Week 11: Understanding Results (Part 1)

What are results?

Results are a way of making sense of data:

- Summarises information from data
- Draws conclusions from data
- Makes estimations from the data beyond the sample
- Results are summarised in the results section of a journal article

Qualitative Results

- **Results are in words**

Quantitative Results

- **Results are in numbers**

Why understand results?

Research in media are typically:

- Health scares
- Miracle cures
- Quirky approaches
- Not accurate or correct
- Prone to overvaluing diagnoses and treatments

Health professionals also underestimate their own cognitive bias that:

- Overvalues results which confirm our thoughts
- Undervalues results which challenge our thoughts

Manipulation of Data

One source said…..
"cocaine floods the playground"
"use of the addictive drug doubles in one year"
"children are out of control… drugs to blame!"
Government report said…
"almost no change in patterns of drug use"

Example

Actual figures:

- 1% in 2004
- 2% in 2005
  - Absolute increase = 1%
  - Relative increase = 100%

Rounding of decimals can also affect results

- Small change of 1.4% versus 1.5% can be made bigger by rounding to nearest whole number (becomes 1% versus 2%)

Therefore:
You must be able to:

1. Read and understand results from pre-appraised sources
2. Read and understand results from original sources

To understand results you first need to understand data.

Level of measurement

Nominal

- Categories only
- E.g. ethnic group (Chinese)

Ordinal

- Categories and ranks
- E.g. University Grades

Interval or Discrete

- Categories, ranks and equal intervals between
- E.g. IQ scores

Ratio or Continuous

- All of the above, plus a true zero point
- E.g. height

Nominal and ordinal data is called categorical (or dichotomous) data

Interval and ratio data is continuous data

Types of Data

Continuous data is an actual value of the measurement

Categorical data is the number of cases that fall into a category

Normal curve

Example:
On a scale of 0-10 what is the worst headache you have ever had, when 0 is none and 10 is the worst you can imagine?
To summarise the data of a group you can:

1. Average headache pain for the group

Average is 7+5+7+2=21/4=5.25
This is continuous

2. Count how many have pain 5 or above

Number with more than average headache pain =3 out of 4 people or 75%
This is categorical

Basic maths, but important as they have different statistical tests used to interpret them

Measure of Central Tendency

These are numeral indices that provide a quantitative summary of the centre of the distribution

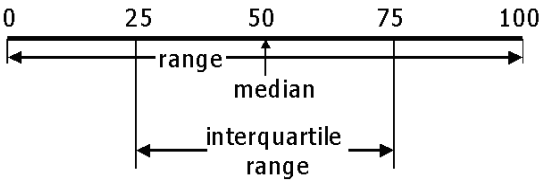
- Mean = average number
- Median = middle number
- Mode = most common number

Measures of Variability

"dispersion or spread of score"

A measure of central tendency means little on its own. It is always reported with the variability around it:

- Range
  - Difference between the highest and lowest score
- Percentiles and quartiles
  - Percentiles divide data into 100 equal portions
  - Quartiles divide the data into 4 quarters
  - Q1=25%, Q2=50%, Q3=75%
  - Q1 to Q3 = interquartile range (middle 50% of scores)



- Standard deviation
  - The most commonly used variance measure
  - The average difference between any score and the mean
  - Important because it includes information on all scores

Descriptive results

- Concerned with organising and summarising information about a collection of actual observations
- Used to characterise the shape, central tendency and variability of the observations
- Applies to known data, usually the sample

<u>How do we describe data?</u>

1. By measures of central tendency
2. By measures of dispersion (variability)
3. By measures of association

<u>Methods of description</u>

With numbers

- Frequencies and percentages
- Measure of central tendency
- Measure of dispersion, spread, variance
- Measure of association

