

Review Article

Presenting graphical information

Bryant TN. Presenting graphical information.
Pediatr Allergy Immunol 1999; 10: 4–13. © Munksgaard, 1999

This article outlines the mistakes that are commonly made when data are presented graphically. Common errors when displaying data are highlighted, and preferred solutions for avoiding these mistakes are described. The preferred approaches form the requirements that authors should meet when submitting papers to the journal of *Pediatric Allergy and Immunology*.

T. N. Bryant

Medical Statistics and Computing, University of Southampton, Southampton General Hospital, Southampton, UK

Key words: data analysis; data display; statistical analysis; computer graphics; data interpretation—statistical

T. N. Bryant, Medical Statistics and Computing, University of Southampton, South Block (Mailpoint 820), Southampton General Hospital, Southampton SO16 6YD, UK

Accepted 8 December 1998

This paper follows on from a paper, published in a previous issue of *Pediatric Allergy and Immunology*, which described ways of presenting and reporting statistical information (1). Here, the graphical presentation of information is described. Graphical software, like statistical software, offers little assistance and guidance to the user on what is considered good practice. The user has the freedom to produce whatever output they desire without any feedback on whether it is correct. This freedom can lead to the production of high-quality figures that are inappropriate for the message that the author is trying to convey.

Tufte (2) states that graphical excellence in statistical graphics consists of complex ideas communicated with clarity, precision, and efficiency. A graph should induce the viewer to think about the substance rather than about methodology, graphic design, the technology of graphic production, or something else. Graphical presentation of information should, according to Tufte (2):

1. Avoid distorting what the data have to say.
2. Present many numbers in a small space.
3. Make large data sets coherent.
4. Encourage the eye to compare different pieces of data.
5. Reveal data at several levels of detail, from a broad overview to fine structure.
6. Serve a reasonably clear purpose: description, exploration, tabulation, or decoration.

7. Be closely integrated with the statistical and verbal descriptions of the data sets.

In many instances graphics can be more precise in revealing information than conventional statistical computations and, given the weak statistical background of many readers, graphical information is more likely to be understood.

Presentation of data

In most figures, the data are typically presented as either one- or two-dimensional diagrams. Often, a figure that makes comparisons between groups may employ what are essentially one-dimensional (1-D) charts presented for a series of categories.

The analysis of data and the graphical presentation of it are closely linked. The first decision to be made is: 'What type of data is being presented?' It may be qualitative data, which can be either Nominal, sometimes called categorical (e.g. allergic, or not allergic), or Ordered Categorical, sometimes called ranked or ordinal data (e.g. improved, no change, deteriorated). The alternative is that the data is quantitative data (e.g. FVC, height or weight). In the exploratory phase of data analysis, in order to decide whether parametric or non-parametric techniques should be utilized, quantitative data should be plotted to decide whether they can be considered to be normally distributed or not.

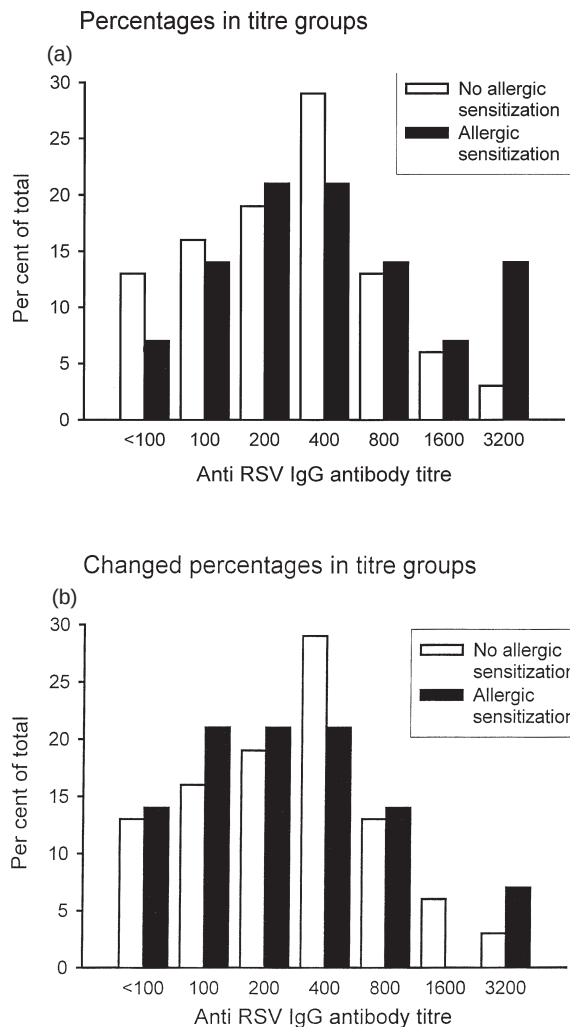


Fig. 1. (a) A bar chart showing frequency distribution for anti-respiratory syncytial virus (RSV) IgG antibody titres in 3-yr-old, non-sensitized and sensitized children. The chart was redrawn from (8). (b) A modified bar chart showing the data from (a) but with two patients re-coded into different titre groups. This illustrates the problem where small numbers are quoted as percentages. The number of cases in each group has been added to this diagram.

Often these plots are not presented in publication but they are very useful nevertheless.

