Scientific Publishing Dos and Don'ts for Authors and Reviewers

Dos and Don'ts for Writing a Scientific Manuscript

From the February 2010 issue of the **AAI Newsletter.**

Based upon a presentation by Pamela J. Fink, Professor, Department of Immunology, University of Washington. Dr. Fink is a former Deputy Editor for The Journal of Immunology *and current member of the AAI Publications Committee.*

Pamela Fink

The task of writing a scientific paper can be quite daunting whether it's your first or your 50th. Learning some basic "rules of the road," however, can demystify the process and provide you with discrete steps in a manageable progression. The following are some dos and don'ts for preparing a scientific manuscript.

Step One: Decide Where to Submit

Don't: Start your paper without a clear plan for where you will submit it.

Do: Decide early in the process where you will submit your work, matching the depth and focus of your studies with those of the chosen journal. Failure to select the journal with the "best fit" can prevent the timely publication of your data and lead to much unenjoyable reformatting of your manuscript.

Consider whether your findings are of broad scientific interest or are very specialized. Is your paper describing a breakthrough finding or a more incremental advance? Research the scope of prospective journals. Such information is generally provided on journal websites. (For *The JI*, visit www.jimmunol.org/misc/infoforauthor.dtl).

Once you have decided on a journal, follow its instructions carefully for how to format your manuscript for submission. (For *The JI*, visit www.jimmunol.org/misc/authorinstructions.dtl)

Step Two: Put Your Figures into Final Form

Don't: Lose sight of your data.

Do: Start by finalizing your figures. All sections of your paper will relate directly to your figures, so putting your figures in final form is the essential first step. (See Chart 1, below) Pay particular attention to the proper use of color, the size of figures, the preferred fonts, and correct positioning of labels and text. Be sure to use an illustration program that is compatible with the journal's format requirements.

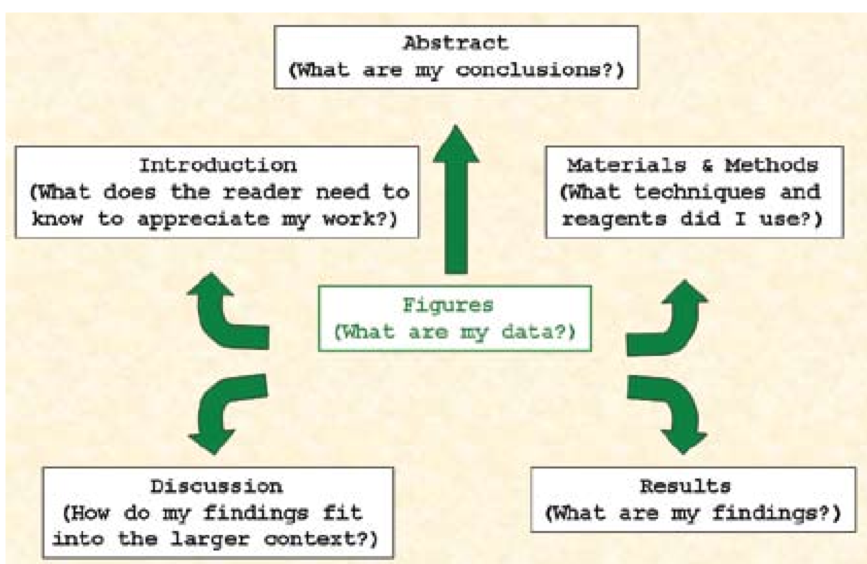
Don't: Make your reader suffer through all of your frustrations and false starts.

Do: Focus on your findings, not your missteps and setbacks. Chronological order may not be your best bet. Arrange your figures to tell the "story" logically.

Don't: Design composite figures with generic titles.

Do: Design each figure to make one clear point and state the point as the title of the figure. For example, in Figure 1A, the figure title is uninformative and the two graphs are unrelated. In Figure 1B, the line graph has been removed and the figure title now concisely states the findings presented by the scatter plot.

