Australian Biochemist Students' Page
Ten Rules of Thumb for the Presentation and Interpretation of Data in Scientific Publications

*David Vaux's quest is to improve the quality of data in scientific publications. The **Australian Biochemist** asked him to provide a brief description of his 'Ten Rules'. Because of space limitations, and the need to avoid litigation, no genuine examples (or counter examples) have been included to illustrate the rules here, but Professor Vaux says he is willing to give the full presentation in person, including many humorous, shocking and libellous examples, whenever and wherever he is invited.*

Science is new knowledge gained through repeated experiment or observation. This knowledge is communicated by publishing papers, which, by convention, include not only the conclusions, but also the data upon which the conclusions are based. The data should be sufficient, and presented in such a way, that those reading the paper are able to interpret them themselves, and come to the same conclusions as the authors. The discussion should include all conclusions that are consistent with the published data.

To convince readers of a paper that a new fact has been discovered, and that the results are not just a fluke or a statistical anomaly, the experiments or observations must be repeated often enough to make it highly probable that the conclusions are correct. Sometimes all of the data are shown, but more commonly they are combined and statistics are used to indicate how they are distributed and how much confidence one should have in any inferences about them.

To be communicated honestly and convincingly and in a way that can be comprehended by those reading a paper, the data must be presented properly. To be convinced that the authors' conclusions are correct, those reading a paper must know how to interpret the data.

This article provides ten rules of thumb for the presentation of data in publications, and each rule can also be considered from the point of view of the reader of a paper, when they are trying to decide whether the data are sound and the authors' conclusions are correct.

Multiple results are often summarised by showing means with error bars. Error bars come in two classes – descriptive error bars and inferential error bars. Descriptive error bars (e.g. range, which encompasses the least and greatest value, and ± standard deviation [SD], which encompasses about two-thirds of the values) show how the data are distributed. For example, if you were told that the mean maximum December temperature in Biddelonia was 35.5 degrees, with a standard deviation of 5.2 degrees, you would know that on ~20 days each December, the maximum temperature would be between 30.3 degrees and 40.7 degrees. Range is a less useful statistic, because, unlike the standard deviation, it gets bigger and bigger the more data that are collected.

It is more common in experimental research to use inferential statistics, rather than descriptive statistics, because they allow you to make inferences about the data. The most commonly used inferential error bars are standard error of the mean (abbreviated confusingly as SE or SEM) and confidence intervals (CI) (**Fig. 1**). The mean of the sample values, plus or minus SEM or CI, gives you an idea of the region in which you can infer that the 'true' mean lies. The larger the number of samples you have, the smaller the interval is, and the greater precision with which you can estimate the 'true' mean. Knowing that the mean maximal December temperature in Biddelonia between 1960 and 2000

was 35.51 with $CI_{95\%}$ of 0.64 would be less help to a tourist than knowing the SD, but might be of great interest to a climatologist who knew that the values between 1920 and 1960 were 34.32 with a $CI_{95\%}$ of 0.62.
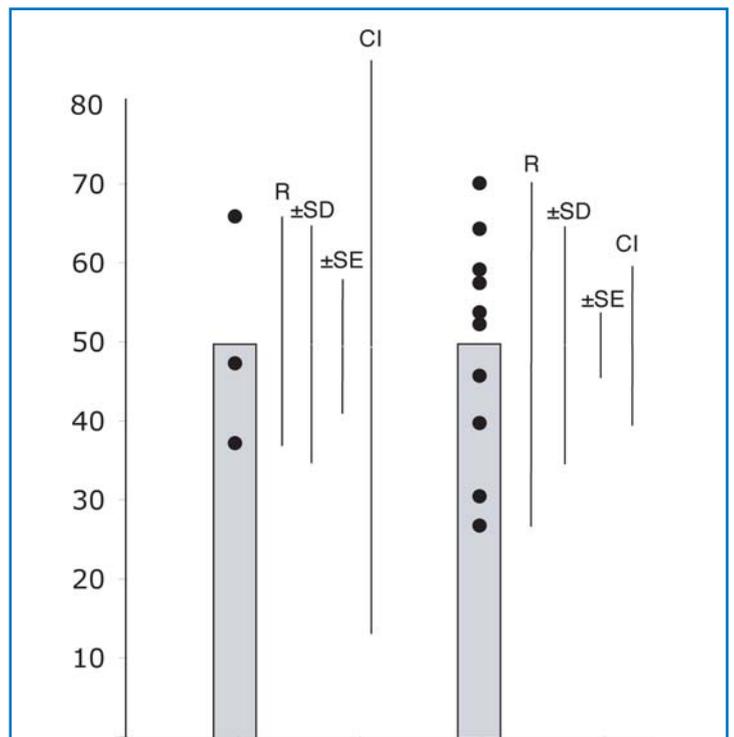
Fig. 1. Error bars look the same but indicate different things. Range (R) and standard deviation (SD) describe how the data are spread. Note that R gets larger, but SD stays about the same as the sample size increases. The 'true' mean can be estimated from the sample mean using inferential error bars standard error (SE) or 95% confidence intervals (CI). With a sample size n=3, there is a ~95% chance the 'true' mean is within the interval mean ± 4xSE, whereas if n is ten or more, there is a ~95% chance the 'true' mean is within the interval mean ± 2xSE. 95% CIs are preferred, because they mean the same thing regardless of sample size: there is a 95% chance the true mean is within the 95%CI bars whether n=3 or n=10. Notice how the interval within which the 'true' mean can be expected to lie becomes smaller as the sample size increases.