Bar charts

Categorical data is often expressed in percentage terms and usually displayed as a bar chart. Typically these are drawn vertically but when the labels for each bar are lengthy, horizontal bars may be more appropriate. Percentages can also be displayed as pie-charts; however, these are not recommended as the eye is poor at making comparisons between the segments of a pie, particularly where there are several categories (segments) whose percentages are similar. When viewing percentages, the reader must be aware of

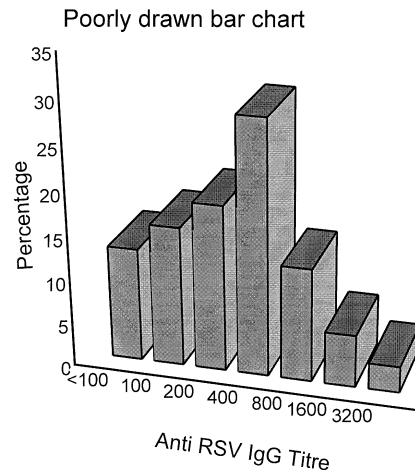


Fig. 2. This illustrates data poorly presented by the use of a three-dimensional bar chart. A two-dimensional bar chart would have been more appropriate, as used in Fig. 1(a). The data show anti-respiratory syncytial virus (RSV) antibody titre in 3-yr-old non-sensitized children (8).

the numbers used to create these figures. For example, 1 out of 2 is 50% and so is 500 out of 1000, but a change of 1 would alter the percentages by 50% (2/2) and 0.1% (501/1000), respectively. Therefore, it is advisable to present the numbers in the diagram. Figure 1 illustrates this problem where data is presented (Fig. 1a) but no counts are given. Switching the two counts from the last two bars (titres 1600 and 3200) to the first two (titres <100, 100) alters the impression drawn concerning the shapes of the distributions (Fig. 1b) but only two data values were changed. Note that in Fig. 1, titre is a categorical variable because the assay method is based on serial dilution. If the exact antibody level was measured, the variable would be quantitative and a histogram would be one way of presenting the data.

Bar charts: two-dimensional, not three-dimensional

It is somewhat unfortunate that plotting routines in many packages offer the option of three dimensional (3-D) bar charts because these can seduce the user into producing charts that are difficult to read. An illustration of this is shown in Fig. 2, using data from the allergic sensitization subjects in Fig. 1(a). It is very difficult to decide from this poorly drawn diagram what values each bar represents. The two-dimensional (2-D) chart (Fig. 1a) is easier to understand. Fortunately, such charts are usually filtered out from papers during the refereeing process but they frequently appear at conferences!

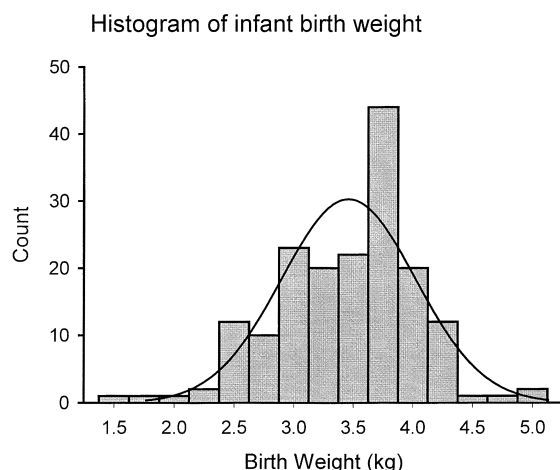


Fig. 3. A histogram of the birth weight of 172 infants, with a Normal curve with the same mean and standard deviation overlaid.

Histograms

Histograms are the most commonly used plot when examining quantitative data. Some packages can superimpose a normal curve on top of the histogram and this can assist the decision as to whether the variable is normally distributed (Fig. 3). A histogram is created by grouping the data into several non-overlapping and equal intervals, typically 5–15. The numbers of observations falling into each category are counted, see (3) for further details. The use of intervals, or bins, results in a loss of accuracy. Decreasing the interval width increases accuracy, but the underlying shape of the data can become lost and, eventually, we end up with a 1-D scatterplot. There should be no gaps between the bars in a histogram; a bar chart has

gaps. Some packages make this error, for example the Excel spreadsheet produces a bar chart, although it is called a histogram, from the Histogram macro that is available from its Data Analysis Tools. Fortunately, the user can change the gap width to zero.

Normal probability plots

These plots are rarely found in the literature; however, they are more useful than a histogram for assessing whether observations from a small sample follow a Normal distribution. The observed values of a numeric variable are plotted against the expected values for a sample from a specified distribution. If the sample is from a Normal distribution, points will cluster around a straight line. The Normal probability plot, which is also known as the quantile–quantile plot, or q–q plot, shows the observed value plotted against the Normal ordinates corresponding to the cumulative probability scores (3). Figure 4(a) shows data for a Normal probability plot of histamine PC 20 challenge. The data points do not fall on the diagonal line; there is distinct curvature. This departure from linearity indicates a lack of Normality. Transforming the data on to a logarithmic scale, and re-plotting this, improves the linearity (Fig. 4b) and suggests that parametric statistical techniques could be performed using the transformed values. Points at either end of the range show some deviation from linearity, and this is expected. However, if points in the middle of the range deviated, this would be cause for concern.