Chart 1—Figures Are Your Foundation



Design Figures to Make Your Point Clear

DON'T

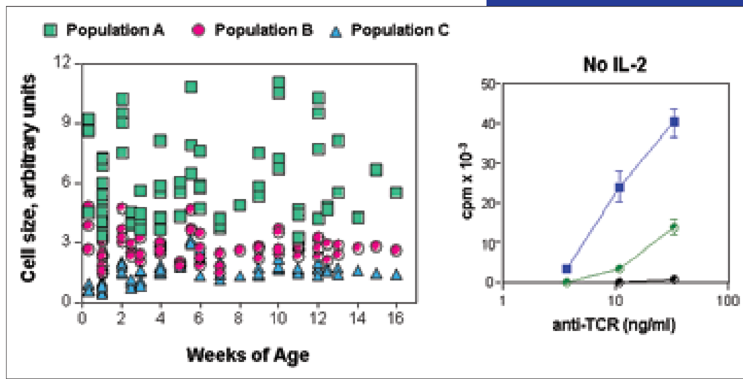


Figure 1A: Sizes of 3 cell populations and cell proliferation after TCR stimulation.

DO

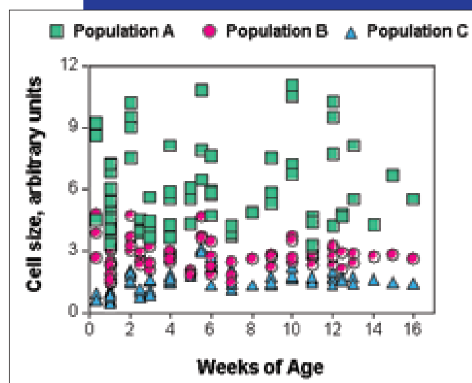


Figure 1B: Population A cells are larger than those of pupulation B and C.

Don't: Make your reader work hard to follow the flow of data in your figures.

Do: Arrange figure panels so the eye naturally follows the appropriate order. In Figure 2A, the order of the panels is illogical. Reorganized, as in Figure 2B, the panels tell the story logically. Arranging the figure panels to generate a symmetrical square or rectangle also eases interpretation. (Compare Figures 3A and 3B on page 3). Keep in mind that figure panels will be sized proportionately for publication. If a larger panel is paired with a much smaller panel, data in the smaller panel may be too small to interpret.

Don't: Cut corners when submitting your figures.

Do: Submit high-resolution figures. If figures are blurry, reviewers may interpret that fact as a sign of haste and sloppiness on your part in the lab as well as at the computer.

Step Three: Write Legends, Materials and Methods

Don't: Wait to write these sections until after writing the article.

Do: Write your figure legends and materials and methods section while you have your figures clearly in mind.

In each figure legend, briefly describe your data in the order it is presented in the figure. Legends should make figures understandable in isolation, but they should not be fully repetitive of the material and methods section. The methods section, on the other hand, should be comprehensive and provide enough detail to allow the reader to repeat the experiments you are reporting.

Step Four: Write Your Abstract and Title

Don't: Lose the focus you have achieved in finalizing your figures.

Do: Write the abstract before you write the results section. Distilling your findings to their essence at the outset will help keep you on track as you write the rest of your paper. Work with, not against, the abstract word limit set by the journal. If your abstract is too long, you are probably including too many subsidiary points!

Once your abstract is finalized, focus on the title. The title is what will or won't draw in your readers. It will be used to index your article, so inaccuracy can reduce its recall in your field. The title should be a concise label, not a descriptive sentence, and it should capture the main points of your manuscript. For example, suppose that your study identified a novel gene that modulates the activation threshold of memory T cells. The title "Studies on threshold modulation of memory T cell activation" accurately describes your study but is too general to be informative. Instead, "Novel gene x modulates the activation threshold of memory T cells" is more informative. If gene x is expressed in only a subset of T cells, namely CD4+ T

