Data Presentation

Aims and objectives

- To review good practice in data presentation.
- To provide examples of different data presentation formats
- To provide sufficient information allowing students to select the presentation format most applicable to their own data

2

What are the purposes of presenting data?

1. To get the message across:
   - Good data deserve to be communicated
   - Did your experiment support or undermine your testable hypothesis?
   - Have you got something important, unique, controversial to say as a result of your work?
2. To allow others to judge our work:
   - Science relies upon peer review, must present work to allow this to happen
   - Others must be able to make their own judgements based on what is presented to them

3

How are data presented?

Data = plural!

- Figures and tables in reports and papers
- Figures and tables in oral presentations
- Figures and tables in posters
- Text, and embedded into text
- Websites

Best practice has resulted in conventions for all of these

4

General Points

- Ambiguity
  - All data must be presented in a way that says what it means.
  - Data in figures and legends **MUST** match text

- Clarity
  - Normally, raw data are not presented
  - Analysed data will be in a format that is clear
  - Use appropriate fonts, legends, symbol size, etc
  - Usually, less is more….

5

General Points

- Format
  - Look at other figures and tables to see what works
  - All figures and tables have legends
  - Figure or table plus legend should be '**stand alone**'
  - Be picky, look for errors before submission
  - Less is more ….

- Work to guidelines
  - Usually have guidelines for authorship and presentation
  - If not followed, material is usually sent back without reading!!
  - If no guidelines (e.g. talks and posters), follow best practice

6

General Points

- Accuracy
  - Should not be a problem with graphics and statistics programmes
  - Many decimal places do not make numbers more accurate (how many to use?)
- Labels and legends
  - Fully informative
  - Stand alone
  - Go for clear, visual impact
  - What Excel and MiniTab call legends are not sufficient!

7

Table or Figure?

- No hard and fast rule but…
  - Information in figures is usually easier to assimilate than in tables. For example:

	Experimental Groups	
	Treated	Untreated
Mean ($\pm$ S.E.) height (cm)	14.75 $\pm$ 0.88	9.01 $\pm$ 0.63
n	12	8
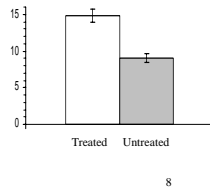
8

Tables – content and style

- Tables suit lots of data that need not be assimilated but needs to be presented.
- Tables suit non-numerical data.

Enzyme	(enzyme number)	Abbreviation	Cleavage site	Enzyme concentration ***	pH	Buffer	Reference
tide-N-glycosidase F	(EC 3.2.2.18),	PNGase F	N-linked oligosaccharides-Asn	50 U	7.5	0.5 M sodium phosphate	(Elder and Alexander, 1982)
tide-N-glycosidase A	(EC 3.5.1.52)	PNGase A	N-linked oligosaccharides-Asn	2.5 mU	5.0	50 mM sodium citrate-phosphate	(Plummer et al, 1987)
α-β-glucosaminidase H	(EC 3.2.1.96)	Endo H	Oligomannose/hybrid oligosaccharides-Asn	50 U	5.5	0.5 M sodium citrate	(Tarentino and Maley, 1974)
α-α-acetylgalactosaminidase	(EC 3.2.1.97/ 3.2.1.110)	O-glycanase™	Galβ1-3GalNAcα1-Ser/Thr	82-112 mU	6.0	10 mM sodium citrate-phosphate	(Bhavanadan et al., 1976)
:acetylgalacto-saminidase	(EC 3.2.1.49)	α-AGA	GalNAcα1-3R/Ser/Thr	625 mU	3.5-4.0	100 mM sodium citrate-phosphate	(Zhu and Goldstein, 1995)
:acetylhexosaminidase	(EC 3.2.1.30)	Streptococcus hexosaminidase	GlcNAcβ1-2*/4/6)** Man-3,6 Gal	6-8 mU	6.0	20 mM sodium citrate-phosphate	(Yamashita et al., 1981)
:acetylhexosaminidase	(EC 3.2.1.30)	JB hexosaminidase	GlcNAc/GalNAcβ1-2,3,4,6R	11-20 mU	4.0-5.0	20 mM sodium citrate-phosphate	(Li and Li, 1970)

9

Table format – some pointers

Table 1. Kinetic parameters of aminopeptidase activity in *Psoroptes cuniculi* soluble extract.