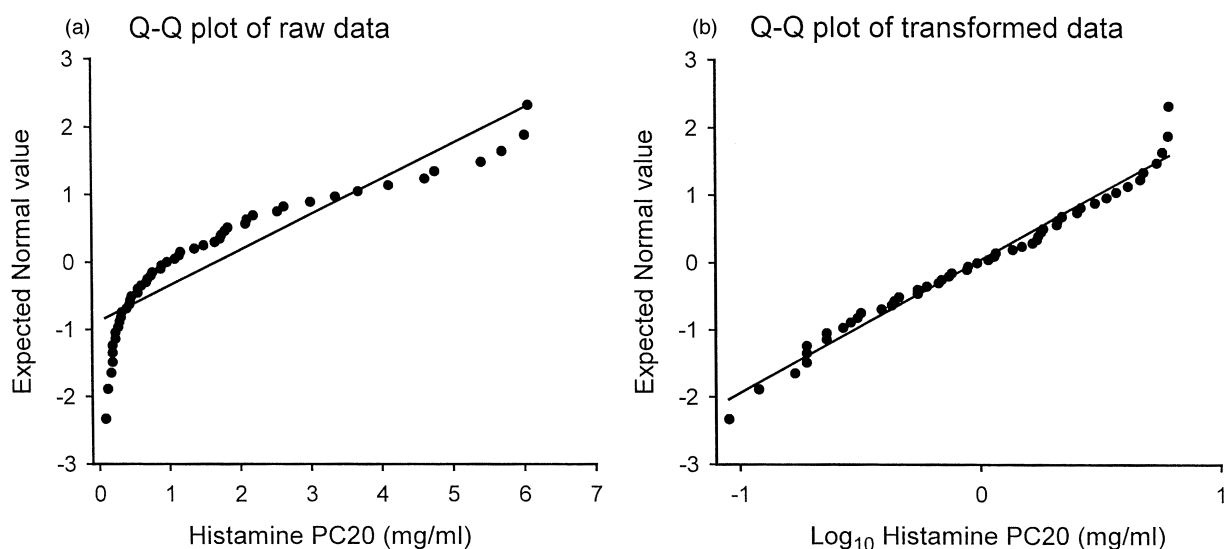


Fig. 4. (a) A Normal probability plot of Histamine PC 20 concentrations. The data points do not fall on the diagonal line, thus indicating a non-Normal distribution. (b) The same data plotted after $\log_{10}$ transformation. The data show a close approximation to a Normal distribution because the points cluster along the diagonal.

Two versions of one dimensional scatterplots of infant birth weight

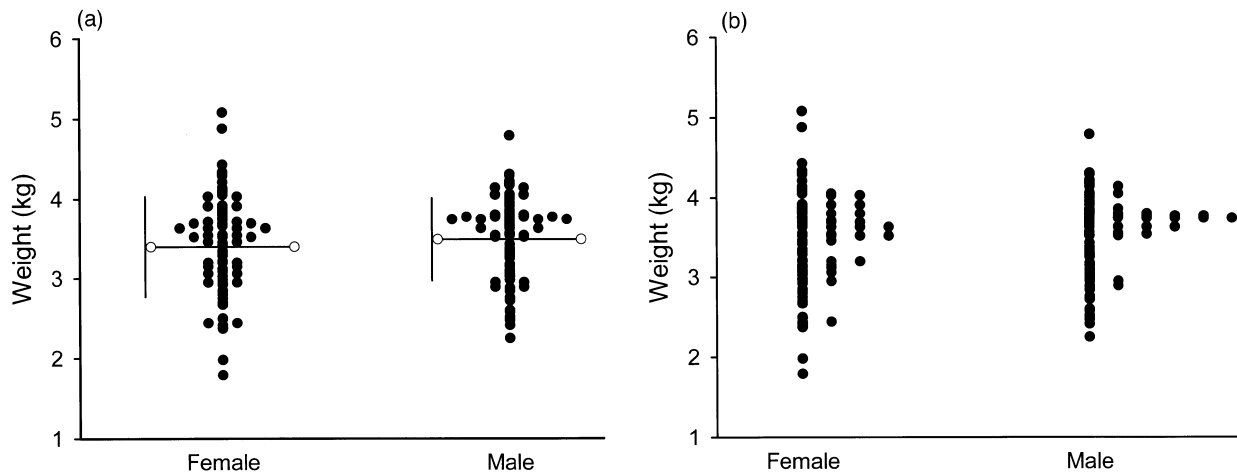


Fig. 5. One-dimensional scatter plots for comparing the birth weight of male and female babies. (a) The data has been plotted symmetrically with the mean weight for each sex drawn as a horizontal line, and plus and minus one standard deviation drawn as a vertical line. (b) Data plotted on one side of a central axis giving a clearer picture of the distribution of the data. Lines for mean and standard deviation have been omitted.

1-D scatterplots

1-D scatterplots are useful when the number of observations is not too large. Figure 5 shows two ways in which the 1-D scatterplot can be presented. Some statisticians argue that Figure 5(b) is preferable because the one-sided representation gives a better and smoother picture of the shape of the distributions. The data points may be offset to avoid overlap, as shown in this figure, or they can be drawn as overlapping data points. However, with overlapping points, if there are a large number of observations, you can end up with what looks like a thick vertical line, which does not present the information effectively. This is where a histogram might be used, but the information displayed is reduced.



Fig. 6. A Box and Whisker plot for comparing the birth weight of male and female babies. The horizontal line in the center of each box indicates the median value.

