Reporting Statistical Results

Presentation Guidelines and Good Practices

Patrick D. Kilgo and Lisa K. Elon

Department of Biostatistics and Bioinformatics

Emory Rollins School of Public Health

Revised April 2022

This document was conceived with the idea of helping public health students learn the basics of reporting statistical results. It focuses on three domains:

1) writing clearly and succinctly about statistical results,
2) organizing statistical reports, and
3) learning the essential elements of preparing clear, easily understood tables and graphs.

Everything covered presumes that you have analyzed the data correctly and responsibly. The elements we discuss are based on the format of a standard journal article in the public health field. However, these guidelines are neither exhaustive nor compulsory - refrain from using any principles that do not apply to your analysis and be sure to tailor your report to the specifics of your study. These are guidelines, not cookie cutters; please use them judiciously.

The basic framework for reporting your results will include these elements:

The Title

The title of your paper should be catchy but practical. Some people choose to pose the title as a question to whet the reader's curiosity:

> Example: **Does the number of grafts influence surgeon choice of off-pump surgery over conventional on-pump coronary artery revascularization in multi-vessel coronary artery disease?**

Others choose to frame the title in terms of the results of the study:

> Example: **Off-pump techniques benefit men and women and narrow the disparity in mortality after coronary bypass grafting.**

Some articles simply list the study subject:

> Example: **A population-based comparison of CIREN and NASS cases using similarity scoring.**

A personal favorite is a hybrid approach where you mix a few of these elements together using a dash or a semicolon:

> Example: **The worst injury predicts mortality outcome the best: Rethinking the role of multiple injuries in trauma outcome scoring.**

Or

> Example: **Did Steroid Use Enhance the Performance of The Mitchell 89? The Effect of Performance Enhancing Drugs On Offense From 1995-2007.**

Of course, be sure the title clearly identifies the topic. This is a title from *Health Education Research (online May 28, 2012)*:
The FLU-FOBT Program in community clinics: durable benefits of a randomized controlled trial.

You might think this is a paper primarily about influenza outcomes; sadly, you would be mistaken. The FLU-FOBT program tried to encourage fecal occult blood testing (FOBT) at the time of annual flu vaccination. A more informative title might have been "**Colorectal cancer screening can be increased by linking it to annual flu vaccination**".

Section 1: Introduction

For a journal article or thesis, the Introduction or Background section serves the following purposes:

- Promotes the general importance of the topic.
- Reviews previous literature related to the subject and points out shortcomings, limitations and needed extensions of those studies.
- Identifies a gap in knowledge of the subject matter.
- Motivates the present research question by justifying the need for the study (what still is not known, needs to be clarified, or needs more evidence).
- Plainly states the intent of the present research study. It should include the objective (what you hope to add to the knowledge base on this topic). This is generally stated in the last paragraph of the introduction.

Section 2: Methods

The Methods section is used to tell the reader *how* a study was conducted. Think of the Methods section as revolving around four sequential items: 1) Subjects 2) Sample and Study design 3) Measurements 4) Analysis. There should be enough information for an a reader to judge the quality of your study design and to replicate the study, if desired.

Subjects:

- Inclusions and Exclusions – what do the subjects have in common? What factors make them eligible for the study? What factors would exclude them from the study?
- Setting – from what setting are the subjects drawn? Are they gathered from the same geographical location? Were they all treated by the same hospital? Give the location of subjects (geographically).

Sample and study design:

- How was the sample identified? Was it a random sample? A convenience sample? Other?
- Specify the timeframe for data collection.
- How many subjects were included in the sample? How was this number determined?
- Study design – describe in detail the design of the study, which might include a brief protocol description for how the study was conducted and a rationale as to why the selected design provides the best measurements.

Measurements: Description of the study variables.

- Outcomes (sometimes called “endpoints” or “the dependent variable”).
- Primary study variables – the independent variables of most interest; this is the “treatment” variable in an experiment.
- Covariates – other variables that need to be taken into account, perhaps confounders that need to be adjusted for in a model. Sometimes called “adjustors” or “control variables”. Often, they might be predictors of secondary interest (still of interest, but not the primary concern).
- Detailed description of the study intervention, if any.
- Describe the scale (numerical or categorical) and sub-scale (ordinal, nominal, dichotomous, continuous, discrete, etc) that was used for each variable <u>if it is not obvious</u>.