Amino acid	K_m (mM)	V_{max} (μmol min^{-1} mg protein^{-1})	Relative Specificity
LpNA	0.09 ± 0.01*	4.25 ± 0.35	47.2
MpNA	1.53 ± 0.04	8.84 ± 2.11	5.8
ApNA	1.62 ± 0.50	6.50 ± 2.07	4.0
PpNA	NH**	NH	-
RpNA	NH	NH	-

A range of concentrations of amino acid p-nitroanilide substrates in 100mM Tris-HCl pH8, 1% DMSO were incubated with *P. cuniculi* soluble extract (total mite protein 7.6µg in each assay). Hydrolysis of the substrates was monitored at 405nm at 30s intervals during the 30 min incubation period at 30°C. Reagent blanks were performed concurrently. K_m and V_{max} were determined as described in the text. Relative specificity was calculated as V_{max} / K_m (Schellenberger et al. 1991). *Data shown are the means of 3 repetitions of each assay ± SEM. The assay was repeated with a different *P. cuniculi* soluble extract with similar results.
** These substrates were not hydrolysed.

- Title (does exactly what is says on the tin)
- Column headers aligned to centre
- Units of measure in column title
- Horizontal lines only. These separate major items, not every line.
- Numbers aligned by decimal point
- Explanatory legend, says what was done to obtain these data
- Descriptive statistics and abbreviations explained below to keep table simple

10

What's wrong with this table?

Time	Protein Concentration	Enzyme activity	Body weight
0	55 +/- 5.0	2.3333 +/- 0.2	100
1	45 +/- 4.1	4.578 +/- 0.241	105
2	33 +/- 2.22	6	117
3	22 +/- 1.21	8.4 +/- 0.222	135

1. Boxes/ vertical lines
2. No alignment of numbers
3. Units missing in title row
4. No consistent use of d.p.
5. Title missing
6. Which drug, which enzyme?
7. No legend
8. Centre titles
9. Some SE's missing
10. +/- instead of ±
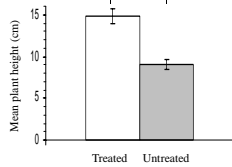
11

Figures – what sort?

12

Figures – what are you plotting?

- Proportional or absolute data?
- Continuous or discontinuous data?
- Correlative/regression data?
- Single data points or mean ± SE?
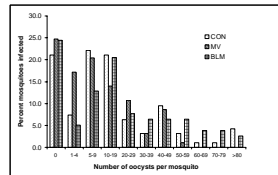- Statistical analyses on plots?
- Specialist types of presentations?

13

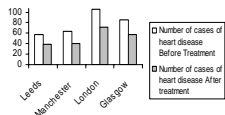# Figures – bar charts

Effect of gibberellin on plant growth

Mean plant height (cm)

Treated   Untreated

Tomato plants were treated from 5 cm seedling stage for 15 days with 0.5 mg/ml gibberellin, 2 ml per day. * = significant difference between treatments, Mann Whitney U test, H = 12.2, p<0.001

- Allows numerical and proportional data
- Easy to compare simple datasets
- Can get very cluttered with several data series.
- Good for general descriptions, basic statistics can be included
- Good for trends, frequency distributions

14

Figures – pie charts

Number of cases of heart disease Before Treatment

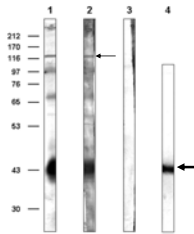Number of cases of heart disease After treatment