It is common for authors to include lines on the 1-D scatter plot to depict the mean and standard deviation. Although 1-D scatterplots are very useful for visualizing data, some statistical and graphics packages do not create them automatically. However, they can be created using an x-y scatter plot where x is a constant value, one, and y is the variable of interest. Some authors call the 1-D scatterplot a 'dot plot'. However, other authors and software packages use this term for a different type of chart.

Box and Whisker

The Tukey box plot, or Box and Whisker plot, provides a good summary of the data where the number of observations is large (Fig. 6). The horizontal line in the box represents the median, the lower and upper quartiles (25% and 75%). The whiskers (vertical lines) represent the range of the data, although different software packages interpret this in different ways. The SPSS package, for example, draws a whisker whose length is either 1.5 times the inter-quartile range (IQR) or the minimum or maximum value if it is smaller. Values greater than 1.5 IQRs but smaller than 3 IQRs, from the end of the box, are identified as outliers, and those that are greater than three IQRs are identified as extreme. Another package, SigmaPlot, uses the 10th and 90th percentiles as the whiskers, with values outside this range plotted individually.

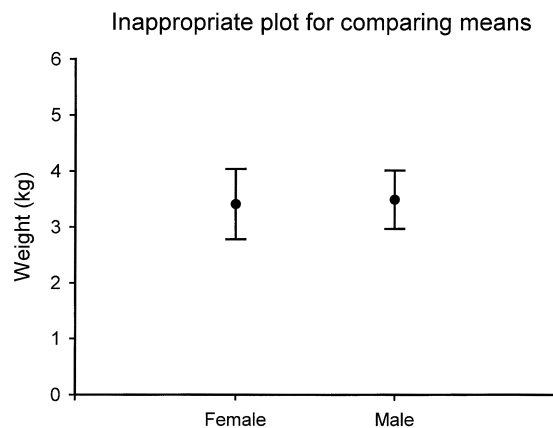


Fig. 7. An inappropriate way to display means and variation between groups. A scatter plot of mean values showing the birth weight of male and female babies. An error bar has been added but is undefined.

Displaying variation between groups

There is great confusion about how to display quantitative data comparing different groups. Cleveland (4) says that although histograms and 1-D scatterplots do a good job of displaying individual distributions, Box and Whisker plots are better for comparing distributions. A quantile–quantile (or Q–Q) plot, is a two-way scatterplot of the quantiles of one variable plotted against the quantiles of another variable. Therefore, the median of the first variable is plotted against the median of the second variable, the lower quantile (25%) of the first against the lower quantile of the second, and so on. If both variables have the same number of non-missing observations, the Q–Q plot is a two-way scatterplot of one variable against the other, after both variables have been sorted in ascending order.

Variation of the raw data

If one wishes to show the variation of the raw data then, where possible, display it using a 1-D scatterplot (see Fig. 5) and add in lines to indicate the mean, median or quartiles. If this is not possible then the box plot (Fig. 6) is the recommended alternative. Drawing means $\pm$ standard deviation (Fig. 7) is poor way of showing variation. It conveys less information and does not show outliers. Equally, the use of a bar chart with an error bar added is not a good way to represent such data (Fig. 8). Sometimes the legend for graphs, such as those shown in Figs 7 and 8, fail to define what the error bar depicts. Comparison of Figs 5 and 6 with Figs 7 and 8 shows how the 1-D scatterplot, or the Box

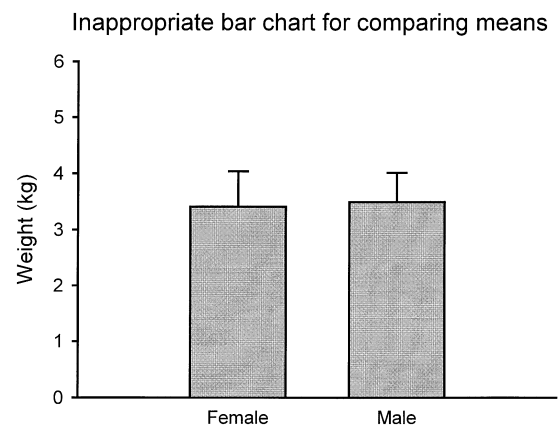


Fig. 8. Another inappropriate way to display means and variation between groups. A bar chart with an error bar added showing the birth weight of male and female babies.

Line graph incorrectly used to show change of Der P1 concentrations over time

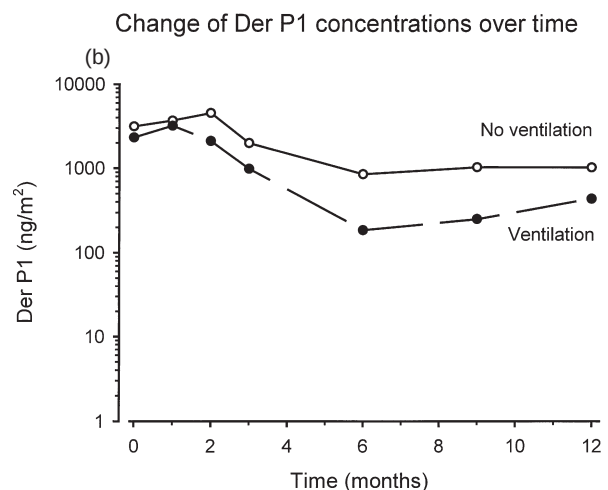
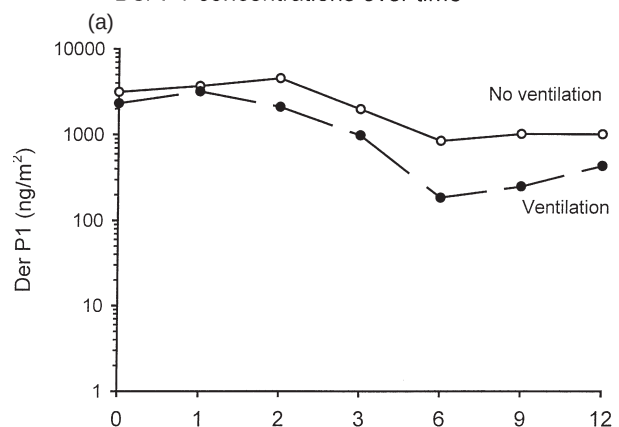


