent, Good, Fair, Poor	E.g. Number of scars in a retina	E.g. Age of the patient

3. The descriptive statistical methods

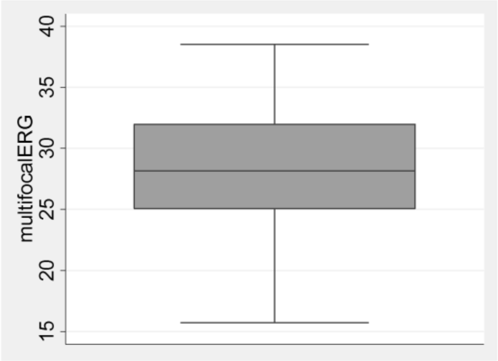
How to do a good *numerical* description of the data measured on continuous scale?

- If the distribution skewed - report median and quartiles
- If the distribution is symmetric - report mean and standard deviations

How to do a useful *graphical* description of data?

- Histogram and boxplot for continuous variables
- Barcharts for nominal and ordinal variables
- Scatter plots to explore associations between continuous variables

Fig. Boxplot of mfERG central density in early DR (n=20)



Why we need descriptive methods?

- Demographics tables for reports
- Quality assurance, e.g. outliers due typo errors
- May find new unexpected patterns or associations
- Explore the distribution of outcome variables, to support selection of inferential statistical methods (e.g. many methods require normal distribution)

Thing to know: Descriptive statistics show or summarize data in a meaningful way such that patterns might emerge from the data e.g. new unknown associations. Descriptive statistics do not, however, allow us to make conclusions beyond the data.

More info: Initiative STRATOS for observational studies, Topic Group 3 "Initial data analysis" http://www.stratos-initiative.org/

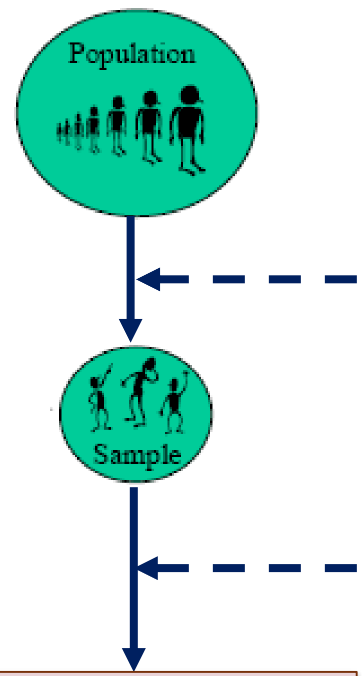
4. The inferential statistical methods

The main principle

We use data collected on a sample to make ***statistical inference*** about the population of interest.

The result of the inference is summarized in

- Confidence interval
- P-value



e.g. All DR eyes in UK

Sampling chance

e.g. a random sample of 202 eyes, one eye per patient

Inferential statistical data analysis method applied

Inference about populations of like individuals

e.g. a statement on the difference in multifocal ERG between early DR and no DR (diabetes only)

Thing to know: Absence of evidence is not evidence of absence.

BunceC, Patel KV, Xing W, Freemantle N, Doré CJ. Ophthalmic statistics note 2: absence of evidence is not evidence of absence, *Br J Ophthalmol* 2014;98:703-705

Nuzzo R. Statistical errors. *Nature*, 506, 2014

I once asked a statistician out.

She failed to reject me!

Thing to know: Absence of evidence is not evidence of absence.

Example: Colour blindness

- We are interested in whether there is an association between colour blindness and gender.
- We asked 240 men and 260 women about color-blindness. The results of a survey are as follows:

	Normal colour vision	Colour blind	Total
Male	221	19	240
Female	254	6	260
Total	475	25	500

Example: Colour blindness

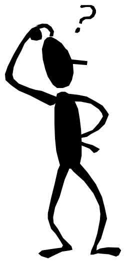
In sample 8% males and 2% females are colour blind

Does this mean that women are at less risk of becoming colour blind than men?

We rewrite research question into null hypothesis

- H_0: there is no association between colour blindness and gender

How to test the null hypothesis?