“Chronic Lung Disease was ordinally classified as None, Mild, Moderate or Severe.”

> "The ABR score has a possible range of 0 to 33, with higher values indicating greater agility."
>
> "Control variables included age (in years), diastolic blood pressure (mm Hg), and binge drinking (drinking ≥5 alcoholic drinks on at least one occasion in the past 30 days, yes/no)."

- Define any variable that is not clear to a lay reader (gender is clear, "adequate nutrition" needs explanation).

- Describe any stratification made on a variable and the subgroups that would result from this.

> "Disease severity was categorized based on the baseline factor level: mild (>20 mg/cm^3), moderate (10-20 mg/cm^3), or severe (<10 mg/cm^3)."

- Be explicit about what is measured, the units, the timing (if relevant), etc.

> "A blood sample was drawn immediately before the vaccination and again 14 days later. A total T cell count (in mm^3) and percent central memory cells were measured via FloJo equipment."

- When appropriate, describe standard definitions of "Normal."

> "Troponin levels were measured in each patient during the peri-operative period; troponin levels greater than 0.05 were considered elevated."

- Briefly mention clinical, lab testing, or survey procedures used (more advanced methods would need more explanation). This might include a description of a survey or protocol, where appropriate.

- Describe any new variables created from the study variables:

> Example: "Race was dichotomized as either Asian or non-Asian. "

- Describe the variables that delineate the study groups (for instance, an experiment with a treatment and a control might employ a single binary variable called 'treatment' which is categorized as either Yes or No). Describe very clearly who/what is in each group.

> "Subjects who exercised vigorously at least 3 times a week for at least 20 minutes each, or exercised moderately at least 5 times a week for at least 30 minutes each time, were categorized as 'following exercise guidelines'. Those reporting fewer or shorter exercise sessions were categorized as 'not following exercise guidelines'. "

- Be certain to describe interventions or protocol-specific schedules in chronological order.

- **Do not** describe the SAS coding steps or use any SAS variable names. Rather, briefly write something like, "Age was categorized as 'under 65' and '65 and older'. Due to small sample sizes, A, B, and C levels of enzyme X were combined to represent a 'low to medium' level."

- **Do not** refer to dataset variable names in the text or the graphs (for example, refer to ‘weight group’, not *wtgrp; ‘*physical activity score’, not *PAItot*).

- Data Quality / Data Cleaning / Data Storage – Often this is not included in a journal article, but we expect to see a brief section on this in your reports. You should describe any checks that were in place to identify problem data and any steps that were taken to “clean” these data. Sometimes cleaning data simply involves correcting typos or checking records. You should mention implausible values and how they were practically (as opposed to statistically, which comes later) dealt with. You can briefly mention the type of database used to store the data. Missing Data – describe which variables have missing data, how many subjects are missing values and what proportion of the sample these represent. Note this in data tables as well.

Analysis

- **Describe the simplest statistical procedures used first and progress to the most complex**. In other words, start with simple descriptive statistics, then relationships between two variables, and finish with those analyses involving three or more variables.

- State the statistical tests used – 2-sample t-test; chi-square test of association; linear regression, etc. You do not necessarily need to *justify* the use of a common method (most commonly, *justify* means stating the assumptions of the test and how you checked that these assumptions were met). Usually, you do not need to put that in a journal article. This is of interest to your teachers but not to the general reader.

  > "To statistically evaluate the effects of GFR class on in-hospital mortality, a multivariable logistic regression model was constructed. The primary variable under consideration was GFR class (Normal, Mild, Moderate, Severe, or Dialysis). The model also contained pre-operative covariates to adjust for potential selection bias; Age, Ejection Fraction, Caucasian Race, Gender, Diabetes, Chronic Lung Disease, PVD, Angina, Arrhythmia, Heart Failure, Previous MI, Left Main Disease and Concomitant CABG. Adjusted odds ratios (AOR) associated with pre-operative dialysis and other covariates, along with 95% confidence intervals (CI), were computed. The reference group for the adjusted odds comparisons was the Normal GFR group."