Fig. 9. (a) has been incorrectly plotted as a 'line chart'. It shows the geometric mean values for concentrations of Der P1 (ng/m^2), measured over 12 months, in houses with and without systems for mechanical ventilation with heat recovery. The choice of a line plot has resulted in each time point being treated as a category. In Fig. 9(b) the same data are plotted but with the correct selection of a scatterplot where the X-axis is a continuous scale rather than a category axis.

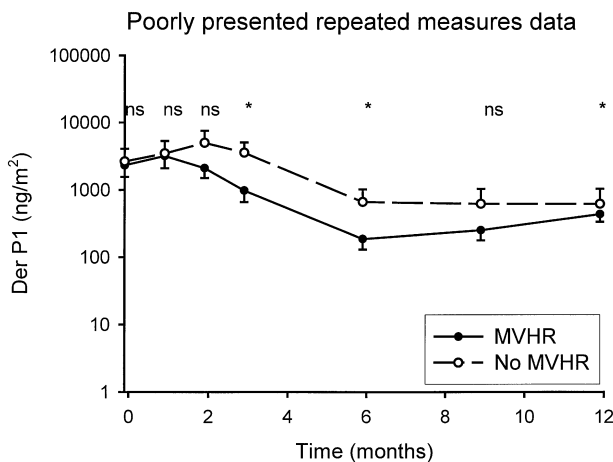


Fig. 10. A badly presented figure of repeated measures data. The inclusion of statistical significance as 'ns' and *, meaning $p < 0.05$, summarizes the result of testing for differences at each time point. This is an incorrect approach for analysis of this type of data. A summary measures approach should be used. Error bars should not be drawn.

and Whisker plot, presents far more information in the same space.

Variation of the means

If the author wishes to show the variation of sample means, it is better to use confidence interval (CI) bars to show this, as opposed to mean $\pm$ a standard error bar. Cleveland (4) has suggested the use of two interval bars: 50% CI and 95% CI. The additional 50% CI gives a middle range for the sample-to-sample variation, analogous to the Box and Whisker plot, although this has not been adopted in many publications.

Scatterplots, not line charts

A common mistake when plotting data in a repeated measures experiment is the incorrect choice of graph. Figure 9(a) shows house dust mite data, sampled from a bedroom carpet, plotted using what many packages call a 'line chart'. Data were collected at months 0, 1, 2, 3, 6, and 12 during the study but the time points are spaced equally. This is because a line chart treats the X -axis as a categorical axis; the chart has no concept of scale. Figure 9(b) shows the same graph plotted as a scatter plot. The problem is that the 'line chart' is a confusing name that has been adopted in both spreadsheet and statistical packages. It is not surprising that the user often makes the wrong choice because they want to draw a graph with lines on so the line chart seems the logical choice. The same problem applies to experimental data where the authors wish to explain results where two or more groups of subject have been exposed to increasing concen-

trations of a substance. Note that in the Excel spreadsheet it is easy to correct this error by editing the chart and changing the 'Chart Type' from 'Line' to 'XY (Scatter)'.

Repeated measures plots

Figure 10 shows an example of a diagram that often appears in the literature. Curves for two groups are drawn with the mean values at each time point plotted with an error bar. The error bar is often undefined in the text and, as in this example, to avoid clutter only the upper or lower bar is drawn. Superimposed on the chart are symbols at each time point indicating the level of significance resulting from, perhaps, t -tests carried out for each time point. This 'zipper plot', as it is termed by statisticians, is an inappropriate diagram for inappropriate analysis of this type of data, as has been outlined previously. For example, what sensible conclusions could be drawn from the study when we observe a difference at months 6 and 12 but not at month 9? Even if testing has not been carried out at each time point, the use of error bars invites the reader to make comparisons between the groups at each time point. The suggested way to display these data is to present them either as a series of individual plots using identical scales (Fig. 11) or as a series of curves on the same plot (Fig. 12). However, if there are too many data points, such plots become unintelligible. In these situations a summary plot using just the mean values, as shown in Fig. 9(b), would be appropriate (5), but the plot of mean values must present a reasonable summary of the trends seen from individual plots. The plot must not hide interesting effects because of the averaging process.

Guidelines for graphs

There are many aspects to consider when creating graphs. The guidelines below for graphs have been taken from three sources (4, 6, 7).