Although it is not necessary to show any error bars, and sometimes better just to plot all of the results, if inferences are to be made about drug vs. placebo, experimental value vs. control, or wildtype mouse vs. knockout, and if error bars are used, it usually is better to use SEM or CI rather than SD or range. However, because all of these error bars are depicted in exactly the same manner, the one essential thing is that if you do show error bars, you must describe them in the corresponding figure legend. This brings us to:

Australian Biochemist Students' Page

Rule Number 1

If error bars are shown, they must be described in the figure legend.

The corollary of this rule is that if you are reading a paper, and you see a figure with error bars, read the figure legend to find out what kind of error bars they are. You can only judge the statistical significance of the results if you know what kind of error bars are shown. If the authors of the paper have drawn error bars, but have not said what kind they are, they must either not understand what error bars are, not have read the paper, or both. If the co-authors (and reviewers) of a paper have not noticed that the error bars are not described, or do not realise why they must be, they have either not read the paper or are incompetent. It follows that if any of these are true, there is no point in reading the paper any further, bringing up:

Rule Number 2

If you are an author on a paper, at minimum you should have read it.

The corollary of this rule is that to determine whether the authors' conclusions are justified, you must read the paper and look at the figures.

In this brief article, I won't go into the details of estimating statistical significance, or how to interpret P values and error bars, I will just refer you to the excellent paper (if I do say so myself, though I should declare conflict of interest) by Cumming, Fidler and Vaux, *J. Cell Biol.* 171, 7-11, which you can download for free from PubMed. However, there are two more rules to do with statistics that should still be mentioned here.

Rule Number 3

N, the number of independent observations, values or samples, should be stated in the figure legend.

Although it is possible to determine statistical significance from P values, SEMs and CIs, the calculations can only be checked by the reader if n is stated. Failure to state n suggests the authors are impolite or incompetent.

## Rule Number 4

Independent results, not replicates.

If you were trying to determine whether a particular gene affected tail length in mice, you might compare the tail of a wildtype mouse to that of a knockout. Measuring each tail multiple times would allow you to put the data into an Excel spreadsheet and calculate the means and SEMs. The more often you measured each tail, the smaller the SEMs and the P values would become. In such circumstances, however, the size of the error bars would reflect on how well you could use the ruler, and the P value would have no bearing on whether the mutated gene affected tail length, because even if you measured each tail 100 times, n would still equal 1. If you do an experiment involving 4 replicate wells and repeat the experiment on 3 separate occasions, $n = 3$, not 4 or 12.

Rule Number 5

Extend axes to zero.

When looking at graphs, readers often assume the axes begin at zero. To avoid surprising, and thereby potentially annoying them, it is best to start axes at zero, unless, of course, they are log graphs. Sometimes this means there is a lot of white space in the graph, or uninteresting long columns, but if the editors feel that this makes your figure unnecessarily large, ask them to recommend how the figure should be reduced in size *after* the paper has been accepted.