- To reiterate, all statistical methods make assumptions. In class, you have learned how to check these assumptions (examples that might apply: study design to confirm independent data, examining plots for linearity, residual plots, thinking about normality, no multicollinearity). Ignoring this step in the analysis plan can cause your results and conclusions to be misleading. Your code should reflect assumption checking.

- Briefly report verification of assumptions for higher-level analyses.

  > "The proportional hazards assumption was verified via a correlation analysis of the Schoenfeld residuals and ranked follow-up time."

- Always mention the alpha level at which all tests were evaluated. When appropriate, note if adjustments for multiple tests were made.

- Mention the software package used;

  > "The data were analyzed using SAS 9.3 (Cary, NC)."

- Statistical Power. Power calculations should be reported for clinical trials and other experiments. This will not be required for your reports unless specifically requested by your instructor.

- Outliers – tell the reader how outliers or implausible data were treated in the study.

Section 4: Results

The main purpose of the Results section is to present the evidence you have found relating to the research question. **It is essential to start populating tables first (see page 14)**. Once the tables are filled in, the writing can flow in a logical way through the table results. Similarly, graphics should be considered early on.

The Text:

Order

- Report results in a logical sequence. Commonly, the first paragraph of the Results section is dedicated to a brief overview of the subject demographics, including the sample size, and a summary of the treatment groups, if applicable. This is usually followed by primary analysis results. In other words, start with the simple descriptive statistics. Second, present the bivariate results. End with the multivariate results (if any).

- Refer to the table (e.g. [Table 1]) at the **first** mention of its contents. This helps the reader tremendously by immediately linking the text to a specific table that they can refer to as they read. Address items in the table in order. If summary statistics on age are in the first row and BMI in the second, discuss age first, then BMI. If you decide you prefer to discuss BMI first, change the order in the table to match. This eases the reader through your results.

- If your document has more than one table, discuss them in the same order that they are labeled (Table 1, then Table 2, then Table 3, etc.). If you change the order in the text, be sure to re-number and re-order the tables—text and tables should be in harmony.

Content

VERY IMPORTANT –INTERPRET FINDINGS FOR THE READER.

At least briefly, write about each table and graphic you include in your report—if it is important enough to include in a table or graphic, you need to **interpret** it for the reader. Both significant and nonsignificant results should be reported.

Saying a result is significant is never enough. For example, consider this statement:
"Pain relief in the two groups was significantly different (p=0.01)"
This statement leaves the reader completely clueless and wondering…how large a difference? Which group had more pain relief? Is the difference in the expected direction? How precise are the estimates? Are my results similar to those in this study?

The following section explains how to address your reader's curiousity.

- Supply the reader with:
    - the **direction** of the effect, if any

      "Group X exercised 10 minutes *more* than Group Y"

      "Among the 184 patients who were measured at 1 year, patients with at least one non-patent graft (n=7) had *significantly worse* survival than those with all patent grafts (n=177, 42.9% vs. 79.7%, p=0.007). Patients with a closed graft at one year have an instantaneous hazard of death *3.76 times higher* (95% CI 1.33-10.59, p=0.012) than those with all patent grafts."

      "The difference in mean test duration between Groups Y and Z was small but statistically different (30 vs. 33 seconds; 95%CI on the difference=1.5-4.5) [Table 1]."

      "Pain relief was *significantly more likely* among those in the treatment group than in the placebo group (50% v 25%; p=.04) [Table 3]."

    - If there was no effect, this should be stated, too.

      "There was *no statistically significant relationship* between age and customer satisfaction (p=.8); among both older and younger patrons, two-thirds reported high customer satisfaction [see Table 2]."

      "There was no evidence of a relationship between transplant success and donor age, donor drinking, or recipient age (all p>0.05) [see Table 3]

      "Average weight was similar in the two groups (group A=49 lbs v group B=50 lbs; p=.55)."

    - **estimates and their units** (e.g., "a mean of 10 minutes of exercise per day"). Add interpretation of the results when this might not be clear. Readers will understand a report of means, but will need the interpretation of regression estimates.

      "Activity increased by 10 minutes for every 1 hour decline in TV viewing (95% CI 8-12 minutes)."

    - measures of **precision,** such as standard deviation for descriptive statistics, standard error or confidence intervals for inferential statistics

      "The average activity level per week was 35 minutes (*SD 15*)."

      "The average activity level per week in the teen population is estimated to be 35 minutes (95% CI 20-50 minutes)."