According to Cleveland (4), the essential points in creating a graph are:

1. Use visually prominent graphical elements to show the data.
2. Use a pair of scale-lines (axes) for each variable, and display the data in a rectangle that is slightly smaller than that created by the scale lines. Tick marks should point outward.
3. Do not clutter the interior of the scale-line rectangle.
4. Do not overdo the number, or have too many, tick marks.

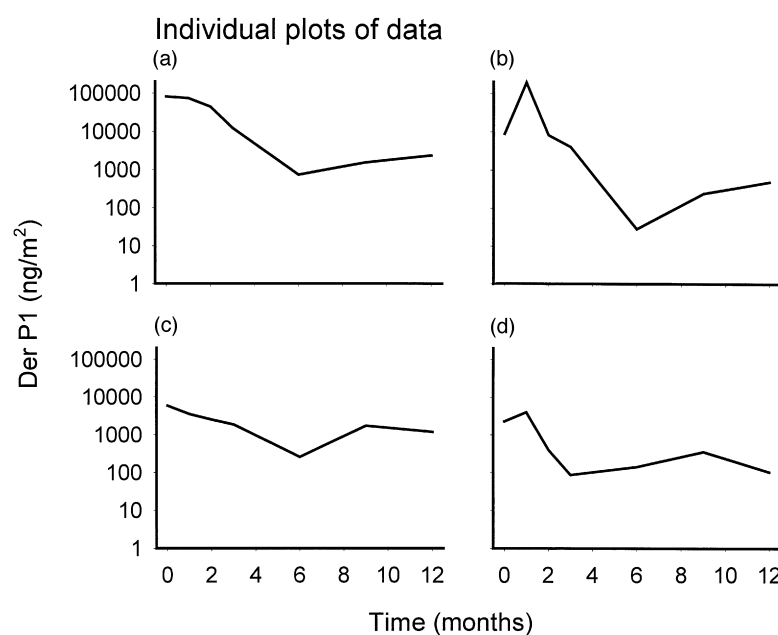


Fig. 11. Individual plots of Der P1 (ng/m^2) concentration, from the bedroom carpet of homes, against time.

5. Use a reference line when there is an important value that must be seen across the entire graph, but do not let the line interfere with the data.

6. Do not allow data labels in the interior of the scale-line rectangle to interfere with the quantitative data or to clutter the graph.

7. Avoid putting notes and keys inside the scale-line rectangle. Put a key outside and put notes in the caption or in the text.

8. Overlapping symbols must be visually distinguishable.

9. Superposed data sets must be readily visually assembled.

10. Visual clarity must be preserved under reduction and reproduction.

Axes labels

Axes should be clearly labelled with both the name of the variable and the units of measurement, using the conventions of the International System of Units (SI).

Shifting the axes

Normally, the axes should meet. However, to make the data that falls on the axis more easily visible, the axes may be offset slightly. For example, compare the initial values at $x=0$ in Fig. 9(a),(b) where the y -axis of the second figure has been slightly offset.

Starting the scales near the data

To save space, scales should be started close to the first data point. Unless it is reasonable to do so, scales do not have to start (0,0) and this is acceptable practice in scientific presentation (4, 6). To make clear that the axes do not meet at (0,0), the axes may be shifted apart slightly. Alternatively, if the plotting software supports it, the axes can be broken by // marks on the axis, which denote where the unnecessary part of the axis, between zero and the first data value, has been omitted. Axes should not extend far beyond the data. Excess white space makes the data seem unimportant and in some cases, such as a scatterplot, excess white space increases the perception of agreement between data, as shown in Fig. 3 of a previously published paper (1).

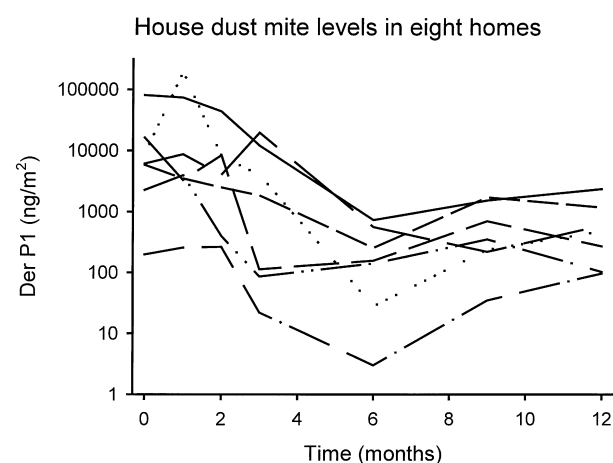


Fig. 12. Plots of Der P1 (ng/m^2) concentration, from the bedroom carpet of seven homes, against time. (The plotting software only provided seven different line types.)

Scales and scale numbers

Scales on each axis should be accurate. The purpose of tick marks is to identify the type of scale, not to identify each data point. Tick marks must be at equal and logical intervals, e.g. 0, 5, 10, 15,... and not 0, 2, 5, 10, 30. Choose the range of the tick marks to include the range of the data so that the data fills as much of the plot as possible, and avoid too much white space. Sometimes it could be helpful to the reader to present two different scales for a variable, for example to show the year of birth and age as top and bottom scales, although this is not often used in the literature.

