Randomised Controlled Trials in Mental Health- AFFIRM Short Course

Analysis and Reporting 2

Jenny Hellier

jennifer.hellier@kcl.ac.uk

Alan J. Flisher Centre for Public Mental Health

AFFIRM
AFrica Focus on Intervention Research for Mental Health

Analysis and Reporting Objectives

- Design stage
  - Specify your analysis
- During the trial
  - Prepare to analyse
- At the end of the trial and beyond
  - Getting to know your data
  - Analysis data sets
  - Hypothesis testing and p-values
  - Some key analysis points

The design stage

- Think about the analysis from the start!

"To consult the statistician after an experiment is finished is often merely to ask him to conduct a post mortem examination. He can perhaps say what the experiment died of. "

Ronald Fisher

The Trial Statistician

- Involve a statistician at the grant proposal stage
- The statistician can provide expertise on randomisation, sample size, blinding and analysis methods, responding to reviewers comments
- The analysis method will determine:
  - The sample size calculation
  - The data collected
  - Data checking required

- http://www-users.york.ac.uk/~mb55/guide/guide14.pdf

During the trial

- Need to input into data management and work closely with the trial manager
- Will prepare appropriate DMC, TSC reports
- In collaboration with the trial team will write the Statistical Analysis Plan (SAP)

The Statistical Analysis Plan

- Definition: ICH Guidelines E9 Statistical Principals for Clinical Trials

"A more technical and detailed elaboration of the principal features stated in the protocol. The plan may include detailed procedures for executing the statistical analysis of the primary and secondary variables and other data and should be finalised before breaking the blind"

Why write a plan?

- To avoid
  - Post hoc decisions which may affect the interpretation of the statistical analysis
  - Torturing the data until it speaks!
- To allow
  - The analysis to be repeated if necessary
  - Documentation of reasoning and choice of analysis and methods
  - Formal agreement within the study team and committees
- Should written in an easy language
- Formal record – publish
- Keep blind writing the SAP and analysing the trial
- Finalise the plan as soon as possible

The Statistical Analysis Plan

- Topics covered include:
  - Objectives and Outcomes
    - Primary, secondary
  - Data collection
    - Baseline and follow up data
  - Blinding
    - Of interventions and analysis
  - Trial Populations
  - Statistical considerations
    - Missing data, non compliance, multiplicity
  - Descriptive Analyses
    - Available data, missing data, baseline comparability of randomised groups, flow diagram

Statistical Analysis Plan

- Interim analyses
  - Rationale, number, stopping rules
- Analysis of primary outcome
  - Method, adjustment, assumption checks
- Analysis of secondary outcome
  - Method, adjustment, assumption checks
- Subgroup analyses
  - Definition, sample size justification
- Analysis of Safety
- Shell tables
- Scoring algorithms

Analysis sets:

- Intention To Treat:
  - All randomised patients are analysed in their allocated group
  - This is regardless of whether they took their treatment, turned out to be ineligible
  - This is an unbiased analysis
- Per-Protocol
  - Complier Averaged Causal Effects
- Complete Case

Analysis Sets: ITT in Reality

- ITT often does not include everybody
  - Participants who provide no data at all

- There are statistical methods that may be able to handle this under certain assumptions
  - Think about whether assumptions are valid
  - Sensitivity analysis

  - http://www.ncbi.nlm.nih.gov/pubmed/21300711

Subgroup Analysis

- Any subgroups should be defined a priori
- Unless the trial was powered to detect a subgroup effect all subgroup analysis should be deemed exploratory
- The number of subgroups tested should be kept to a minimum
- Multiple testing should be accounted for

Analysis sets

- Advantageous to demonstrate a lack of sensitivity of the principal trial results to alternative choices of the analysis set used

- When the intention to treat and CACE lead to the same conclusion confidence in the trial is increased

At the end of the trial: Data Locking

- The last patient has completed the trial
- Data checking takes place by the trial team and the data is frozen
- Data checking takes place by the statistician and the database is locked
- Data is extracted for final analysis
- Statistician performs a blinded analysis if possible
- Once complete everyone is unblinded

Preparing for the analysis

- Data cleaning
  - Is the data we have plausible?

- Checks may include
  - Range checking
  - Logical checks
  - Date checking
  - Outliers
  - Missing data

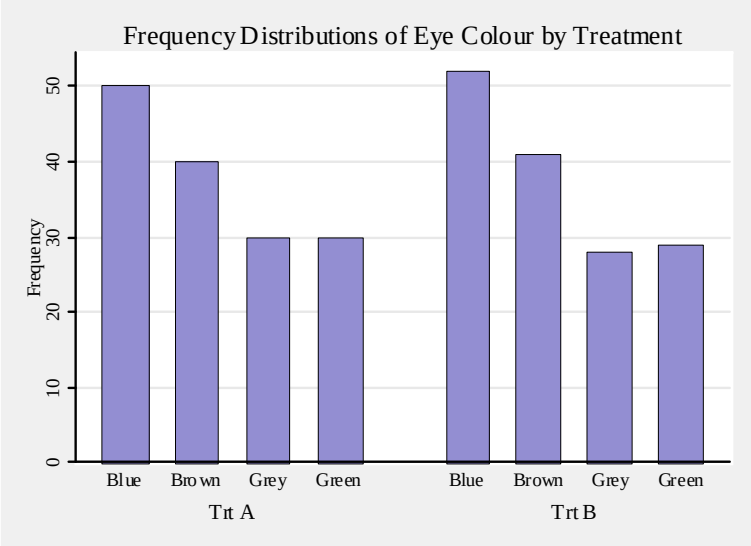
Getting to know your data

- It is important to understand your data before rushing to analyse it

- Get familiar with
  - Ranges and distributions
  - Types of data
  - Coding used
  - Where missing data is and how much

Categorical data

- Many be binary or category
- Summaries by n(%)
- Be aware that categorical data can be ordinal (eg severity, none, mild, moderate, severe) or nominal. Orderings contains useful information we should use in the statistical analyses.