- We can use a method of comparison of proportions
- Result will be reported as Confidence Interval and P-value

Example: Colour blindness

- Confidence interval: range of plausible values for the "true" difference (usually use 95% confidence)
- Method of comparison of proportions
  - General formula (in large samples)

estimate $\pm$ 1.96 · standard error

95% CI for $p_1 - p_2$ is (2%, 9%)

$\Rightarrow$ there is a significant difference in the proportion with colour blindness between the 2 groups at a 95% confidence level *(because value 0 is not in CI)*

Note: P-value <0.05 *if and only if* the 95% confidence interval does not contain the hypothesised value. Here p-value is 0.007

5. Clinical and statistical significance are different concepts

Statistical significance

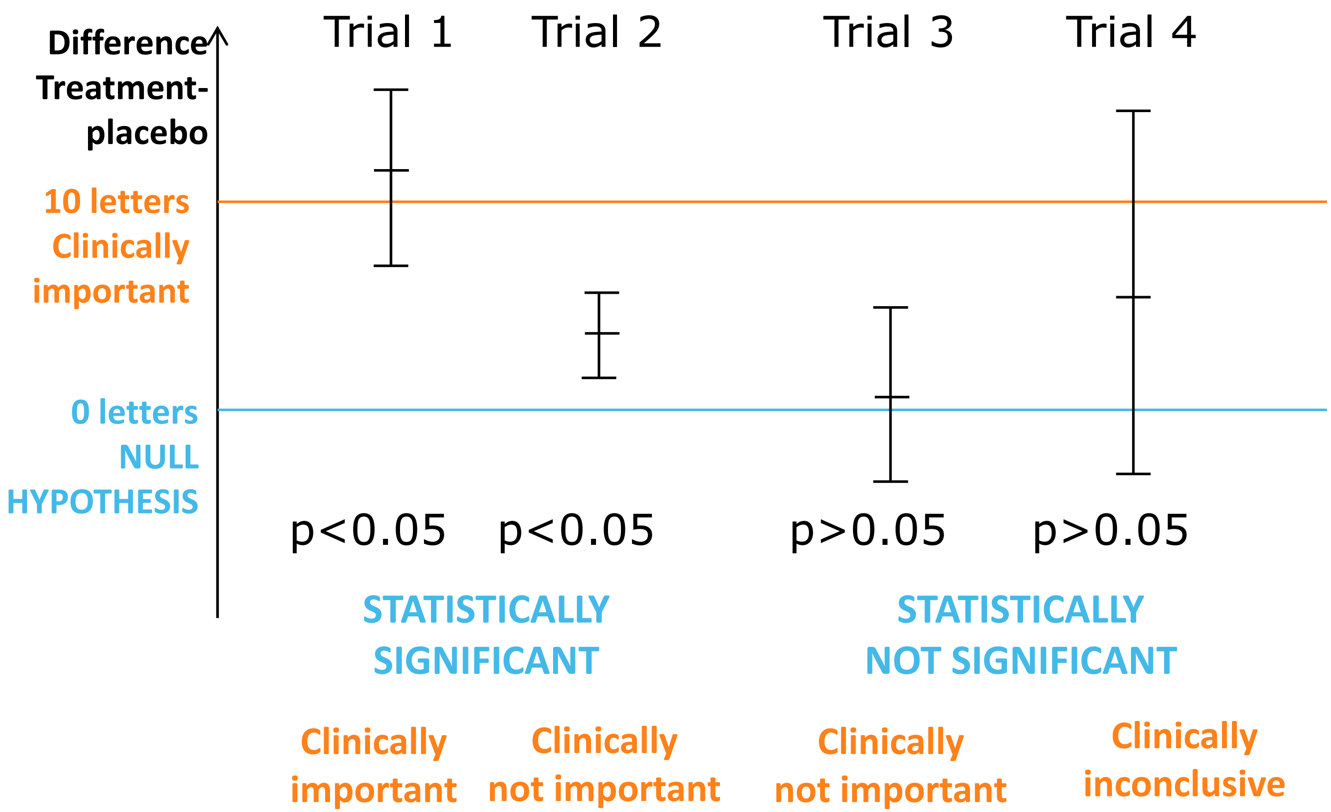
- It rules out chance as an explanation for the observed difference of means
- Summarized by p-value

Clinical relevance

- It tells if the observed difference is of clinical value
- Summarized by confidence interval

Kay R, 2014. *Statistical Thinking for Non-Statisticians in Drug Regulation*, Wiley Blackwell, 2nd edition.

Example: In a collection of 4 placebo-controlled trials in BCVA a difference of 10 letters in terms of mean increase of visual acuity is to be considered of clinical importance; anything less is unimportant. The results are given in the figure below.



Thing to know: Presentation of p-values together with confidence intervals is good practice.

