Presenting Statistical Methods and Results

Our analyses are only as good as our ability to communicate them. We need to develop skills at comprehensively and efficiently presenting statistical methods and results to communicate our findings to readers.

Presenting Statistical Methods

You obviously need to provide sufficient information on all methods for the reader to be able to interpret your findings, but here we will deal only with statistical methods.

Sampling. Explain to the reader what your sampling space was (i.e. the statistical population from which you want to draw your inferences). Were there any limitations or restrictions on your sampling space (e.g. only live trees > 5cm dbh)? How were samples collected? This does not refer to practical issues like "..using a net...". This refers to whether samples were collected randomly, systematically, arbitrarily... Was there any stratification to the sampling? What was your unit of replication? Was there any sub-sampling? How many treatments were there and how were they assigned? You should then provide a *post mortem* on your sampling design. Were there any unplanned incidences that cause you to deviate from your initial sampling plan. For example, "We randomly sampled seedlings from each of 5 sites that varied in the number of years post-fire. No seedlings were found in the one site that was most recently burned so we present results from the four sites for which we were able to collect data on seedlings."

> **Data versus Datum**. In my world, the word 'data' is plural. The singular form of data is 'datum'. So while datum is and data are, data ain't never is! That said, in many circles it is now acceptable to use data as a singular noun. I don't agree with this, but language evolves and I am not in charge.

Data Manipulation. Provide an explanation for how raw data were manipulated to derive the data that were actually analyzed. This might involve some sort of simple or complicated calculation. It might also involve the exclusion of some data for biological or statistical purposes. This is also where you would indicate whether you transformed any of the data prior to analysis (e.g. log transformation). If so, are you presenting values on the raw scale or on the transformed scale? Did you do anything else to the data before analysis (e.g. averaging within groups, or perhaps standardizing). If you excluded any outliers

you MUST report this in your methods. Failing to report the exclusion of outliers is unethical.

Analyses. Explain how you analyzed your data. Especially if you used some complex models it can be useful for non-expert readers to start with a general and non-technical description of basically what you did and why. For example, "In order to identify environmental factors affecting the abundance of deer mice in Algonquin Park, we fitted a statistical model that included six environmental measures as predictors of the density of mice." Be clear on which statistical approach you used (general linear model?), which predictors were categorical and which were continuous. Also, if you are using a mixed effect model then be sure to specify which factors were considered to be fixed effects and which were random effects. We will discuss fixed vs. random effects in more detail later. For more complex statistical models you might find it helpful to explicitly write out your statistical model.

How did you assess the assumptions of statistical approach that you used? If you performed a transformation to your data because one of the initial assumptions were not met in a preliminary analysis, did this transformation actually work? That is, did the new model satisfy the various assumptions of the approach?

It is both useful and advisable to make specific reference to the statistical package that you used to analyze your data. This is also true for packages within R that you might make use of. Although it rarely occurs, there are instances where particular software packages (commercial or otherwise) were found to contain errors. Nevertheless, do NOT let references to specific packages (e.g. I used proc MIXED in SAS) replace a clear description of what you actually did. The details of a specific program will not make sense to readers who use a different program. Canned packages will be long forgotten in 50 years, but hopefully your research will not. You also want someone to be able to replicate your research in whatever software package they choose.

An Example of Statistical Methods

A generalized linear model (GLM) with a logit link function and Poisson error distribution was used to analyze the relative abundances of all EMF species to determine whether there was a significant effect of CO_2, block (which represents differences in mean potential net N mineralization rates), or a $CO_2 \times$ block interaction. Fisher's exact tests were conducted for the four most abundant taxa to compare their frequencies between ambient and elevated CO_2 plots. Significance values were adjusted using the Bonferroni correction to account for multiple comparisons. GLM analysis was performed using R version 2.0.1 (R Development Core Team 2005). All other analyses were carried out using JMP version 5.1 (SAS Institute 2003).