- Only proportional data
- Need additional legends and additional information
- Good for non-numerical data (e.g. surveys)
- Not always useful for showing trends

Before Treatment

After treatment

- Leeds
- Manchester
- London
- Glasgow

15

Figures – XY and line graphs

- Excellent for showing trends
- Line graphs
  - X axis not necessarily proportional
  - Data plotted against single variable (e.g. time, concentration
- XY graphs
  - X axis always proportional
  - Can be interdependent variables
  - Essential plot for correlation and regression
- Can show individual data points or means with standard errors
- Can take a LOT of datasets in a single graph if plotted well

16

What's wrong with this graph?

Effect of drug on enzyme activity

1. Boxed
2. Grid lines
3. Units on both axes
4. Standard errors
5. Concentration of drug
6. Which drug, which enzyme?
7. No legend
8. Can't distinguish the two groups
9. Numbers not below a 'tick'
10. Title in wrong place

17

Dealing with sequences

- **A lot of conventions** – need to be aware of these and follow them.
- Primers underlined
- * for conserved sites
- Areas of interest (in this case active sites of two closely related enzymes) boxed in
- Rarely use colour (publication costs)
- Species name or sequence code at left end of each line
- Often have position in sequence numbered on right side

18

Pictorial data: gels

- Each lane labelled with unique code or treatment
- Marker lanes shown at the edge
- Bands of interest can be highlighted with arrows; different for different bands
- Make sure that the contrast allows all relevant data to be seen
- Usually have rather detailed legends

19

Pictorial data: microgaphs
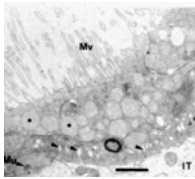
- Make sure micrographs show what you think they are showing
- **MAGNIFICATION**
- Clear labelling contrasting with the figure.
- Make sure all labels are clarified in legend.
- Often detailed legends.
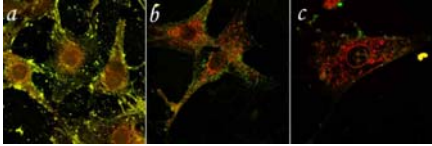
PtdIns(4,5)P₃ and PtdIns(4,5)P₃-NBD colocalize in 3T3 L1 preadipocytes. 3T3 L1 preadipocyte cells were pre-loaded with PtdIns(4,5)P2-NBD, then stimulated with 60nM insulin. Cells were fixed at different time points and immunofluorescence performed. The primary antibody was a anti-PtdIns(4,5)P₃ monoclonal. Secondary antibody has a Texas Red fluorophore attached. Panel a is time 0. Paned b is 30 seconds after insulin stimulation. Panel c is 90 seconds after insulin stimulation. The green is derived from PtdInd(4,5)P₃-NBD. Red is from the Texas Red secondary antibody. Yellow indicates colocalization of the red and green fluors. 20

Combined graphics
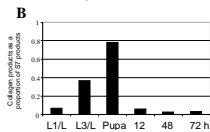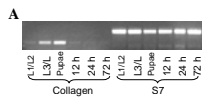
- Can combine different presentation types in a single figure
- Use to support qualitative data (top) with quantitative data (bottom).
- May need to stretch some of the rules of presentation
- Again, detailed legends will be needed for complex figures

21

Presenting data – final pointers

- Look at journals and papers to find examples that work and that you like
- **NEVER NEVER NEVER** use the excuse that a software programme would not let you format the figure the way it should be.
- Remember that the figure/table with it title and legend should be able to stand alone without other text.
- Get someone unfamiliar with the work to look it over.
- Get someone familiar with the work to look it over.
- If you see an error, correct it before someone else sees it.
- Presenting data takes time and effort: don't shy away from it
- Prepare your figures and tables **before** you write about them. (Hint: this is a very useful tip!)

22