6. Choosing the right statistical inferential method

Several factors to consider

- ***Objective*** of the analysis
  - e.g. Estimating prevalence, comparison of groups, instruments comparisons
- ***Type of data***
  - E.g. how the data are measured
- ***Correlations***
  - Data are ***independent*** if each patient was given only one treatment
  - Data are ***paired*** if each patient was given both treatments (cross-over design)
  - If BCVA of an eye measured at two time points
- ***Distribution of the data***
  - i.e. is it ***symmetric*** or ***skewed?*** *Is it* ***Normal?***

Example:

We want to compare IOP for female and male. We collected IOP on one eye per patient, 20 eyes in each gender group. What statistical method to choose?

Answer:

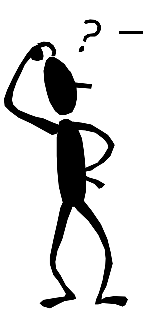
Objective: compare means in two groups

Type of data: continuous

Correlations: none

Distribution: needs to be checked with measured IOP data in each group.

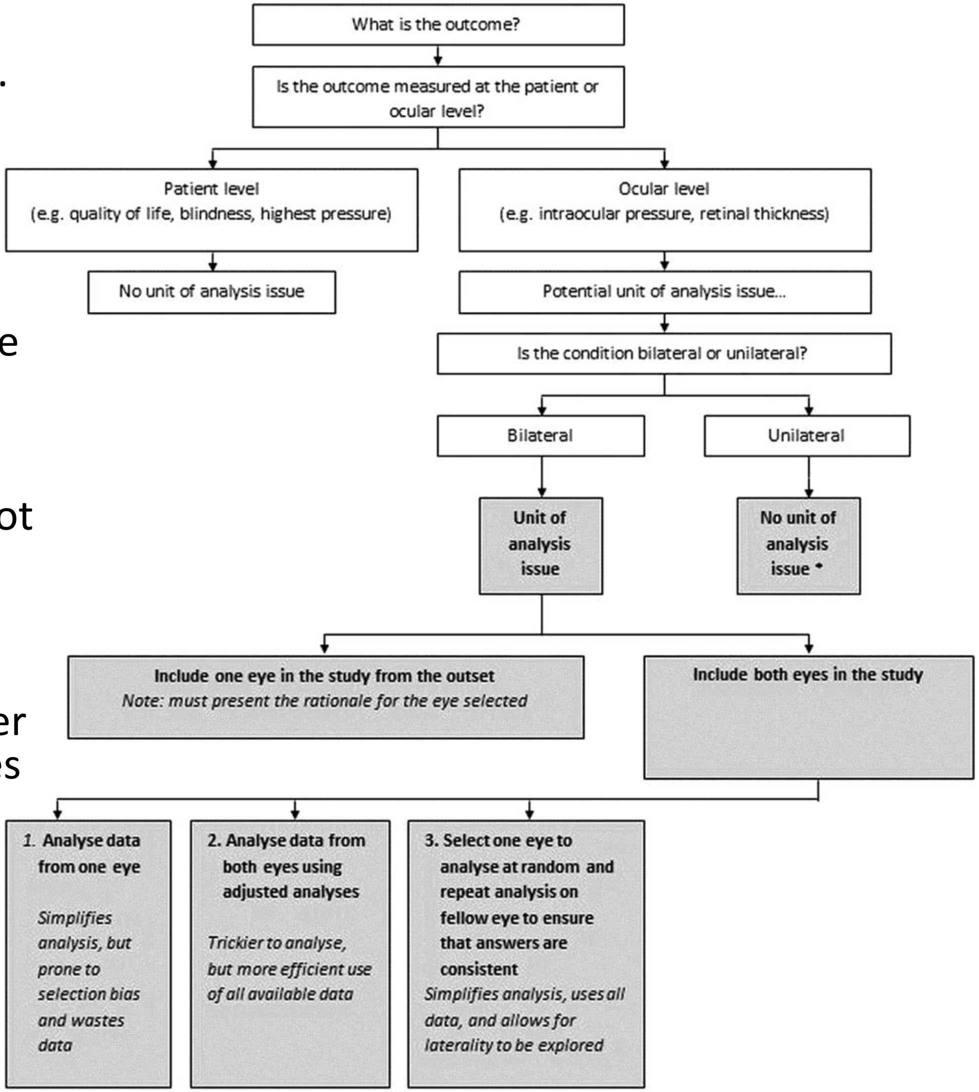
What method: we may use two-sample t-test if the histograms of IOP confirm normality of data



Thing to know: Always check if the *assumptions* of the statistical method are *satisfied.* If the assumptions are violated, then the estimates, confidence intervals and p-values can not be trusted.