    - statistical **significance** can be added parenthetically [e.g., ''On average, Group X exercised 10 minutes more than Group Y (*p=0.03*)'] or directly ["Exercise was significantly higher in Group X

compared with Group Y (mean difference=10 minutes)" [editorial note: the methods section should have noted the alpha level used to demarcate statistical significance].

- If findings are "statistically significant" but perhaps not practically meaningful, you could point out that the difference is small or not clinically meaningful.

- Avoid redundancy in reporting dichotomous variables – You can usually just report the percent for one of the groups ("among women, 38% exercise"; you do not need to add 'and 62% do not", it is understood). This advice applies to the content of tables as well.

- Not every number in the table has to be discussed; point out the most important or 'surprising'. By giving some numbers and referencing the table, readers will know where to go to find similar data. The variables of primary interest should be discussed with all the details mentioned above, while the secondary variables might be reported more generally. *However, if a variable was important enough to include in a table, it should be at least briefly mentioned.*

- **Nonsignificant results are also important (really!)**. If you tested the relationship of Y with five variables and found four with p>0.05, do not snub these poor slackers—this is useful information. Report that they were not related to the outcome.

  "Popularity was not significantly related to hair length, clothes brand, gpa, nor geekyness (all p>0.05)."

- Numerical values may be reported parenthetically (i.e., within parentheses), especially when they are reported elsewhere in tables.

  "Mortality was significantly higher in those with preoperative dialysis (18.3% vs. 5.2%, p<0.001)."

- Refer to tables and graphs correctly and in order. Again, not every number in the table needs to be explicitly highlighted but readers need to know where to go to find the types of comparisons being discussed in the text. You are the reader's guide through the information—help them navigate swiftly and to correct conclusions!

- Acronyms are OK if they are defined earlier in the paper (usually at the term's first use).

- Do not regurgitate all of the numbers that are in the table into the text. Interested readers can look more closely at the tables, if they wish. *Sample sizes and measures of precision, while not always mentioned in the text, should be in the table.*

- **Write succinctly.**

  Example of wordiness from a student's paper:

  "If there was a significant change in the radiology parameter estimate, the variables were kept in the model. A significant change in the parameter estimate is deemed a change greater than 10%.

> During the removal process, if there was a change in the parameter estimate larger than 10% then the variable was added back into the model."

An acceptable revision of this statement, using about half the verbiage:

> "If eliminating a variable from the model resulted in the radiology parameter estimate changing by greater than 10%, the variable was considered a confounder and was kept in the final model."

Rounding values—i.e. how many decimal places to use?

Use as many as you need to convey the message, but no more than that!

1. With 10,000 people reporting, we found that 49.33% of cat owners enjoy computer games while 55.99% of dog owners do.

2. With 10,000 people reporting, we found that 49% of cat owners enjoy computer games while 56% of dog owners do.

Which version do you prefer to read? Even if the sample size is large and can justify the use of several decimals, do you really need them to get the message across?

Each variable is different. The appropriate rounding should be determined separately for each. You make the decision, not the statistical package. For example, if age is recorded in whole numbers, do not report mean age as 44.567 years just because that is what the program spits out.

For continuous variables: A rule of thumb is that if the original measurement was recorded to one decimal place (ex. 3.4), then you may report summary statistics such as means to two decimal places (mean=3.45).

For percentages: consider the size of the denominator. Generally reporting whole percentages (ex. 27%, not 27.3%) is adequate. Do not use decimal places in percentages unless the denominator in the upper hundreds.

For P-values: typically rounded to the third digit; therefore, the smallest p-value you would report would be denoted as $p < 0.001$. This is not a strict rule.

