

Overview

The purpose of a Results section is to illustrate to the reader the evidence for your specific conclusions through statistics quantified on the data. In genomics applications and methods papers, it is often, but not always, the case that the Results section comes before the Methods section. The Results section should be (mostly) self-contained: do not assume the reader has read your Methods section, and start by motivating and describing your approach. Present the data you are working on in a clear and quantitative way. Describe the results of applying methods to those data. Finally, describe what the results individually and in aggregate suggest about your conclusions. Each paragraph may have the following general form: "In order to address the question of XXX, we applied our statistical model to the YYY data. We found that ZZZ. These results suggest that AAA." Note that the Results section is written in past tense, describing experiments that were done in the past.

Structuring your Results section

Results on Simulated Data; Methods Comparison

Results on simulated data should contain the following information:

- how were the simulated data simulated? (if using an existing simulator, describe and cite)
- what other methods were run? How were they run? Was there any special post-processing or preprocessing of data or results necessary to make these results comparable in your analysis pipeline?
- what statistics are you using to compare the results from these methods on your data? why are those appropriate? what do each of them tell us about the methods (e.g., FDR vs power vs mean squared error, etc.)? why might we need more than one?
- how do each of the results from the methods compare? what assumptions implicit in the different methods are borne out in the results? What structure or predictions are hard or easy for each method? What are hard?

Keep in mind: we are not comparing different methods, we are comparing results from different methods. Here, it is essential to be fair and honest. It is not useful to say that your method is better than all others in all tasks, but it is essential to show that your approach has some favorable qualities with respect to the other approaches. The authors of the related methods will be reviewers for this paper.

Results on Real Data

As with the simulated data (but often with fewer comparative methods) your Results for a real data application should include the following information:

- What do the real data look like (features, samples, missing data, outliers)?
– cite original work if relevant
- What questions were we asking when we applied our method to the data?
- What were the results of applying this method to the data? This is often overlooked and instead the answers extracted directly. It is often useful to characterize what the results looked like: number of selected features, plots of linear regression, imputed time-series data
- What specific aspects of the results can be used to answer our domain-specific problems of interest?
- If an alternative method was run, why does the method in the paper produce results that are superior in terms of addressing specific questions about the data?
- How can we quantify the results without ground truth? What experiments (e.g., GO analyses, eQTL mapping, enrichment analyses) that can serve to validate our findings in the absence of experiments?

Presenting results in the Results section

Making an observation about your data or results.

A statement of the general form:

“We found that there were substantially more quarks in a bloogle than in non-bloogles.”

must be accompanied by a p-value or some measure of how likely it is that this observation could be made by chance. For example:

“We found that there were substantially more quarks in a bloogle than in non-bloogles ($p \leq 5.6 \times 10^{-4}$; Fisher’s exact test).”

Some notes about this significance presentation:

- always accompany the p-value statement with a short reference to the test. If it is your own test (e.g., from a logistic regression equation), cite that equation number, which should be written clearly in your Methods section.

- always write the quantified significance the same way throughout your results. Here, I specified p-value as p , and used the less than or equal sign.
- do not cut and paste this p-value from R. R uses the awful format $4.5E^{-4}$. This is meaningless in writing. Use base 10 notation.

In general, words like *more*, *most*, *fewer*, *some* should be avoided; instead opt for a quantitative treatment of the observation or comparison.

Referring to a figure or a table

Refer to results from a table or a figure, not to figures or tables themselves.

- Bad: “As you can see in Figure 4, the first and second PCs clustered the ancestral groups.”
- Instead: “The first and second PCs clustered the ancestral groups (Figure 4).”
- Bad: “Table 2 shows the accuracy of our method against all of the alternative methods.”
- Instead: “We compared the accuracy of our method against four alternative methods on the simulated data (Table 2).”