From Parrent et al. 2006. CO2-enrichment and nutrient availability alter ectomycorrhizal fungal communities. *Ecology*, 87: 2278-2287.

Presenting Results

The party line is that we only report results in the results section and that any discussion or interpretation of the results is left for the discussion section. This is fine, but some small interpretation and discussion will help the flow and readability of the results section. For example, "In order to make sure that treatment and control plots did not differ prior to our experiment, we tested…".

Rule #1. Make BIOLOGY and not STATISTICS the subject of your sentences. I will repeat - make biology and NOT statistics the subject of your sentences! Although this website is about statistics, I will most vigorously argue that we are in the business of understanding biology and that we simply use statistics as a tool to aid our understanding. Stats are the means to an end, not the end themselves! So be sure to do more than report statistical results. Your job is to tell a biological story that is supported with statistics.

For example, replace...

"ANOVA indicated significant differences among treatment means (stats). Subsequent post hoc tests revealed that significant differences occurred between groups A and B (stats) and A and C (stats) but not between groups B and C (stats)."

with...

"Plant biomass differed significantly among treatments (one-way ANOVA: stats). Biomass was highest in treatment group A (mean ± se = …; n = x), followed by treatment groups B (mean ± se = …; n = x) and C (mean ± se = …; n = x) respectively. Post hoc tests revealed that the biomass of plants in group A was significantly higher than both B and C (post hoc test: stats), which did not differ from one another (stats)"

An Example of Statistical Results

In the southwest Yukon, spring temperature has increased by nearly 2 °C (regression: $F_{1,25} = 3.6$, $p = 0.07$; $b = 0.074 \pm 0.039$ °C yr^{-1}) and there has been no particular trend for precipitation over the last 27 years (regression: $F_{1,25} = 0.5$, $p = 0.49$; $b = -0.6 \pm 0.9$ mm yr^{-1}). Over the past 10 years the average number of cones available over a female's lifetime has increased by over 35% (figure 1*a*; $F_{1,8} = 2.2$; $p = 0.056$; $b = 0.053 \pm 0.024$ cone index cohort^{-1}; $n = 10$ cohorts). During this same 10-year period, mean lifetime parturition date of female squirrels advanced from *ca.* 128 days from 1 January (8 May) for females born in 1989 to *ca.* 110 days from 1 January (20 April) for females born in 1998 (figure 1*b*; $F_{1,8} = 5.9$; $p = 0.0003$; $b = -2.02 \pm 0.34$ days cohort^{-1}; $n = 10$ cohorts). This represents a change of over two weeks in just 10 years or *ca.* 6 days generation^{-1} (figure 2).

From Réale et al. 2003. Genetic and plastic responses of a northern mammal to climate change. Proc. Roy. Soc. L., B, 270: 591-596.

Note that this is laziness and not self-promotion. It is easiest for me to grab text from my own papers!

What Statistics to Include in your Results Section?

The specifics of what you do or do not report will depend on the particular statistical test that you used. There are a few general components that should be reported for any test:

1. Parameter of interest (e.g. means, median, slope, effect size) with units
2. Uncertainty/variability (sd, se, 95%CL, IQR)
3. Sample size (with units where ambiguous)
4. Test statistic (e.g. $F_{1, 32}$, X^2, t_{12})
    - Remember to report the associated df
    - Don't need many sign. digits. t = 13.41 versus t = 13.42 will have no effect on the resulting P-value.
5. P-value
    - Avoid P > 0.05. Better to provide a more precise P-value if the journal style will let you.
    - Be as precise as possible when 0.2 > P > 0.001
    - Two decimal places is usually sufficient
        - i.e. P = 0.249 is not needed -> P = 0.25
    - P is always > 0 and < 1. P is never = 0 or 1.

Significant Digits

When reporting data it is a good rule of thumb to report standard errors to two significant digits. Report means or other parameters to the same precision as long as this does not exceed the precision of the original measurements.