With tables and figures

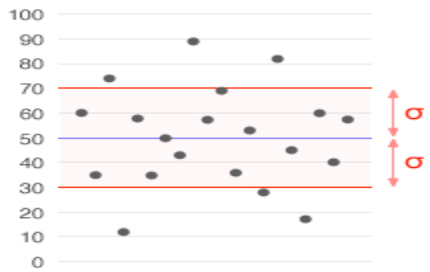
- Frequency distribution
- Bar chart, histogram, polygon
- Scattergram

Frequency Measures in Epidemiology

Incidence (incident cases)

- The frequency of <u>new occurrences</u> of disease, condition or death in a define population <u>over a period of time</u>

Example:



In this example, the mean = 50 and the SD = 20
In the results sections and tables this is represented as 50 (20)

	Group 1 Mean (SD)	Group 2 Mean (SD)	Group 3 Mean (SD)
Height (cm)	166.0 (10.97)	165.7 (29.7)	168.3 (11.1)
Weight (kg)	92.3 (17.9)	100.5 (24.9)	97.3 (22.7)
BMI	33.7 (6.6)	34.3 (7.4)	34.6 (7.1)

Important Terms

- The mean or median is called the **point estimate**
- The standard deviation or in the inter-quartile range is called the **measure of variability.**

The point estimate and its measure of variability are probably the most common statistics. These terms will be used frequently

Take home message

Step 1: Decide if data in study is continuous or categorical
Step 2: Decide if data is descriptive or inferential

Prevalence (prevalent cases)

- The number of persons in a defined population who have a specified disease or condition at a point in time

Measure of Association

- Measures the strength of linear relationship between 2 variables
- Measures association, agreement or correlation
- Used to measure *reliability* amongst other application
- Result is usually called a coefficient
- Figure is called a scattergram

Correlation

- Is common measure of association
- Can only be used for variables that are continuous
- Positive correlation e.g. age/weight
- Negative correlation e.g. age/exercise
- Ranges from 1 tp -1
    - 1 = full positive correlation
    - -1 = full negative correlation

Week 12: Understanding Results (Part 2)

Inferential Results

Inferential Statistics

"inferential statistics are about *what can be inferred (means estimated, expected, suggested, generalised, approximated)* from the sample to the population"

- Necessary for answering questions for those beyond the sample
  - Cause
  - Prevention
  - Diagnosis
  - Treatment
  - Prognosis

How do we estimate population characteristics from sample data?

- We use the ***normal curve*** as a model for making statistical assumptions and estimations
- When data is not normally distributed different methods (ranks) are used

Understanding Results

- Statistical Significance
- Clinical Significance

RCT Result Example

The results from an RCT about a new weight loss medication has provided a result of:

Difference in means = 8kg (95%CI 6kg-10kg)

$p<0.05$

This is the difference between the intervention group and the control group

Answer:

Might be presented in 3 ways:

1. Weight loss difference in means =8kg
2. Weight loss difference in favour of intervention group = 8kg
3. Difference in means = -8kg (-10 to -6)

Point estimate for this example = difference in means = 8kg

95% Confidence Interval (CI) means we are 95% certain that the true result for the population lies in this range

Hypothesis Testing

Null Hypothesis (H_0): proposes that there is no effect

Alternative Hypothesis (H_a): the opposite of the null hypothesis

For an RCT this means whether the group outcomes (results) are the same or different

"The process of deciding statistically whether the findings of an investigation reflecting chance or 'real' effects at a given level of responsibility"

There are 2 possible explanations for a positive outcome (groups are different) in a study:

1. The research hypothesis (alternative hypothesis) is correct
2. The observed difference between groups occurred by chance (result was a fluke)

Probability

Probability (p) is the likelihood an event will occur, given all possible outcomes. Value ranges between 0.0 and 1.0 (0-100%)

For example: what is the probability of getting tails if you flip a coin?