Annotated examples:

Patients with a closed graft at one year had an instantaneous hazard of death 3.76 times higher than those with all patent grafts (95% CI on the hazard ratio, 1.33-10.59; p=0.012) [Table 3]. {*comment: gives the point estimate and its precision, gives the direction of effect (higher), gives additional information on the probability of seeing such a HR if there is truly no difference in the population (i.e, the p value), and references the table number*}

The mean test duration was significantly shorter in Group Z than Group Y, but only by 3 seconds (30 vs. 33 seconds; 95%CI on the difference=1.5-4.5) [Table 1]. {*comment: gives direction of difference, says it is statistically significant and gives precision of estimate, but also points out that the actual difference is small (this does border on "discussion/opinion" but is probably justified here; this point can be elaborated in the discussion section. It references the table.*}

There was no statistically significant relationship between age and customer satisfaction (p=.8); among both older and younger patrons, two-thirds reported high customer satisfaction [Table 2]. {*comment: author notes insignificant result, gives p value to substantiate the statement; goes on to report the overall high satisfaction rate, which may be of interest to the reader even if the rate doesn't differ by age group—use your judgment about what information to highlight from the tables.}*

Pain relief was significantly more likely among those in the treatment group than in the placebo group [50% v 25%; p=.040] [Table 3]. The treatment group also had higher average satisfaction scores compared to the placebo group [5.2 (sd 1.1) v 3.3 (sd 0.7); p=.003]. There were no significant differences between the groups in average length of hospital stay, number of office visits, or depression index score (all p<.4). {*comment: Gives the table # <u>immediately</u> so the reader can find the data. There were five variables in the table. The two significant items are reported in detail first, while the three insignificant items are then just listed—the interested reader can go to the table to get more detail. This is one way to succinctly deal with non-significant items while not ignoring them. If a variable is important enough to make it into a table, it should be mentioned at least briefly in the text. The level of detail is a judgment call. Sample sizes and measures of precision, while not always mentioned in the text, should be in the table.}*

Tables:

Each table or graph should be able to stand on its own. If readers were to only have the table or graph without any accompanying text, they should still be able to understand the data. The formatting and all other details of the presentation should permit easy navigation through the data and promote quick understanding of the results.

For Each Table:

- Write meaningful titles. ["The relationship between weight gain and dietary habits among teenage girls in the Southern US", not "weight gain and dietary habits"; "Demographic and clinical characteristics of subjects in the 2012 IPIRC study.", not "Descriptive statistics"]

- Be selective in the data you present. Do not use everything SAS gives by default, just what is relevant to describe your data. If your data are normally distributed it may be enough to report a mean and standard deviation. If the distribution is skewed you may choose to report a 5-number summary or a median and interquartile range (IQR).

- Do not copy and paste SAS output. Design your tables thoughtfully. All aspects of the format and content should be intentional to guide the reader through the material. The data can tell a story—but a well-told tale is carefully planned. A demographics/characteristics table should be the first. If there are two or more groups of primary interest, this table should contain information describing distributions in these primary predictor groups.

- Put some thought into the order of the variables—what order will best tell the 'story' of the data? Some possibilities: Importance, Size of the outcome, Time-based order, Geographical order, etc. Alphabetical is usually the least useful order.

- Consider subheadings/groupings in longer tables. For example, subsections for variables representing demographics, health history, clinical data, activities, etc.

- Refer to variables with explicit terms when they would not be obvious to the lay reader or are ambiguous. ["physical activity > 3 days/wk", not "adequate physical activity"; "Herpes simplex virus", not "HSV"]

- Include n. The reader wants to know how many observations are being summarized.

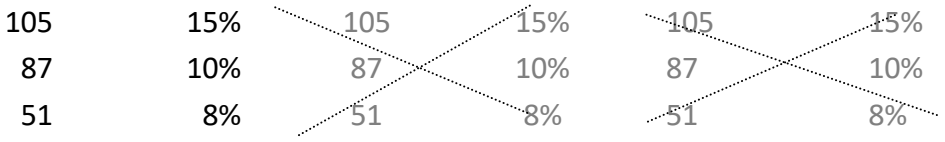
- Include units where appropriate. ["age, yrs", "physical activity, days/wk"]

- Limit decimal places—usually none, one, or two is appropriate, depending on how much data you have, how accurate the measurement is, etc. Review page 12.

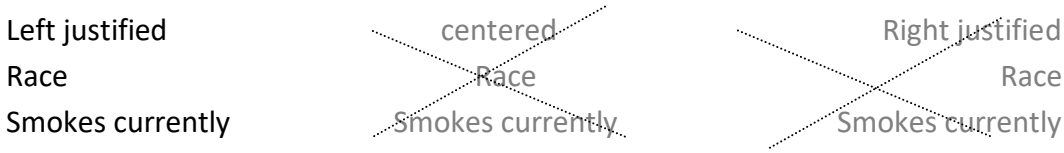
- Clearly label the statistics —mean, std, %, 95%CI, etc. These labels might be best at the top of the column, or they might be appropriate next to the row's label. Use your best judgement.