In my view it is never enough to simply report P-values. These on their own do not provide any information on which statistical test was used (in a glm are you referring to the significance of the parameter - t statistic - or the significance of the factor - F statistic?). Including information on the degrees of freedom also help to make clear what the unit of replication was that was analyzed. For example, in a study of 10 plots where each plot contained 10 quadrats that could have sampled 10-100 individuals we can tell a lot about what was

exactly analyzed in the model by looking at the degrees of freedom. If the error df = 5 they they probably used plot means or somehow accounted for the sub-sampling. If the error df is 9543 then they used each individual as a unique replicate. All of this should be clear from the methods, but including df helps to make sure that this is clear.

Should I Report my Data in Tables or Figures?

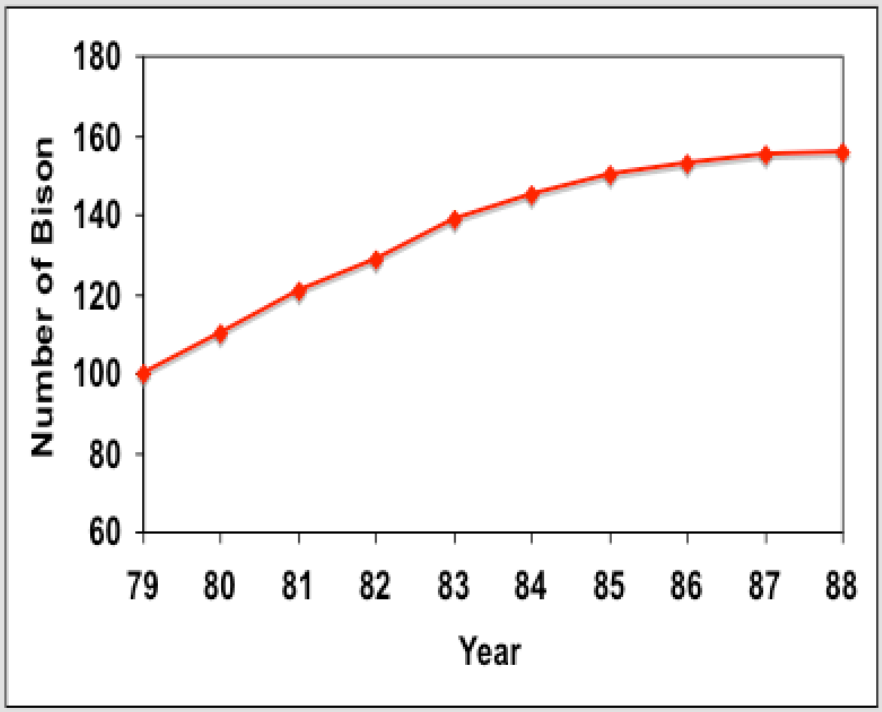
First of all, good for you for knowing that this is an OR question! Data should be reported in tables, figures or in the text but not in multiple places. Since space in journals is limited tables and figure need to be justified. Generally, tables and figures are the main means by which your story is told. Tables are usually more effective when the actual values have some importance, whereas figures are often better at showing more general patterns in the data. Overall, I would also say that figures have more impact than tables (think about how many figures you see in oral presentations versus tables). So choose your figures wisely! This will represent your story.

Formatting Tables

Data are easier to compare down columns rather than across rows. This means that tables are better oriented such that rows represent categories across which you want to compare and columns represent variables of interest. It is also often effective to organize a table such that predictors are in the left-most columns and the response is in the right-most column. Here are some other tips when preparing tables:

1. Align decimal places within a column
2. Tables need to be able to stand alone. Make sure there is enough detail in the title such that the table can stand on its own.

Year	Number of Bison
1979	100
1980	110
1981	121
1982	129
1983	139
1984	145
1985	150
1986	153
1987	155
1988	156



Whether you think this table is more effective or the figure is more effective will probably depend on whether you think the actual values matter more than their temporal trend.

Compare these two tables.

The first table is oriented with comparable data in columns.

The second table is oriented with comparable data in rows.

