

Displaying data and interpreting results

Displaying data

- Visual methods can make the point much stronger than simply describing the data
- Appropriate use of tables and graphs can enhance the message you are delivering
- But they have potential to confuse or convey wrong messages

Guidelines for making good graphics

- Decide on the point you wish to present, then chose the appropriate method
- Emphasise one idea at a time in a figure
- Use conventional graphing methods (eg; time almost always plotted along the X axis)
- Pay careful attention to the scaling of the graph
- Always include a title – which should focus on interpretation of the data, not data itself

Guidelines for making good graphics, cont

- Graphs and tables must be self-contained, and stand on their own without reference to text
- Clearly label to indicate values
- Specify the units being used clearly
- If possible mention the total sample size of the data set for which the graph or chart is made
- Be sparing and consistent with use of colour and fonts

Data graphics - tables

- Display numbers or words arranged in a grid
- Good for when exact numbers need to be presented
- Useful for:
 - ◆ Displaying pre- and post- results
 - ◆ Presenting correlations or comparisons

Data graphics – bar graphs

- Show quantities represented by horizontal or vertical bars and are useful for displaying:
 - ◆ The activity of one thing through time
 - ◆ Several categories of results at once
 - ◆ Data sets with few observations
- Standard deviations may be displayed using a deviation bar, extending beyond the top of the data bar

Data graphics – line graphs

- Show sets of data points plotted over a time period, connected by straight lines
- Useful for displaying
 - ◆ Any set of figures that need to be shown over time
 - ◆ Results from two or more groups compared over time, within age groups, sex differences, etc
 - ◆ Data trends over time

Data graphics – pie charts

- Show proportions in relation to a whole, with each wedge representing a percentage of the total
- Useful for displaying:
 - ◆ Component parts of a whole in percentages
 - ◆ Budget, geographic or population analysis

Interpretation

- STEPS fact sheet and reports use percentages and means to discuss and interpret data
- A percentage is a way of expressing a proportion, a ratio or a fraction as a whole number
- Mean and median are measures of central tendency
- Mean provides average of a set of values
- Because of its simplicity 'mean' used more than any other measure of central tendency
- Median is the middle in a distribution of values (used in reporting the PA data on MET mins)

Interpretation

- Look for overall pattern, then deviation from the pattern
- Look for extreme values and gaps
- Locate center and spread of distribution
- Compare graphs with same scale - look for max, min

Confidence Intervals

- Measure of precision in your data
- Method for testing subgroup differences
- The more precise your results, the more confidence you can have in them
- All sample-based surveys lack some amount of precision due to non-sampling error and sampling error
- A measure of sampling error is called standard error for a particular percentage or variable
- Precision is measured by standard error and illustrated with confidence intervals

Standard Error

- Square root of the variance of a percentage or variable across all possible samples of equal size and design
- Calculated by Epi Info which takes into account the cluster sample design
- Used to calculate confidence intervals

Confidence Intervals (CI)

- Calculated by multiplying the standard error (SE) by 1.96

$$CI = SE * 1.96$$



- Expressed as a range around the percent -

42% (40% - 44%)

42% (±2%)

- The range contains the average value of the percent which would result if all possible samples were used
- A 95% CI suggests that if 100 samples were drawn the average value of the percent would be contained in 95 of the 100 CI's

Confidence Intervals - Method for Testing for Subgroup Differences

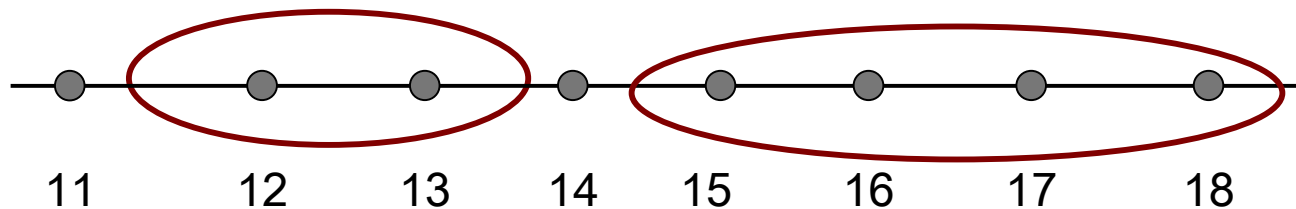
- Possible subgroups
 - ◆ Males and Females
 - ◆ Age groups (25 – 34, 35 – 44, etc)
 - ◆ Two countries, two regions, two cities
- **Question** – Are they different?
- **Answer** – Look at the 95% confidence intervals (CI's)
 - ◆ If the CI's overlap – they are not statistically different
 - ◆ If the CI's do not overlap – they are statistically different