7. Choice of inferential statistical method depends on *unit of analysis*

- **Eyes** exist in **patients**,
  - and typically patients have **two** eyes.
- Challenge for the study design
  - Do we use one eye?
  - Or both?
- Challenge for analysis
  - Measurements from two eyes of same patient may be ***correlated*** and hence carry less information than two eyes from two different patients; and standard data analysis methods are not appropriate.
- If eye is the unit of analysis
  - Then correct statistical method to analyse your data depends on whether you did measurements from both eyes or just one eye.



*unless repeat measures on or within an eye

Thing to know: Always report if the unit of your analysis is eye or patient.

Bunce C, Patel KV, Xing W, Freemantle N, Dore CJ, Ophthalmic statistics note 1: unit of analysis, *Br J Ophthalmol* 2014;98: 408-412.

8. Statistical methods *to compare groups* with one variable

Comparing two groups (e.g. two treatments)

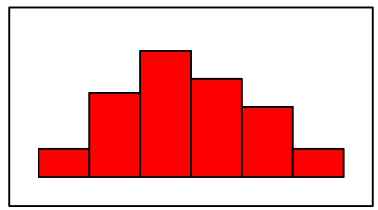
Type of Data

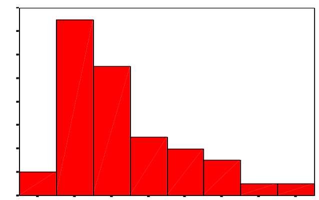
Categorical

Continuous Normal

Continuous Skewed

e.g. Color blindness yes/no or 3 severity grades of diabetic maculopathy





Tests for differences, independent data:

Chi-square χ^2 test

t-test

Mann-Whitney U test

Fisher's exact test

Tests for differences, non-independent (paired) data:

McNemar's test

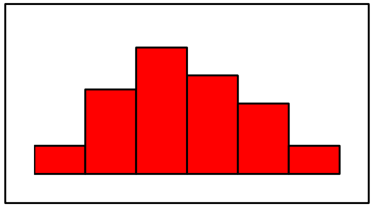
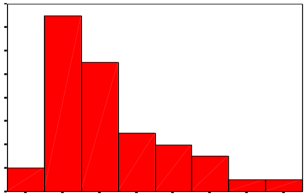
Paired (single sample) t-test

Wilcoxon test

Thing to know: If continuous data are skewed then applying a transformation (*log* or *sqrt*) can make them Normal, in such case a t-test can be used which has stronger power (as a parametric test).

Comparing three of more groups (e.g. multiple treatments)

Type of Data

Categorical	Continuous Normal	Continuous Skewed
e.g. Color blindness yes/no or 3 severity grades of diabetic maculopathy		

Tests for differences, independent data:

Compare two groups at a time. This leads to multiple comparisons.	One-way ANOVA **Assumptions**: Normal distribution of data in each group; Independence of units of analysis; Equal variability in each group	Kruskal-Wallis test

Thing to know: Be aware of multiple comparison problem and seek proper adjustments.

Cipriani V, Quartilho An, Bunce C, Freemantle N, Dore C. Ophthalmic statistics note 7: multiple hypothesis testing-to adjust or not to adjust. Br J Ophthalmol 2015;99:1155-1157.

9. Statistical methods for *association between two variables*

Association

- It is a synonym for ***dependence*** or ***relation*** (e.g. relation between age and choroidal thickness on OCT)
- It is a very general relationship: one variable ***provides information*** about another.
- One particular type of association is a linear association, also called ***correlation***.

Type of Data

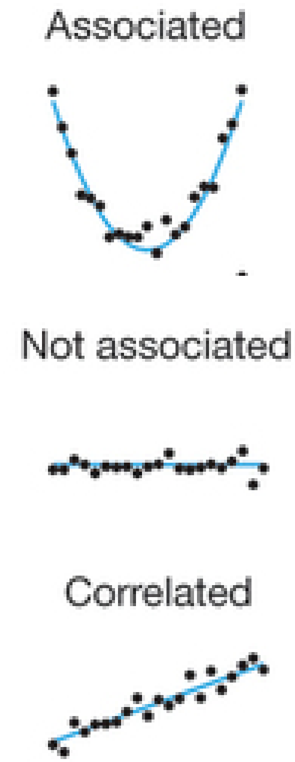
Both variables ordinal

Spearman rank correlation coefficient to assess an increasing or decreasing association

Both variables continuous

Pearson correlation coefficient and Linear Regression to assess a linear association

Statistical modelling for more complex associations



Things to know: 1. High correlation or association does ***not*** mean causation. **2.** Be aware of ***confounders*** when assessing the correlation. If confounders not taken into account, then paradoxical correlations values may arise. **3.** When assessing agreement between two measurement methods, the correlation should ***not*** be used.