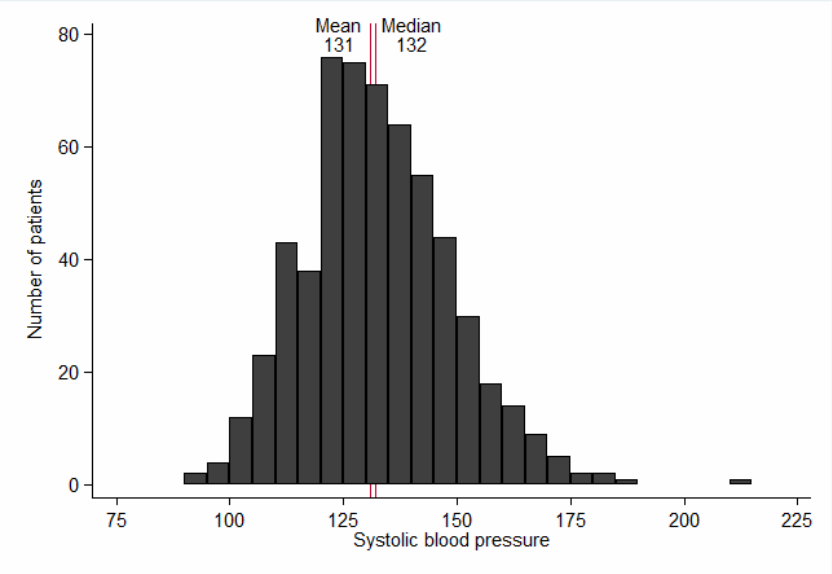


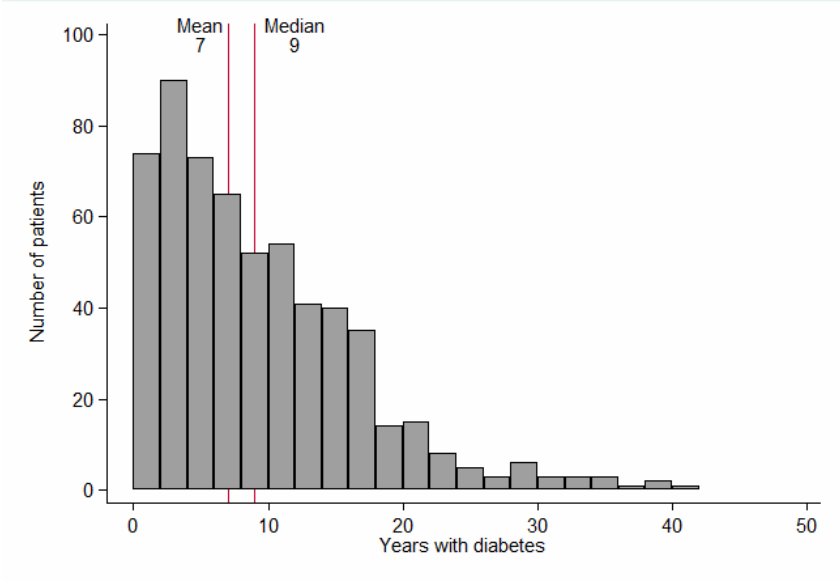
Numerical Data

- Data can be discrete: can only take a limited number of distinct values
  - Number of children 1,2,3,4 etc
- Data can be continuous: can take on any numerical value
  - Height, weight, blood pressure
- Summarise by
  - A measure of location (mean , median)
  - A measure of spread (standard deviation, IQR)
  - Dependent on the distribution of the data

Numerical Data: Graphical Summaries

- Histogram





Baseline comparability

- Baseline imbalance may occur because the randomisation was incorrect or just by chance
- Should be discussed from a clinical point of view not a "statistically significant" view
- You may want to adjust the analysis
- http://www.thelancet.com/journals/lancet/article/PIIS0140-6736(00)02039-0/abstract

Statistical Hypotheses about the outcome

- Null hypothesis (no difference):
  - The mean blood pressure (BP) in group A is the same as the mean BP in group B.
  - Mean BP(A) - Mean BP(B) = 0

- Alternative hypothesis (there is a difference):
  - The mean blood pressure (BP) in group A is different to the mean BP in group B.
  - Mean BP(A) - Mean BP(B) ≠ 0

Hypothesis Testing

- We calculate a test statistic using information from our sampled data
- The way the test statistic is calculated depends on the type of data we have
- The test statistic is assumed to have a particular distribution when the null hypothesis is true

P-values

- By comparing the calculated test statistic with the assumed distribution of the null hypothesis we generate a probability statement about the likelihood of the observed data
- It is the probability that you would see as much (or greater) difference between the treated and control groups by chance alone
- 0.05 is an arbitrary cut –off
- Sometimes known as statistical significance
- Errors
  - To conclude “no treatment effect” when p>0.05
  - To conclude that when P<0.05 this will always equate to a real and clinically important effect.

Estimation and Confidence Intervals (CI)

- Our research objective is to use the estimates (e.g. mean BP) found in our sample to make inferences about the population.
- If we were to repeat the study our estimate of the mean BP would be slightly different.
- To express this uncertainty around our estimate we calculate a confidence interval
- The confidence interval is a range of values which we are fairly sure includes the true population parameter.
- Narrower interval = more precise estimate
- Usually calculate 95% CIs
  - If we took 100 samples then 95 of the CIs would contain true population parameter
  - Corresponds to p-value of 0.05