Organize Your Data Panels Logically

DON'T

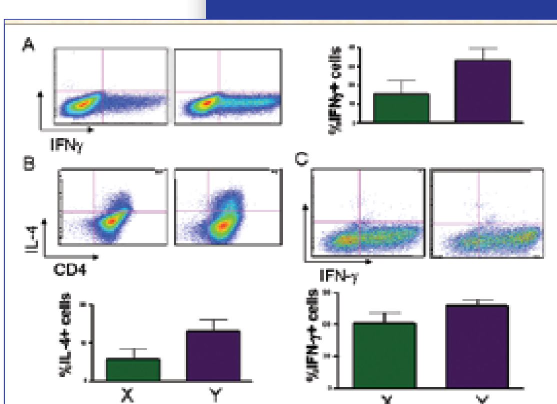


Figure 2A

DO

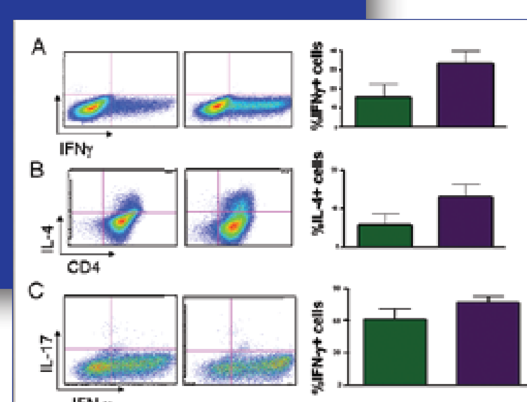


Figure 2B

cells, consider this title: "Novel gene x modulates the activation threshold of memory CD4+ T cells." If you cloned gene x from chickens and there are no homologues in other animal species, or you simply do not know, define the species "...activation threshold of chicken memory CD4+ T cells." If novel gene x is a member of a previously defined gene family, include that information as well: "Novel alphabet family member gene x modulates…." Pay attention to syntax. In the title, "The threshold of the memory T cell was studied and a novel gene x was determined to control its modulation," "its" may refer to gene x or memory T cell. The title is the lead into your manuscript, so be sure to spend plenty of time getting it right.

Step Five: State Your Results

Don't: Repeat the materials and methods section here or assume that this is the place to explain the significance of your findings.

Do: Briefly describe your data in the order in which it is presented in the figures. If possible, divide the results into subsections with subheadings very similar to your figure titles. Including a one-sentence conclusion at the end of each subsection is a huge help to readers. For example, "These data indicate that gene x constitutively associates with the T cell receptor complex." You can also use the final sentence to explain your rationale for the scientific question addressed in the next section. For example, "Our observation that decreased levels of gene x expression correlated with decreased frequencies of memory T cell activation led us to investigate whether the expression levels of gene x modulated the activation threshold of memory T cells."

Step Six: Write the Discussion

Don't: Repeat the results section or emphasize results that may be perceived as incidental findings.

Do: Place your research findings in the greater scientific context. Discuss how they advance the field and offer explanations for any data that contradict published work. The discussion should be a scholarly piece of writing. It is your opportunity to place a personal stamp on your paper. Expect to write many drafts to get it right!

Step Seven: Write the Introduction, Cite References

Don't: Fail to emphasize the relevance of your research.

Do: Define the unanswered questions that determine the focus of this research.

Use the introduction and supporting references to show the reader which work you place at the center of your field. Begin by describing the current state of the scientific field that you are investigating. Cite key original scientific reports, not just reviews.

To introduce your study regarding the threshold of memory T cell activation, you might start by reviewing key discoveries that have led to present-day understandings of T cell activation. Discuss what is known about threshold determination. Narrow your introduction to a review of the known differences between naïve and memory T cell activation. You may need to discuss models and competing theories. Be sure to establish why the questions your study answers are significant. The final paragraph can offer a brief summary of your findings.

Don't: Be careless in providing author information, assigning proper credit, or identifying potential conflicts of interest.