Altman N, Krzywinski M. Points of Significance: Association, correlation and causation. Nature Methods 2015 12, 899-900.

10. Statistical methods for *relation between* *__several__* *variables*

These are more advanced statistical methods used in ophthalmic clinical research:

- Called ***"statistical modelling"***, or ***"multivariate modelling"***
- **Most commonly used methods**
  - One –way, Two-way ANOVA
  - ANCOVA
  - Simple and Multiple linear regression
  - Logistic regression
  - Survival models
  - Linear and nonlinear mixed models, also called "hierarchical models"
- **Why to use them? Can allow to:**
  - adjust for confounders
  - more complex (than linear) relationships
  - non Normal outcome
  - better utilisation of missing data
  - imbalanced designs…

Main principle

Characterise the relation between outcome and predictors via a model. The model is fitted to the data. E.g.

BCVA = 50 - 8 Diabetes - 0.4 Age + Error

Outcome **Predictors (or Covariates)**

Assumptions of the model are checked

Use the model:

- **To answer research questions**

  E.g. Effect of diabetes on BCVA?
- **To make predictions for individual patients (or eyes)**

  E.g. BCVA at age 41 yrs in diabetes

Thing to know: More sophisticated methods can answer same research questions as simple methods (such as t-test), while adjusting for the effect of confounders.

Summary

- The best practice is to apply statistical data analysis methods in three stages:
  - **Stage 1.** Do descriptive summaries of data
    - Numerical and graphical summaries
    - E.g. two histograms of BCVA: one for each Diabetes group
  - **Stage 2.** Do statistical inference on one variable at a time, or do analysis of association between two variables (depending on your research goal).
    - This means that you are not adjusting for confounders
    - Called **unadjusted** inference
    - e.g. two-sample t-test for BCVA of two diabetes groups
  - **Stage 3.** Do statistical inference while **adjusting** for confounders.
    - This is typically the main result for your report
    - e.g. linear regression of BCVA with Diabetes as main risk factor while adjusting for Age

! However, your report needs to also contain the results from all stages.

Discussion

- **Good design of the study is important**
  - Statistical data analysis (inference) methods however sophisticated can not “rescue” a poor study design
  - Consult a statistician as early as possible !
  - Sample size calculations need to be done to assure required power of analyses
- **Choice of statistical inferential methods**
  - It is a complex task
  - Decision need to be done early at designing stage, and described in the statistical analysis plan
  - Two types of statistical errors: Type I error “false positive” and Type II error “false negative”
- **There are other statistical methods not mentioned here**
  - Bayesian statistical methods, Methods for genetics data
  - Multivariate statistical methods (Principal factor analysis, Cluster analysis, Factor analysis)
  - Statistical methods for imaging data … and many more!

Discussion

A quote from feedback forms:

- “I think this needs to come across that statistics is a difficult branch of science, …
- … that people need support to make sure that they are doing the correct tests.
- Otherwise doctors think they can do complex stats, but that would be like asking a statistician to do cataract surgery”.

Recommended reading

Books

- *Practical statistics for medical research* by Douglas G. Altman [A general and intermediate level book.]
- *Medical statistics from scratch* by David Bowers [A perhaps more readable intermediate level book.]
- Kay R, 2014. Statistical Thinking for Non-Statisticians in Drug Regulation, Wiley Blackwell, 2nd edition. [Very thorough intermediate level book.]

Journals' with series on how to do statistics in clinical research

- BJO has ***Ophthalmic Statistics Notes*** (8 published, http://www.brcophthalmology.org/ophthalmic-statistics-group-osg)
- American Journal of Ophthalmology has ***Series on Statistics***
- British Medical Journal has series ***Statistics Notes***

Guide on how to do clinical research in ophthalmology

- Clinical Research. A primer for Ophthalmologists. International Council of Ophthalmology. February 2009, http://www.icoph.org

Nature notes

- http://www.nature.com/collections/qghhqm/