Symbols for data points

The recommended symbols for data points according to Wallgren et al. (7) are: ●, ○, ▲, △, ■ and □. The symbols ▼, ▽, ◆ and ◇ could also be used but they can be difficult to distinguish from ▲, △, ■ and □. The authors do not recommend as it does not reproduce clearly. However, Cleveland (4) refers the use of the following symbols: ●, ●, ○, ⊙ and ⊙ in preference to ▲, △, ■ and □. Few publications routinely include papers with color figures, but black and white can be used symbolically: for example, ○ no treatment, ● treatment and, if data is grouped, then △ and ▲ can be used for no treatment and treatment, for a second group.

Overlapping data points

Several options can be used to overcome a mass of coincident points.

1. *Jittering*, where a small amount of random noise is added to the data before plotting. This could be achieved for a spreadsheet containing 10 values in columns A by using the formula: $=A1 + ((0.5 - \text{RAND}()) * \text{AVERAGE}(A1:A10)/10)$, which adds or subtracts a random amount of 0 to 5% of the mean to the original data value. Jittering is best done with circles, which tolerate this much more than squares, triangles or diamonds. The use of jittering should be reported in the figure legend.

2. *Moving*, where the overlapping points are moved by a consistent amount, as shown in Fig. 5. Unless it is obvious, the legend should inform the reader of this.

3. *Sunflowers*, where more than one point is replaced by a line for each radiating from the dot position, e.g. four points would equal a cross.

Log scales

Use a logarithmic scale when it is important to understand percentage change or multiplicative factors. Exponential growth shows that doubling is going on at the same rate when data are plotted as logarithms. Showing data on a logarithmic scale can cure skewness towards large values and enable linear trends to be displayed. Using a different logarithm base ($\log_{10}$, $\log_2$ or $\log_e$) only changes the values of the tick marks; it does not change the position of data. The difference is a scaling factor. Use $\log_{10}$ where the data goes through many powers of 10, but if the data ranges over two or fewer powers of 10 then $\log_2$ or $\log_e$ may be more appropriate.

When logarithms of a variable are graphed, the scale label should correspond to the tick mark labels. Sometimes two scales can be shown; one the original units and the other presenting the data as logged units.

Lines

If more than one line is to be displayed in a chart then the use of different line styles will be required where the lines overlap. Various permutations of dots and dashes can be employed (Fig. 13a) but approximately five lines may be all that can be distinguished. Where lines do not overlap and are clearly distinguishable, each one can be plotted using an unbroken line.

The thickness, or width, of lines is another consideration and will depend on the published size of the diagram Fig. 13(b). Different elements of a chart will use different thickness, depending on their importance. For example, a curve and axis lines are more important than tick marks and grid lines.

Patterns and shading

Graphics packages often delight in the number of patterns they can offer to the user. However, there are probably only five or six useful patterns. Patterns should be easily distinguishable from each other and restful on the eye (Fig. 13c), and not difficult to distinguish (Fig. 14). The alternative to using patterns is to use different colors on the gray scale (shading). However, it is important to choose shades that are can be distinguished from each other. Most software packages provide 16 basic colors and, from these, white, light gray, dark gray, and black can be selected. Other shades of gray can sometimes be defined if a 'custom colors' option is available. If color is used to differentiate parts of a graph, it

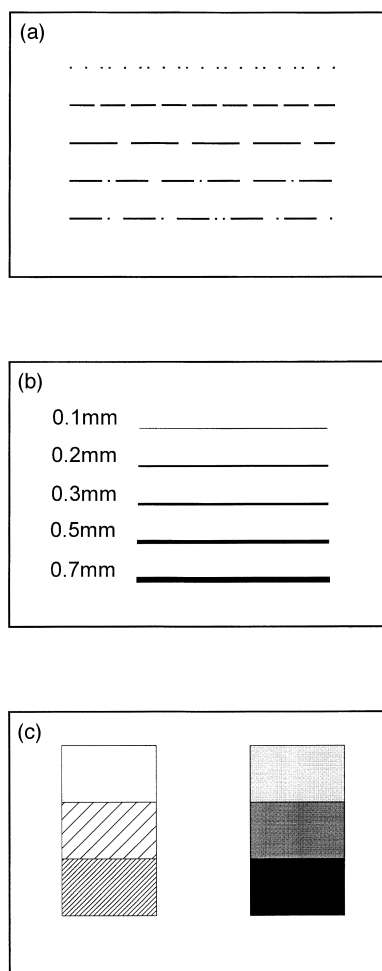


Fig. 13. Examples of (a) line patterns, (b) line widths (thickness), and (c) patterns and shading that can be used in charts.

must be remembered that these colors look the same shades of gray when photocopied.

One of the basic rules when stacking items that are patterned or shaded (e.g. in bar chart) is to put the lightest pattern or shade at the top and the darkest one at the bottom with a gradation between them (Fig. 13c). See Wallgren et al. (7) for more details of this.

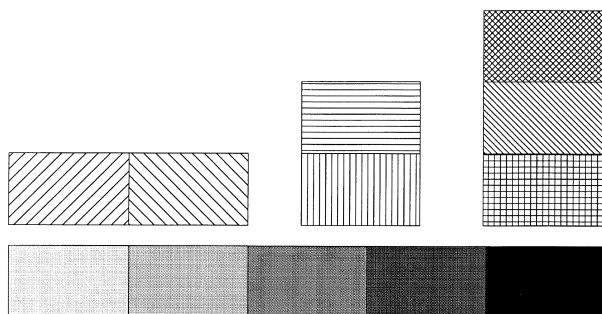


Fig. 14. Examples of poor combinations of patterns and shades for charts. If many shades of gray are used, the distinction between them is lost on reproduction.