Are they different?

- Participants who are overweight
Males - 16.4% (14.5-18.2)
Females – 12.6% (11.4-13.7)

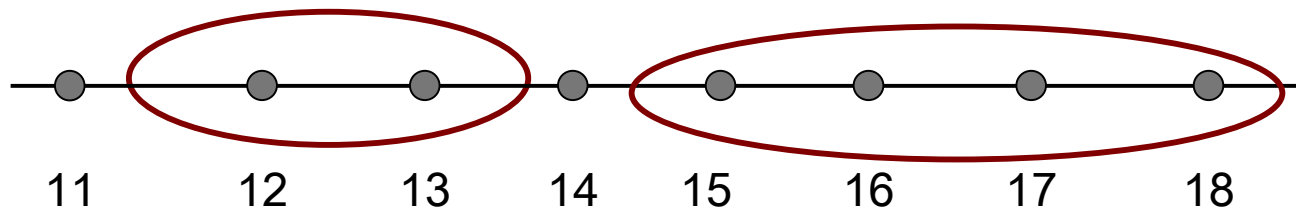
Are they different?

- Participants who are overweight
Males - 16.4% (14.5-18.2)
Females – 12.6% (11.4-13.7)



Are they different?

- Participants who are overweight
Males - 16.4% (14.5-18.2)
Females – 12.6% (11.4-13.7)



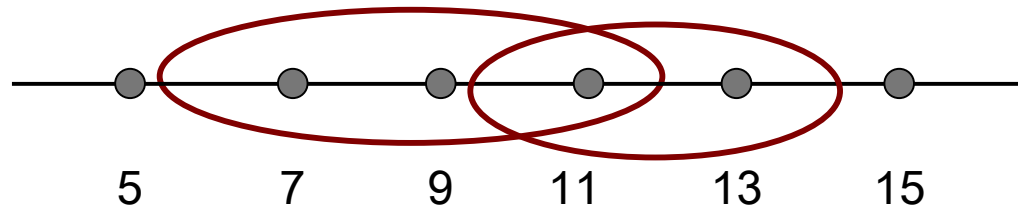
Confidence intervals do not overlap = results are different

Are they different?

- Participants who are have raised blood pressure
Males - 8.4 (5.4-11.4)
Females – 11.2 (9.2-13.2)

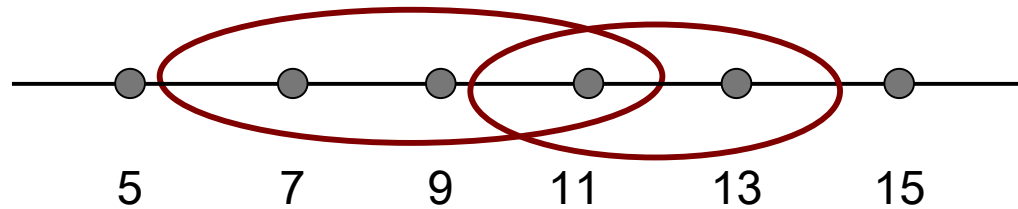
Are they different?

- Participants who are have raised blood pressure
Males - 8.4 (5.4-11.4)
Females – 11.2 (9.2-13.2)



Are they different?

- Participants who are have raised blood pressure
Males - 8.4 (5.4-11.4)
Females – 11.2 (9.2-13.2)



Confidence intervals overlap = results are NOT different