- Be consistent – use the same page justifiers, fonts, font sizes, etc. for all cells. If variable X is reported to one decimal place in table 1, it should continue to be reported to one decimal place in tables 2 and 3, and in the text. However, variables X and Y do not have to have the same rounding—rounding depends on the raw measurement's precision. Review page 12.

- Right-justify all numbers. Glance down a column; notice that it is much easier to comprehend and compare the values when the digits are lined up.



Left justify text



- Use footnotes for
  a) information on the number of observations that are missing values,
  b) long variable definitions, and
  c) to provide details on the statistical tests. Footnotes are important but keep them to a minimum—consider alternatives if they start multiplying.

- Tables should be kept small with respect to the number of columns (i.e., prefer a long table format to a wide one). Large, unwieldy tables with lots of columns are irritating and difficult to read.

- Avoid redundancy in reporting dichotomous variables – You can usually just have one row to report the percentages for one of the groups ("among women, 38% exercise"; you do not need to add 'and 62% do not", it is easily understood). For studies with treatment groups: Each "treatment group" should have its own column and the risk factors / covariates /outcomes should be listed in rows.

- Each treatment group column should have a **bolded** header with the sample size listed.

When There Are Multiple Tables:

- Order tables from simple to complex (i.e., univariate -> bivariate ->multivariate)
- Write about the results in the same order (simple, progressing to the most complex).
- Explicitly refer to each table … [Table1], [Table2], [Table 3], etc.
- Place all tables and figures in an appendix for journal article submissions. Intersperse them with the text in grants and simple reports.

The following section has examples of many types of tables. We provide comments to help you recognize what makes a good table.

Annotated Examples of tables

Table 1. Patient clinical condition and hospital admittance status for Southern Hospital's Emergency Department, July1 to Aug31 2001.

Characteristic	Overall n=677	
Patient symptoms	**n**	**%**
Chief complaint		
Abdominal pain	260	36.6%
Genitourinary problems	105	15.5%
Nausea/vomiting	87	13.0%
Back pain	51	7.6%
Fever	40	5.9%
GI bleed	37	5.5%
Diarrhea	27	4.0%
Gynecological problems	27	4.0%
Other	16	2.4%
Post-operative problems	14	2.1%
Constipation	8	1.2%
Any secondary symptoms	448	66.5%
Abdominal or pelvic pain	493	73.3%
Nausea	341	50.9%
Vomiting	248	37.0%
Diarrhea	132	19.6%
Fever	102	15.2%
Vaginal bleeding	31	4.6%
Patient history		
Previous abdominal pathology	150	22.3%
Previous cancer history	129	19.1%
Immunocompromised	81	12.0%
Physical examination and lab tests		
Positive signs in abdominal exam	434	64.4%
Focal tenderness	228	34.7%
Distension	40	6.1%
Abnormal bowel sounds	36	5.5%
Mass	22	3.3%
Rebound Tenderness	19	3.0%
Rigidity	19	2.9%
Any positive lab test--blood	367	62.7%
Any positive lab test--urine	262	41.6%
Received any ancillary imaging	97	14.5%

Comments on this table:

1. Informative title.
2. Uses multiple visual guideposts—structural elements such as bolding, white space, line weight, subheadings to define different elements of the table
3. Includes overall sample size
4. Thoughtfully organized on several levels: first, temporally and topically (from admittance, physical exam, and then types of medical tests); second, by frequency within topics.
5. Numbers within a row can be easily distinguished as n and %.
6. Columns of numbers can be scanned easily because of spacing and alignment (right-justified).
7. Uses appropriate number of decimal places.
8. Decimal places are justified based on the overall n of 677.

Table 1. Characteristics of Participants by Gym Membership, The GPA Study 2022.