Banking to 45°

To assist clarity, the slope of a 2-D scatterplot should be made to appear as close to 45° (or -45°) as possible (4). This is controlled by use of the aspect ratio, height of the data rectangle divided by the width. The importance of the aspect ratio has been known for many years and it affects the way in which we visually decode information. The aspect ratio of a graph is an important factor for judging the rate of change. Care needs to be taken to distinguish the true slope from that on the diagram (as determined by the aspect ratio). There is also a special case when X and Y are both measured in the same units. In this case, scales should be equal, irrespective of where the scatter plot actually lies. Departure now from a true 45° tells us something about the relationship between X and Y , especially if X and Y are measuring the same thing.

When a figure is made up of several graphs, the overall banking should be 45° . This is done by supposing that all the data are plotted on a single graph, using an aspect ratio that produces 45° banking for all these points. This aspect ratio is used for the each individual graph.

Fonts

It is recommended that a *Sans Serif* font, such as Arial or Helvetica, be used to label diagrams. Titles, if used, should be written in a **bold** font and placed above the chart, using either left or center justification.

Final size

The final size of text and numbers in a figure should be no less than 8 point type (1.5 mm), although axes numbering might require 7 point type. Therefore, the original figures should be drawn to allow for reduction during the printing process. The original figures produced for this paper occupied about one-half of a page in portrait orientation. The size of the text was: figure title 18 point, the axes labels 16 point, and legends and scale numbers 14 point. Lines were drawn with a thickness of 0.5 mm. The figures were reduced to 50% of their original size using a photocopier to assess their intended publication size.

Figure legends

As the major conclusions of research will probably be presented in graphical form, it is important to make figure legends (captions) comprehensive and informative. For a graph to

Table 1. Software packages for producing publication-quality figures

Type of Package	Name	Web site
Spreadsheet	Microsoft Excel	www.microsoft.com
	Corel Quattro	www.corel.com
	Lotus 123	www.lotus.com
Statistical	Minitab	www.minitab.com
	SAS	www.sas.com
	SPSS	www.spss.com
	Stata	www.stata.com
	Systat	www.spss.com/software/science/systat/
Scientific	Axum	www.mathsoft.com/axum
Graphics	FigP for Windows	www.biosoft.com/figpwin.htm
	Origin	www.microcal.com
	SigmaPlot	www.spss.com/software/science/SigmaPlot/

be understood correctly, there must be a clear, direct explanation of the data that are graphed and of the inferences drawn from the data. The legend should:

1. Describe everything that is graphed.
2. Draw attention to the important features of the data.
3. Describe the conclusions that are drawn from the data on the graph.

Error bars should be clearly explained. Error bars are used to convey variability in the data, but they can be misunderstood if inadequately described because they can illustrate different things.

1. The sample *standard deviation* of the data.
2. An estimate of the standard deviation (or standard error) of a statistical quantity, e.g. the mean.
3. A *confidence interval* for a statistical quantity.

Software packages for creating graphics

Many packages can be used to produce publication-quality figures. At the basic level, spreadsheets such as Microsoft Excel and Corel Quattro offer a wide variety of charts that are easy to create, yet offer a flexible editing so that one can improve on the 'instant' charts created. Statistical packages comprise the next level. These typically offer a wider range of charts for medical and scientific usage but some packages are limited in their editing and enhancement capabilities. At the top end of the range are scientific graphing packages that offer features that may not be required for routine work but provide features that allow the user to produce the best-quality charts, suitable for publication. Table 1 lists some currently available software for producing figures.

Summary

Good presentation of results is essential for publications. Inappropriate charts are still too common in the literature. When presenting data, clear vision is vital (4), and two broad strategies should be adopted when presenting graphical information: make the data stand out, and avoid superfluity.

Acknowledgments

I should like to thank David Machin, Mike Campbell and Mark Mullee for helpful comments during the preparation of this article.

Technical note. All of the figures presented in this paper were produced using SigmaPlot, version 4.0.1, running in Windows 95, and were printed using an HP LaserJet IV printer.

References

- 1 BRYANT TN. The presentation of statistics. *Pediatr Allergy Immunol* 1998; 9: 108–15.
- 2 TUFTE ER. The Visual Display of Quantitative Information. Connecticut: Graphic Press 1983.
- 3 CAMPBELL MJ, MACHIN D. Medical Statistics – A Commonsense Approach, 2nd edn. Chichester: John Wiley, & Sons 1993.
- 4 CLEVELAND WS. The Elements of Graphing Data. New Jersey: Hobart Press 1994.
- 5 MATTHEWS JN, ALTMAN DG, CAMPBELL MJ, ROYSTON P. Analysis of serial measurements in medical research. *Br Med J* 1990; 300: 230–5.
- 6 International Committee of Medical Journal Editors. Uniform requirements for manuscripts submitted to biomedical journals. *Ann Int Med* 1997; 126: 36–47.
- 7 WALLGREN A, WALLGREN B, PERSSON R et al. Graphing Statistics & Data. Thousand Oaks: Sage Publications 1996.
- 8 STRANNEGARD O, CELLO J, BJARNASON R, SIGURBERGSSON F, SIGURS N. Association between pronounced IgA response in RSV bronchiolitis and development of allergic sensitization. *Pediatr Allergy Immunol* 1997; 8: 1–6.