Do: Take time to ensure that everyone mentioned in your article, the coauthors and other contributors, are properly identified. On the title page, the names of coauthors should be written as they prefer (e.g., with

Panel Arrangement Facilitates Interpretation

DON'T

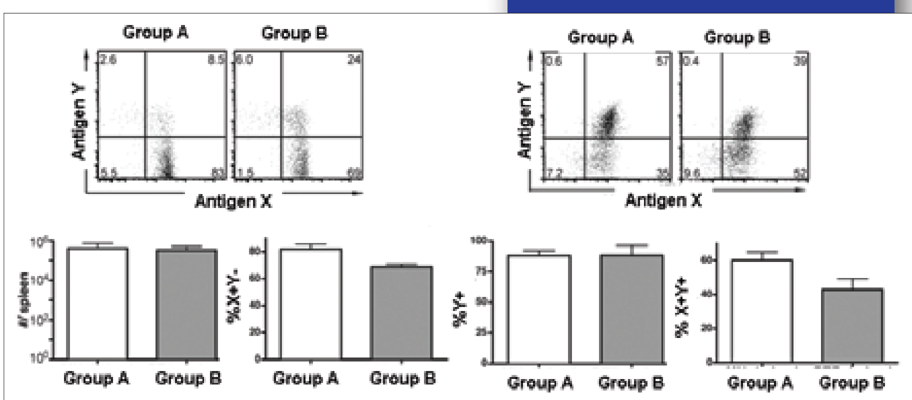
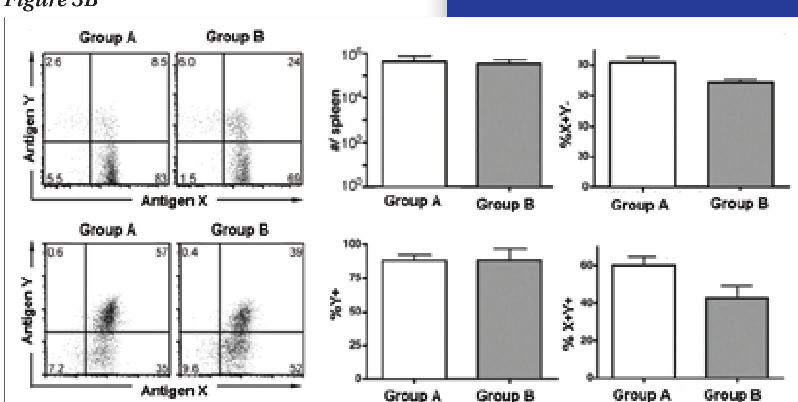


Figure 3A

DO

Figure 3B



middle initial(s) and/or formal first name) and their institutional information, properly stated. Be sure that all coauthors are in agreement on the inclusion and order of the names. In the acknowledgements section, recognize those who gave technical assistance, supplied reagents, offered helpful comments and/or suggestions. Some journals designate the acknowledgements section for presentation of information such as grant support or the disclosure of potential conflicts of interest, such as commercial affiliations, consultancies, or stock holdings. Be sure to check the journal's instructions regarding these matters.

Step Eight: Compose the Cover Letter

Don't: Treat the cover letter as a formality of superficial importance.

Do: Spend time crafting the cover letter. This is an opportunity for you to address the editor(s) and reviewers directly—to explain to them on a more personal level why you believe your work is of great importance and merits publication in their journal. By convention, the cover letter is addressed to the journal's editor in chief. Explain why you believe your manuscript is appropriate for this journal and highlight the article's main points. This information can also help the journal staff direct the manuscript to the most appropriate editor(s) and reviewers.

Don't: Simply reuse previous cover letters.

Do: Edit recycled cover letters, as necessary. If this manuscript was previously rejected by another journal, make sure the cover letter has been appropriately edited to eliminate any references to the previous journal. Be sure to modify statements about your research, if necessary, to fit the scope of the new journal.

Step Nine: Submit the Manuscript

Don't: Submit without outside input.

Do: Once you have a solid draft, solicit comments from colleagues and then revise and edit accordingly.