	Gym Membership		
Characteristic	**Yes n=100**	**No n=100**	**p-value***
Gender			0.043
Female	67%	53%	
Ethnicity			0.010
White	37%	21%	
Black	51%	72%	
Other	12%	7%	
DBP mmHg mean (sd)**	115.2 (10.2)	118.3 (10.1)	0.813

*** Gender, Ethnicity: Chi square test; DBP: 2 sample t-test;**
****Diastolic blood pressure**

Comment on this table:

1. Explanatory title.
2. Estimates are based on the column as the denominator. This is correct *in this case* because the author wants to compare demographics of non-members to demographics of members.
3. Provides units for DBP. Defines DBP in a footnote.
4. Defines the values as either mean and sd, or %.
5. Gives n for each column (if there were missing values for any of the characteristics, a column for the row n's would be appropriate.
6. Lists the names of the statistical tests in a footnote.
7. Digits are lined up for easy scanning down the column.
8. Decimal places are consistent within a row.
9. One decimal place for average DBP is appropriate since it is measured in whole units.
10. Whole numbers for % are appropriate for this sample size (n=100 in each group)

Example of another way to cleanly present a mix of continuous and categorical data by using separate columns:

Characteristic	mean ± sd	n (%)
Demographics		
Age	38.9 ± 8.2	
Gender		
Male		76 (40%)
Nationality		
United States		124 (65%)
Guatemala		34 (16%)
Peru		26 (13%)
Mexico		3 (2%)
Other		8 (4%)
Years of education	6.5 ± 3.5	
Health-related factors		
BMI		
Male	27.9 ± 4.2	
Female	29.2 ± 4.5	
BMI ≥ 30		
Male		20 (26%)
Female		47 (41%)
History of hypertension		12 (6%)
Blood pressure		
Normal (<120/80)		63 (33%)
Pre-hypertension (120-139/80-89)		84 (44%)
Hypertension (<140/90)		45 (23%)

Yet another way:

Characteristic	mean± sd median [q1, q3] or % (n)
Demographics	
Age, yrs	38.9 (8.2)
Male	40% (76)
Education, yrs	6.5 (3.5)
BMI	26 [22, 31]

Demographics	
Age,yrs, mean (SD)	38.9 (8.2)
Male, % (n)	40%, (76)
Education, yrs, mean (SD)	6.5 (3.5)
BMI, median [q1, q3]	26[22, 31]

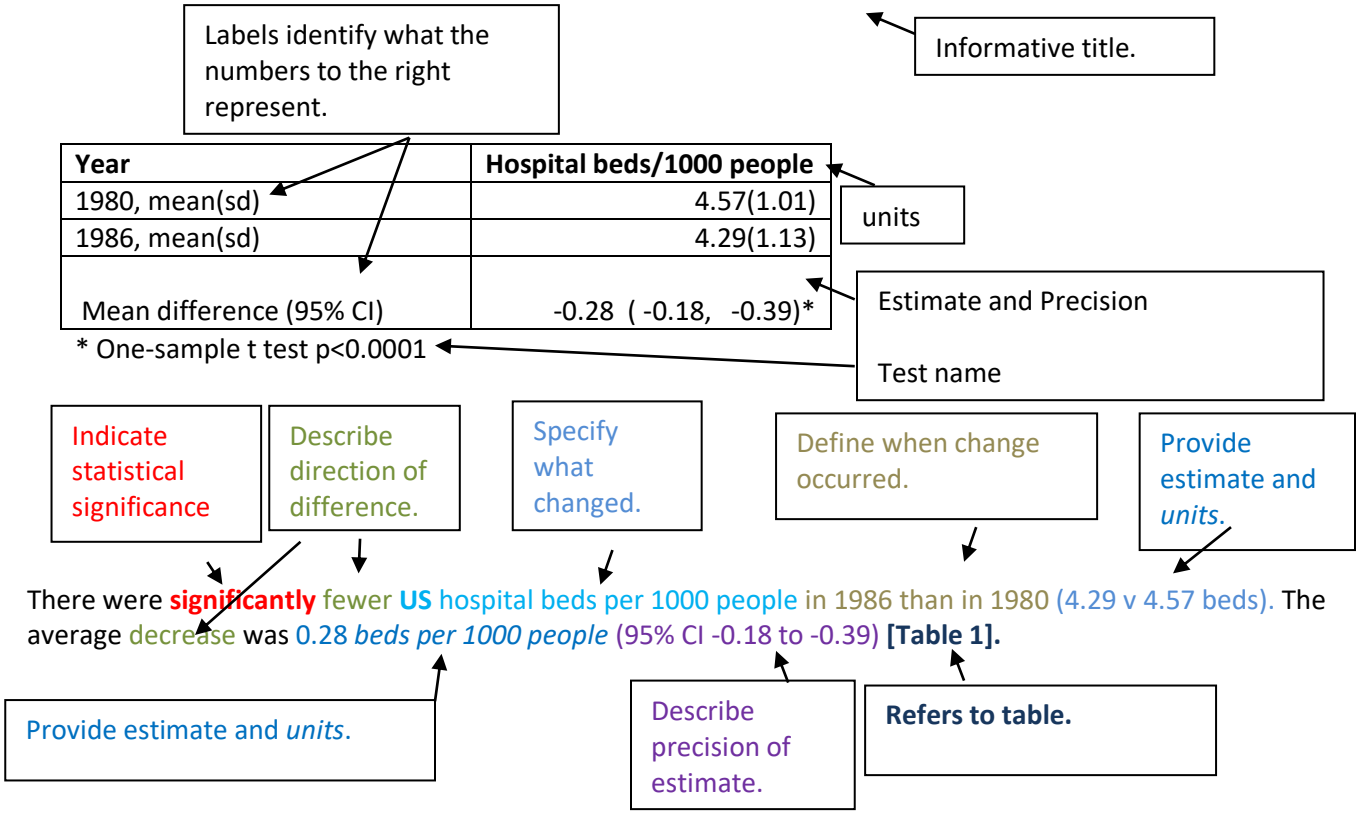
Put mean and SD in separate columns.
Right justify for the mean and left justify for the SD. Similarly for % and n.
This allows the reader to glance quickly down the column of neatly lined up estimates.