Table 1. Spring density, trapability, dates of first and last parturitions, and length of breeding seasons for *Peromyscus maniculatus* populations supplemented with high energy food, high protein food and unsupplemented populations in the Kananaskis Valley, AB

Year	Treatment	(*n*)[a]	Spring density (individuals ha^{-1})[b]	Trapability[c]	Date of first parturition	Date of last parturition	Length of breeding season (days)
1996	Control	2	6·57	0·69	163·5	229·0	65·5
	Energy	1	15·79	0·82	—[d]	191	—[d]
	Protein	2	8·50	0·80	147·5	228·0	80·5
1997	Control	2	13·46	0·74	132·5	215·5	83·0
	Protein	3	10·69	0·72	138·0	216·7	78·7

Note: Data are presented as means of all populations within each treatment.
[a]Number of monitored populations.
[b]Number of over-wintered individuals known to be present in population during the first 2weeks of trapping.
[c]Number of captures (minus first and last) summed for all individuals in the population divided by the number of potential captures (minus first and last) summed for all individuals in the population.
[d]Young-of-the-year were present at the start of trapping in the energy-supplemented population.

TABLE 1. EMF community response to elevated CO_2.

	Ambient CO_2 plots				Elevated CO_2 plots			
Parameter	1	5	6	Mean (SE)	2	3	4	Mean (SE)
No. samples	50	64	81	65 (9.0)	70	73	73	72 (1.7)
Phylotype richness	18	17	26	20.3 (2.8)	25	17	24	22 (2.5)
Mean rarefied richness		14.7 (12.2, 17.1)†	20.6 (17.5, 23.8)†	17.8 (1.7)	21.7 (19.0, 24.4)†	15.0 (12.6, 17.2)†	20.3 (17.5, 23.1)†	19.0 (2.0)
Unique species (%)	33.3	29.4	30.7	31.1 (1.1)	32	47.1	45.8	41.6 (4.8)
Rarefied Shannon diversity		2.16 (2.0, 2.3)†	2.68 (2.5, 2.9)†	2.48 (0.16)	2.85 (2.7, 2.9)†	2.09 (1.9, 2.3)†	2.66 (2.5, 2.8)†	2.53 (0.20)

Notes: Richness and diversity data were rarefied to a samples size consistent with plot 1, the plot with the smallest number of samples.
† Numbers in parentheses are 95% CL.

3. Most publishers do not allow vertical lines in tables.
4. Notes go below the table and can provide additional details on measurements.

Making Effective Figures

Use lines graphs for continuous predictors. It is usually better to adjust your scale or orientation to emphasize your result, although some journals will require you to have your y-axis start at zero for a variable that can take on a value of zero. This will often make the result look less impressive. Lines

Examples of a bar chart and a chart with point estimates.

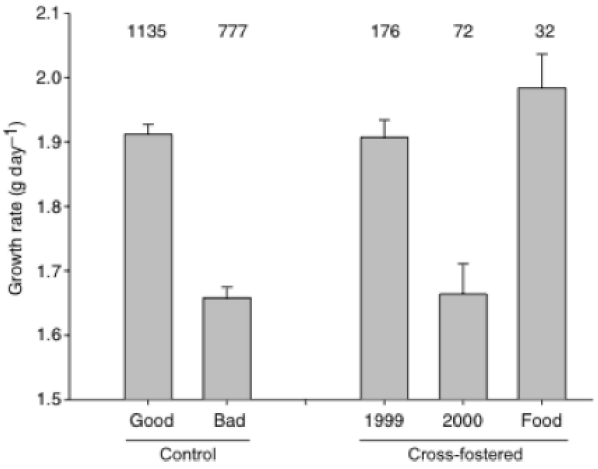


Fig. 1 Rates of growth in body mass (g day^{-1}) of nestling red squirrels (*Tamiasciurus hudsonicus*) in good and poor conditions for growth. Control growth rates were measured on nestling squirrels born in this population over the past 16 years (1987–2002). Growth rates of cross-fostered offspring in 1999 and 2000 correspond to good and poor conditions for growth, respectively. Some cross-fostered offspring received a food supplementation (*Food*) in 2000. Values are presented as means + one standard error. Sample sizes (number of juveniles) are indicated above each bar.