Don't: Submit without obtaining necessary consents.

Do: Provide a copy to all coauthors and obtain their consent to publish. Most research institutions and private companies also require in-house approval before a manuscript can be submitted, so be sure to follow your particular organization's publication rules.

Don't: Forget to broach copyright issues.

Do: Include a statement asserting that the manuscript is not currently under review or submitted to another journal. Indicate that the manuscript has been approved for publication by all authors and state that there has been no previous publication (unless in a meeting abstract) of the material within the manuscript.

Plan for Writing Your Paper:

- Decide Where to Submit
- Put Your Figures into Final Form
- Write Legends, Materials & Methods
- Write the Abstract and Title
- Write the Results Section
- Write the Discussion
- Write the Introduction, Cite the References
- Write the Cover Letter
- Submit Your Paper
- Respond to Reviewers' Comments

Once all of these steps have been completed, review the targeted journal's policies and procedures. Following the journal's instructions, upload your cover letter, manuscript, and figures and SUBMIT!

Step Ten: Respond to Reviewers

Once the article has been reviewed,

Don't: Look for bias and intent to cause pain. While conspiracies probably do exist, you are not likely in the midst of one!

Do: Begin with the presumption that the reviewer was unbiased and put substantial (unpaid) effort into understanding your work. This is, after all, the most likely scenario. Besides, assuming otherwise is simply not productive. If the reviewer misunderstood some of your data or experimental design, do not focus on assigning blame for the misunderstanding but, rather, on what you can do to prevent other readers from experiencing a similar fate. If a comment sounds snide to you, ignore the reviewer's tone and focus instead on the point being made. Don't expend any effort trying to identify your anonymous reviewers. There really is no productive point to this exercise, and in my experience as an editor, authors are rarely correct in their assumptions.

Don't: Respond to these comments hastily, even to your coauthors. Do not inflame potentially raw feelings.

Do: Draft a measured reply to the critique; make a list of changes to be made and additional experiments to be performed. Wait a day before revisiting your rebuttal letter and discussing the plan with each coauthor. In my experience, revised manuscripts ARE better than the originals. The process, while painful, does work!

The Appropriate Use of Statistics in the Biological Sciences

From the March 2010 issue of the **AAI Newsletter.**

Based upon a presentation by Pamela A. Shaw, Mathematical Statistician for the Biostatistics Research Branch, National Institute of Allergy and Infectious Diseases, NIH.

Pamela Shaw

Mathematical confirmation of research findings is an essential component of science. While biological researchers often find the application of statistics to their results a daunting task, there are steps researchers can follow to help ensure that their data are properly analyzed and presented in their publications.

Careful Planning and Design

Researchers can sometimes find themselves unable to make sense of their results because of unexpected factors arising during their study. When such frustrated researchers come to me looking for post-study help with their statistical analysis, I often find myself playing the role of a confessor, asking "What was your *a priori* hypothesis? How did you plan to test it? What actually happened?" Occasionally, by going back over what was planned and how it was executed, researchers can fit an appropriate statistical model to their data. Too often, however, a flawed design cannot be salvaged.

How can you avoid such an unpleasant scenario? Meet with a statistician early on and include a statistical analysis plan in your study design. Employing a statistician early is somewhat like taking out an insurance policy. Biostatisticians can help you design your experiments in such a way that unexpected variables or events are less likely to sabotage your study. A statistician can help you determine the number of measurements necessary for the level of evidence desired and determine the right experimental design for desired comparisons. The statistician can also help you think about possible sources of variation, such as learning effect, varying background, batch effect, edge effects, and the need for randomization. Finally, knowing up front the statistical model you will employ enables you to determine the proper controls. So, in another sense, meeting with a statistician before you embark on data collection is like practicing preventive medicine.