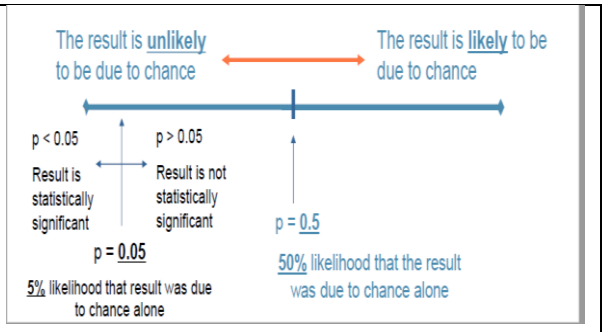
P= ½ = 0.5 (same as 50%)

p-value

hypotheses are accepted and rejected at a threshold of pre-set probability that the result is due to chance

p= probability due to chance

*in research, we often **5%** as an amount of error/chance. This is the **threshold** for whether a difference between groups is due to the group being researched, or just chance.*



- $p < 0.05$: a *small probability* that the difference between groups is due to chance *"statistically significant"*
- $p > 0.05$: a *high probability* that the difference between groups is due to chance *"not statistically significant"*

A positive finding is a small value – you want the difference to be due to the treatment, not chance.

Clinical Significance

Limitations of the p-value

A result is statistically significant if the p-value is less than 0.05.

- The result is partly dependent on sample size: harder to find statistical significance with small sample sizes: easier with large sample sizes
- Is like a yes/no answer
  - Gives no information about how large the effect size is
  - Gives no information about whether the effect is clinically meaningful

Point estimate

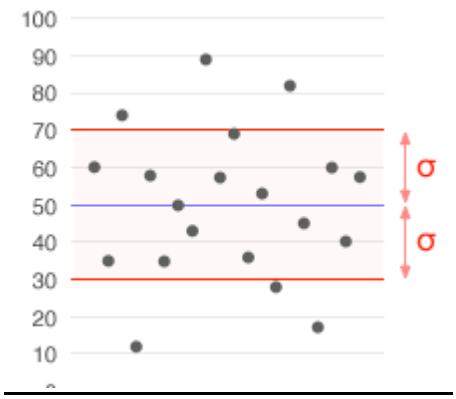
"a single number regarded as the most plausible value from the sample data"

Estimation will provide a result AND a degree of certainty:

- In the sample the results are 100% certain
- In the population we can't be completely certain
- We choose a pre-determined level of uncertainty – usually 95%

Measure of Variability

Standard deviation



Confidence Interval

- The most common result you will see is the 95% confidence interval
  - This is equal to 2 standard deviation
  - The % suggests the degree of certainty of the estimation. As usual we accept 5% chance of uncertainty (95% chance of certainty)
  - A range will be given for the interval of plausible values

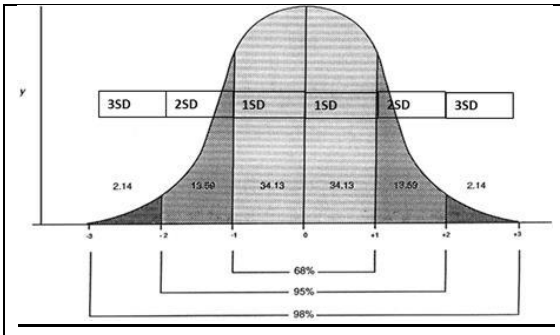
For our example:

For weight loss in favour of the intervention group:

Difference in means = 8kg

(95% Confidence Interval is 6kg to 10kg)

We are certain that the true result for the population lies between 6kg and 10kg



For our example:
95% confidence interval = 2 standard deviation

- Turns the sample result (point estimate) into what we estimate for the whole population
- Usually the most important statistic to find in a results section for the evidence professional
- Most common measure of variability used for estimation is the confidence interval

OR
We are certain that your client will lose between 6 and 10kg from this intervention

8kg (95%CI 6kg – 10kg)

Point estimate

Probability level

Measure of variability

Clinical Significance

Example 1

Is resistance training effective for weight loss in adults?

Study type: RCT or SR

Result: the weight loss in favour of the intervention group was 8kg (95%CI 6kg – 10kg)

Interpretation:

In this study the treatment group lost 8kg more (on average) compared with the control group. We can estimate, with 95% certainty, that the true difference in the population would be between 6kg and 10kg. My clients would be expected to lose between 6 and 10kg with the same training program.