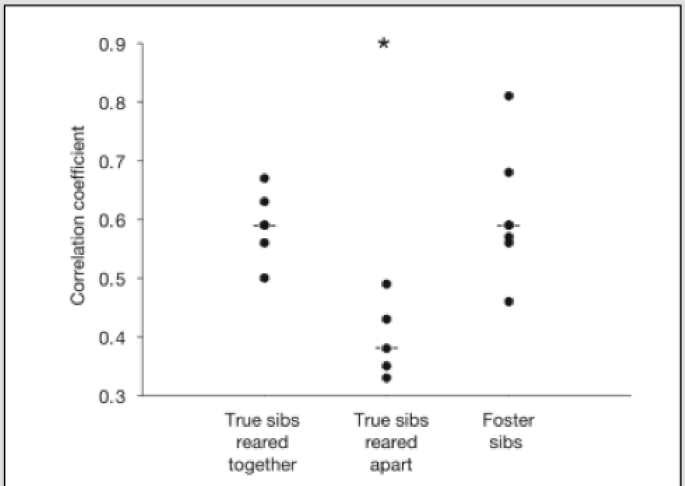


Figure 2 Call similarity between different groups of siblings. Correlation coefficients (dashed horizontal bars indicate means) for pairwise comparisons of churr calls in each group were obtained using SPCC (Kruskal–Wallis test, $\chi^2 = 9.752$, $P = 0.008$). Asterisks indicate significant differences after treatment comparison tests[28].

show trends (perhaps through time). Only show lines that are significant. If the relationship is not significant then do not include a line. If you have not fit a function to the data then simply connect the dots when there is a logical connection between successive points. Do not fit some sort of a smoothed curve that connects all the dots (Excel likes to do this!).

Bar graphs are useful for depicting data that are grouped by categories. Use bars for variables that can take on any value between zero and the mean value. Another way to think of this is that the value starts at zero and increases until the mean value (e.g. growth measures, days since events, number of offspring produced, amount of methane released). For point estimates it is better to use isolated points +/- SE rather than a bar +/- SE.

Recently there has been some backlash against bar charts since they can hide weird patterns in the data, and can, therefore, be misleading. Violin plots or box and whisker plots where the data are shown (but jittered) are two alternatives that are preferred because the distribution of the underlying data is more transparent. I really do need to write a new section on this because this is an important issue!

Partial Plots

In a general linear model we are often interested in assessing the effect of some variable, X1, after controlling for one or more other variables (e.g. X2, X3, etc.). This is what the parameter in the statistical model represents - the effect of that variable after controlling for the effects of all other variables. The problem is that if we are using only our raw data then we can only plot the effect of X on Y. One way to deal with this is to use partial plots. Partial plots

show the relationship between some predictor of interest and 'corrected' values of Y. These are partial residuals. They are 'partialed' or 'corrected' for all the other terms in the model.

For example, if we have some model

$$y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + e_i.$$

and we are interested in showing the effect of X_2 on y after controlling for the effects of X_1, X_3 and X_4 then we need to plot the partial residuals for X_2 against X_2. In this case we can calculate the partial residuals as

$$e_{partial,\ i} = \beta_2 X_{2i} + e_i.$$

We will discuss this in more detail when we review general linear models.

Appendices and Supporting Material

In recent years publishers have greatly increased the amount of material that can be included in online-only supplemental material or appendices. These are cheap for the publisher since they do not appear in the printed article, so they encourage them. Beware though that these are often not read as closely as they should be. Remember that 1) publishers' and authors' best interests are often in conflict and 2) readers are lazy. Online appendices can be quite useful for providing additional data that might not be typically included in a basic article but which might be useful for some future meta-analysis. This is also changing as many journal are now requiring that you upload your original data to some data archiving site, such as Dryad.