Avoiding Common Pitfalls

Proper randomization can maintain balance of several factors simultaneously, including unexpected ones such as changes in reagent lots or technicians mid-experiment. Randomization can be a very powerful tool, but ask yourself, "Am I really randomizing?" For example, taking mice from a cage in an apparent random fashion to deliver an experimental or control treatment might not result in a randomized sample. If you give the first half of the mice pulled out of the cage one treatment, you may have inadvertently introduced a difference between treatment arms, as more docile or less vigorous, less healthy mice will likely be taken first. Also consider sample size. A sample size should be large enough not only to detect the desired effect reliably, given the underlying variability in the population, but also to absorb the possibility of outliers and unexpected losses of study samples or subjects. The number of measurements that are required will depend on the level of evidence desired (e.g., $p<0.05$), the size difference that you are interested in detecting, and the desired power to detect that difference (i.e., the probability that the test will reject the null hypothesis given the alternative is true). Meeting with a biostatistician before you start collecting data can help ensure that your experiment has adequate power and a design that is robust enough to withstand unexpected sources of variability.

Choosing the Right Statistical Method

Generally, we use statistics to test a "null hypothesis," that is, the assumption that there is no difference between what we are comparing. For example: *The infection rate is the same for vaccine and placebo group.* Your statistical test will determine the strength of evidence present in the data to reject the null.

Selection of the appropriate test will depend upon whether your data are continuous, categorical, or dichotomous. If you are willing to make assumptions about the distribution of your data, such as whether they follow the normal distribution, you can choose a parametric test like the t-test. Otherwise, a nonparametric test like the Wilcoxon may be a better option. The proper test is also determined by sample size. With small sample sizes, exact tests may be the only option for proper inference, for many of the usual tests (chi-squared or t-test) rely on large samples or strong assumptions about the distribution of data to be valid. Finally, consider whether there is correlation in your data. Correlation, or the degree to which two or more attributes or measurements show a tendency to vary together, can come in many forms. Multiple measures on the same individual, litter effect, plate effect, batch effect, and experiment day are factors that can induce correlation, to name a few. Calculations ignoring correlation or clustering can result in the wrong conclusions.

Drawing the Right Conclusion

Much confusion for scientists surrounds the question of whether to use standard deviation (SD) or standard error (SE) when reporting results. SD is a measure of the data spread or the variability in a population; SE is a measure of the uncertainty in your statistic, or of, say, how precise your measure of the mean is. Report the one of these two quantities that is of greater scientific interest, given your type of experimental data.

Another occasionally problematic aspect of statistics for scientists is the interpretation of the "p-value." The p-value is the probability of observing, just by chance, results at least as extreme as those in your data, given the null hypothesis is true. If you have a small p-value, chances are low. For example, if you get a p-value= 0.03, there is only a 3 percent chance of seeing data this extreme or more extreme simply by chance. By convention, p-values of less than 0.05 are considered strong evidence against the null and the result is declared statistically significant. Note, though, that "not significant" is not the same as "not different." You cannot claim to have proven there was no difference unless the study was designed to test for equivalence. The lack of statistical evidence for a difference may be due to a true difference being much smaller than the study was designed to detect or it may simply be due to bad luck.

It is useful to provide confidence intervals (CI) in addition to p-values for your principle statistics of interest. The CI is an interval of plausible values for the parameter of interest. The usual 95 percent confidence interval for the mean is the mean ± 1.96 SE. Confidence intervals provide more information than p-values provide. They provide an estimate of both the magnitude of the effect and the uncertainty in your estimate. Statistical significance is not the same as scientific significance! With lots of data, small (biologically unimportant) differences can be statistically significant. A p-value provides information only on statistical significance, whereas a confidence interval can be used to judge both significance and scientific relevance.

Reporting Your Results

Within your manuscript, all graphs and tables should be self-explanatory. Graphics should reflect your design and be informative. Key summary statistics can be useful to display on the graph and should be clearly noted in the legend. For example, in the figure legend, be sure to explain whether error bars denote SD or SE, give p-values, and/or CI when appropriate. Any omitted data points should be noted.