If you are reading a paper and notice axes that do not start at zero, be doubly suspicious of all the results in the paper (**Fig. 2**). Are the authors trying to exaggerate or spin their results, or make non-significant differences appear significant? Mentally extend the graphs to zero and reassess before you allow yourself to be convinced. Authors who buff, exaggerate or spin their results cannot be trusted, so look with extra care at all the other results.
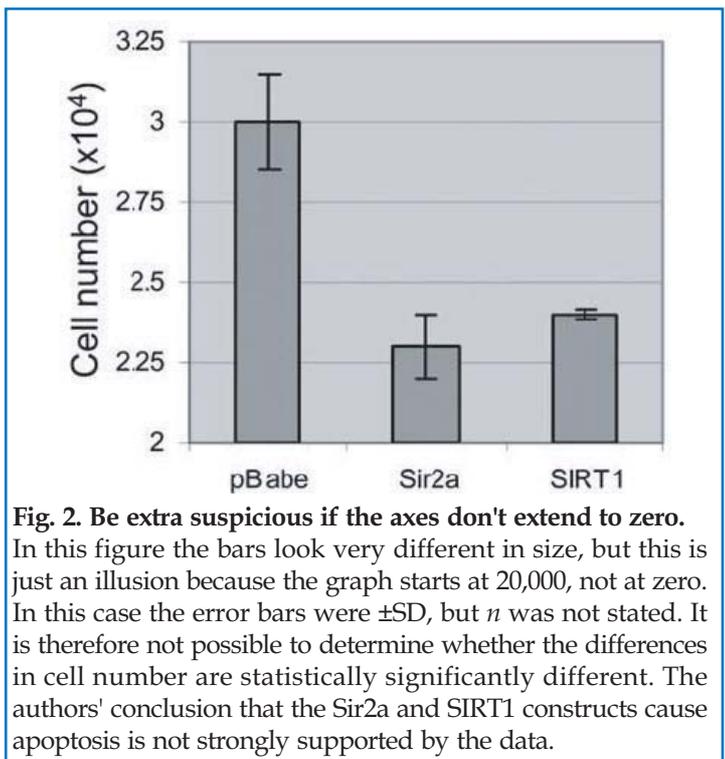
Fig. 2. Be extra suspicious if the axes don't extend to zero. In this figure the bars look very different in size, but this is just an illusion because the graph starts at 20,000, not at zero. In this case the error bars were ±SD, but n was not stated. It is therefore not possible to determine whether the differences in cell number are statistically significantly different. The authors' conclusion that the Sir2a and SIRT1 constructs cause apoptosis is not strongly supported by the data.

Rule Number 6

The contrast and brightness can be changed, but only equally across the whole image, and not so much that any band or spot, or the background, is no longer visible.

Now that everyone puts together their own figures using computer software, it is easy to modify images to make them look good. All reputable journals have instructions to authors that describe what can be done to images, and what is going too far (see Rossner and Yamada, *J. Cell Biol.* 166:11-15, for more details). In general, modifications such as changes to the contrast or brightness can only be made when they *increase* the clarity of an image; no change should be made that makes something less apparent. Changes must be made uniformly across the entire image, so that band intensities can be compared. The background should remain visible, and an artificial grey background should never be pasted in.

Australian Biochemist Students' Page

When reading a paper, it is best to download the PDF or the highest resolution images available from the publisher's web site and view them on a computer screen. Use the 'magnify' and 'select' tools to look more closely. Images can be 'grabbed' off the screen and put into photo analysis applications to look for subtle differences in the background and to look for pixel intensity distributions that can give away illegitimate changes.

Rule Number 7
Show as much of the gel or blot as you can.

Those looking at your figures will be more readily convinced if they can see most of your blot or gel than if they can only see individually cut out bands. Large blots with streaks and spots might even engender sympathy from reviewers. On the other hand, when looking at a paper, if you see a blot cut into a narrow strip, or cropping around individual bands, ask yourself why this might have been done. Were some other bands removed? How specific was the antibody?

Rule Number 8
If joining images, leave a gap or add a line where the join was made.