CI Activity

Question:

Does anabolic steroid use affect school grades compared to those who do not use anabolic steroids?

Study type: RCT or SR measure as marks/100

Result: 3 marks (95%CI 1-5 marks) in favour of the intervention group)

What does this mean?

Week 13: Interpreting Results

Minimal Important Difference

Minimal Important Difference (MID) is the *smallest worthwhile difference* (improvement) expected by a patient to proceed with a treatment

Example:

Pain scale 0-10 (0 is no pain, 10 is the worst pain imaginable)

You have a heel pain and have marked this as 8 on pain scale

Treatments options are:

1. $10 comfy cushioned heel pad
2. $300 custom orthoses, daily strengthening and stretching
3. $5000 surgery, 6 weeks' functional interruption, post-operative discomfort and some risk of harm

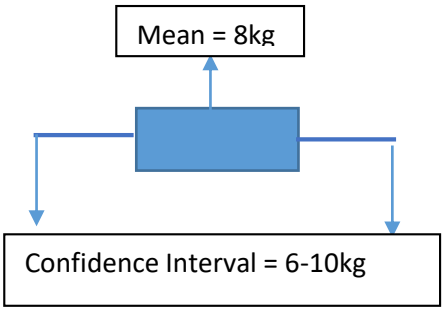
You compare the treatment effect size with the minimal important difference.

MID for having surgery = 80% improvement

95% CI = between 50% and 60%

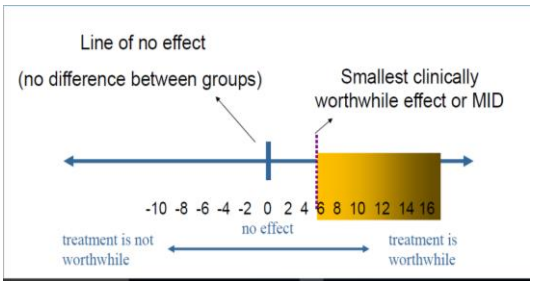
This means the best evidence suggests average improvement in pain following this procedure is between 50% and 60%. The patient does not proceed because their minimum improvement wanted to undergo surgery is at least 80%

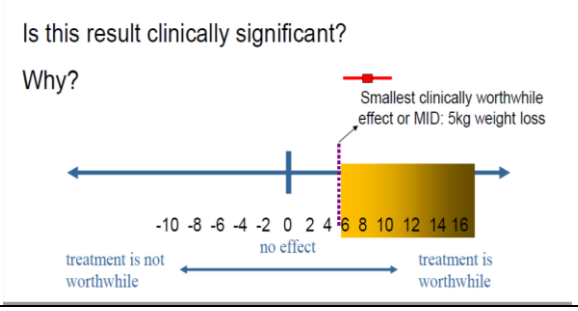
Tree plot



The box represents the point estimate

The line represents the confidence interval

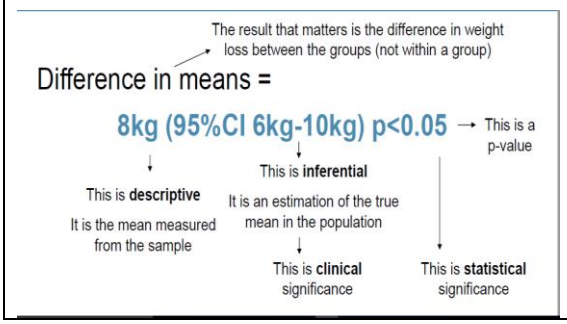




Clinical Significance

Statistical significance is that p-value <0.05

Clinical significance is when the treatment effect (confidence interval) is equal or more than the MID



Revision

Continuous data involves an actual measurement e.g. average student mark on assignment = 62 marks

Categorical data involves a proportion of events e.g. how many fails compared to passes = 5 fail grades compared to 10 pass grades

Both examples could provide information about student performance in an assignment