The way you display your data is extremely important. The wrong type of display can be misleading and can obfuscate your results. When you have binary data (yes/no, dead/alive, success/failure, positive/negative), a 2x2 contingency table can be useful for both paired and unpaired data. For continuous data, if the data are unpaired, Figure 1 is an appropriate graph, while the upper graph in Figure 2 is a more appropriate graph for paired data. As Figure 2 illustrates, the significance level can be very different for paired and unpaired data. Therefore, it is important to graphically represent the correlation in the plot. When presenting a plot of data with many data points, a box plot with bars to indicate the summary statistics, such as median, inter-quartile range (IQR), and range, can be a more useful representation of your data. If there are not enough points (fewer than 10-15 points), box plots may be misleading, as they emphasize summary statistics, such as the IQR or minimum/maximum. These values will generally be very poorly estimated with only a few data points. Take care to match your data with the correct data display so that your results are easily interpretable and clearly substantiate your claims.

Summarizing Your Methods

In the materials and methods section of the manuscript, be sure to describe how you collected your data and the statistical methods of analysis used. If a non-standard statistical method is used, it should be fully explained or adequate reference provided. If there is a fair amount of data processing, such as normalizing the data or excluding data points, this should also be explained. In the methods section, be sure to

- Describe the key attributes of the population/material studied, such as the
  - Size, demographics
  - Nature of the control group (matched? historical?).

- Describe the key aspects of your experimental design, such as whether
  - Multiple treatments are employed. If so, describe your method for assigning different treatments (randomized? blocked?)
  - Experiments were done on different days
  - Assays were analyzed in batches
  - Fresh versus stored samples were used.
- Describe the type of quality-control assessments that you used

If your data analysis has involved several statistical tests, your section should describe each of the tests that you performed.

Excluding Data from Results

Avoid throwing out data points. Any data exclusions should be reported in the manuscript and be subject to peer review. Realize that the decision to exclude data is often influenced by expectations of what the result should be. Know, too, that excluding data introduces the risk of destroying randomization, and that omitted data can arouse suspicion about the integrity of your data. It is best to

- Keep outliers in the analysis and choose statistics less influenced by outliers (e.g., rank methods). This is generally preferable to throwing out data.
- Have a robust design.
- Report excluded data. If some data are excluded, calculate and report results with and without exclusions for comparison. (For example, present both the intent-to-treat and per-protocol analyses.)
- Describe your reasons for exclusion of any data.

Summary

Involve a statistician in the planning stages of your study to ensure that you can make sense of your results. At the completion of your experiments, consult again with a statistician to select the correct statistical method(s) for analyzing your data. In your manuscript, be sure to describe how you acquired your data and report the methods of statistical analyses employed. For the presentation of your results, choose data displays that properly represent your data and include supporting statistics. Finally, when you are set to submit your completed manuscript, consider seeking a statistical review of your interpretation of the data. Presenting a sound statistical summary of your findings will best communicate the scientific importance of your work.

Figure 1—Data are shown by group. The group means and error bars (± 1 SD) are denoted by horizontal lines.

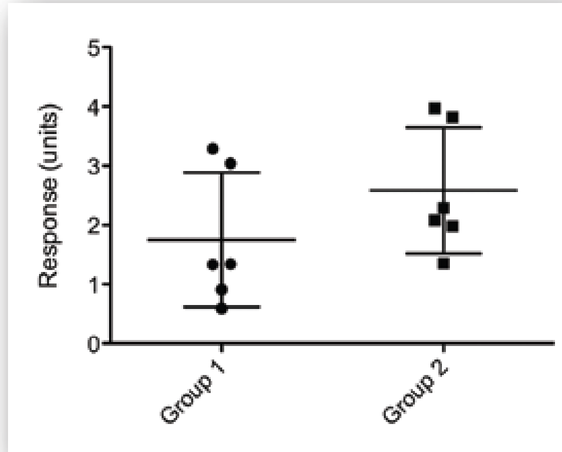


Figure 2—Below are two plots of the same data reflecting different designs. The mean difference and p-value from the associated Wilcoxon test are given.