Figures are included in scientific publications so readers can be convinced that the evidence supports the authors' conclusions. So that readers can know whether they can compare bands on blots, they need to know whether they came from the same gel. Whenever two or more images are put into a single figure, a line or gap should be left so that it is clear that the parts had separate origins. If part of an image is removed (eg. an irrelevant or empty lane), the other parts of the image should not be joined together without leaving a line or gap, because if the join is seen, it might arouse suspicions unnecessarily. A note stating that the images are from the same blot or gel can be included in the figure legend. If images are boxed with black lines, there should be a black line wherever an image has been cut, cropped or truncated.

When you are reading a paper, look at the PDF on a computer screen and magnify the images. With practice you can learn what sort of unnatural appearances are due to illegitimate image manipulation and which are merely artefacts due to compression. If you see images that have been cut and joined, do not trust them or any others in the paper (**Fig. 3**).

Rule Number 9
Indicate molecular weights.

If you see a band on a Southern, Northern or Western blot, the easiest way to get a preliminary indication that it is the one you expect it to be is to check that it is the right size. To do this you need some sort of standard such as molecular weight markers, 18 and 28s rRNA, or Ig heavy and light chains. Just as these can provide some insurance for the experimenter against mistaken identity, they can also help reassure reviewers and readers that the band they are looking at is the real thing. Molecular weights should be indicated on all gels and blot as a courtesy to the reader. If you see a figure in which the molecular weights are not marked, there is a greater chance the band is not what it is made out to be, and the authors have deluded themselves and have been discourteous to the readers.

Rule Number 10
Include loading controls.

In order to know whether differences in band intensity reflect differences in molecular concentration, or are just due to unequal loading of the gel, it is necessary to have loading controls, such as antibodies to actin on Western Blots, or a probe for GAPDH on Northern blots. When looking at figures with blots probed for several molecules, make sure there is a loading control for each sample and look for evidence that the same blot has been reprobed. Finding that loading control bands smile upwards, whereas similar molecular weight bands smile downwards can show that loading controls have not been used appropriately.
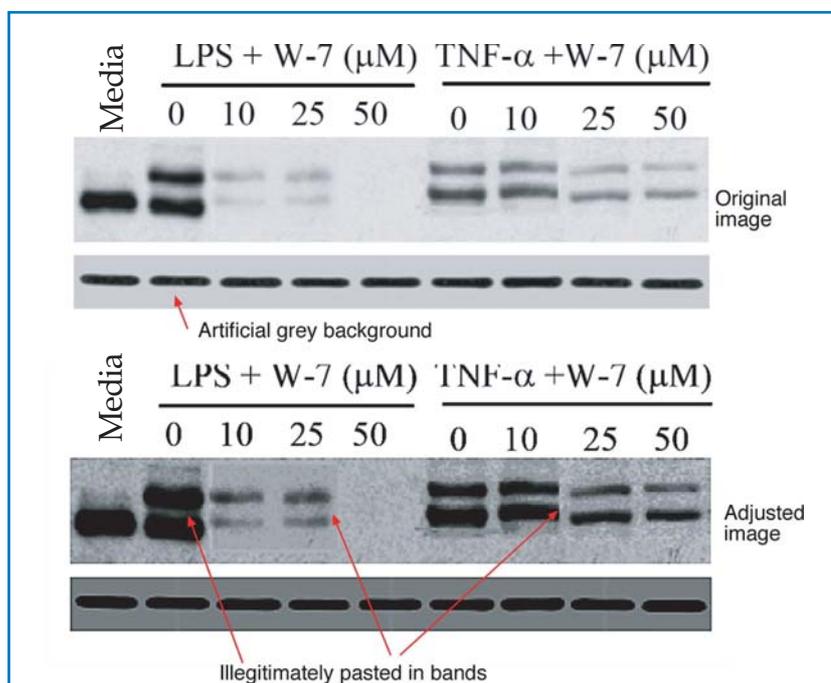
Fig. 3. Look carefully at images of gels and blots.
Magnify the images on a computer, rather than routinely printing them out. Photo manipulation software can be used to alter the contrast and measure the pixel intensity distribution to reveal otherwise hidden features. Compare the width and orientation of the bands to those in the loading controls. Check that the molecular weights are indicated.

Professor David Vaux
Department of Biochemistry
La Trobe University
d.vaux@latrobe.edu.au