It is best to remove the column and row borders when using this presentation.

Compare this to the greyed out version, which is quite 'busy'.

Here is an annotated example of a simple table with its accompanying text.

Table 1. Change in the number of US hospital beds per 1000 people between 1980 and 1986 (n=50 states).



Notice that the table also provides the number of hospital beds/1000 in each of the two years (4.57 and 4.29). Although the focus of the study was on whether there was a change between 1980 and 1986, most readers would also be interested in the actual values for each year and not just the change and confidence interval on the change. You might mention the group values in the text, as done above, or you might just leave it to the interested reader to find them in the table.

An example of a table for simple linear regression results

Table 2. The association of factor Y with several patient characteristics, assessed with simple linear regression; the Zip Study, 2021, n=421.

Characteristic	Intercept estimate (se)	Slope estimate (se)	p-value for test of slope	R^2
Urban Area (ref=NYC)	1.00 (0.05)			0.18
Atlanta		-0.19 (0.04)	0.023	
Chicago		-0.21 (0.05)	<0.001	
Drinks any Alcohol (ref=no)	1.00 (0.05)	0.35 (0.04)	<0.001	0.35

Notice that the reference level is specified by (ref=). Confidence intervals could have been supplied instead of standard errors (use the clb option in the model statement to output 95% CI: model y=race1 race2/clb;)

Below we provide three examples of text describing the results for the variable *urban area.*

1. "Both Atlantans and Chicagoans had significantly lower average values for factor Y compared to New Yorkers (est. β (se): -0.19 (0.04) and -0.21 (0.05), respectively) [Table 1]. Urban differences explained 18% of the variability in Y."

2. "The average factor Y of Atlanta residents (0.81 units) was lower than that of New Yorkers by 0.19 units (standard error for the difference = 0.04 units, p-value=0.023). Similarly, the average Y of Chicagoans (0.79 units) was lower than that of New Yorkers by 0.22 units (standard error for the difference = 0.05 units, p-value<0.001). Urban area differences explained less of the variability in Y than did alcohol drinking (18% v 35%)."

3. "Urban area differences were significant (p-value < 0.001) and explained 18% of the variation in Y R^2=0.18). New Yorkers had significantly higher average factor Y than both Atlantans (95% CI for the difference xx to xx) and Chicagoans (95% CI for the difference xx to xx)." [Editor's Comment: if reporting confidence limits, here denoted by xx, report them in the table]

Notice that each example provides quantities and interpretation of the results--- significance, direction of the effect for each comparison (i.e., higher/lower), and precision---without repeating every number from the table.

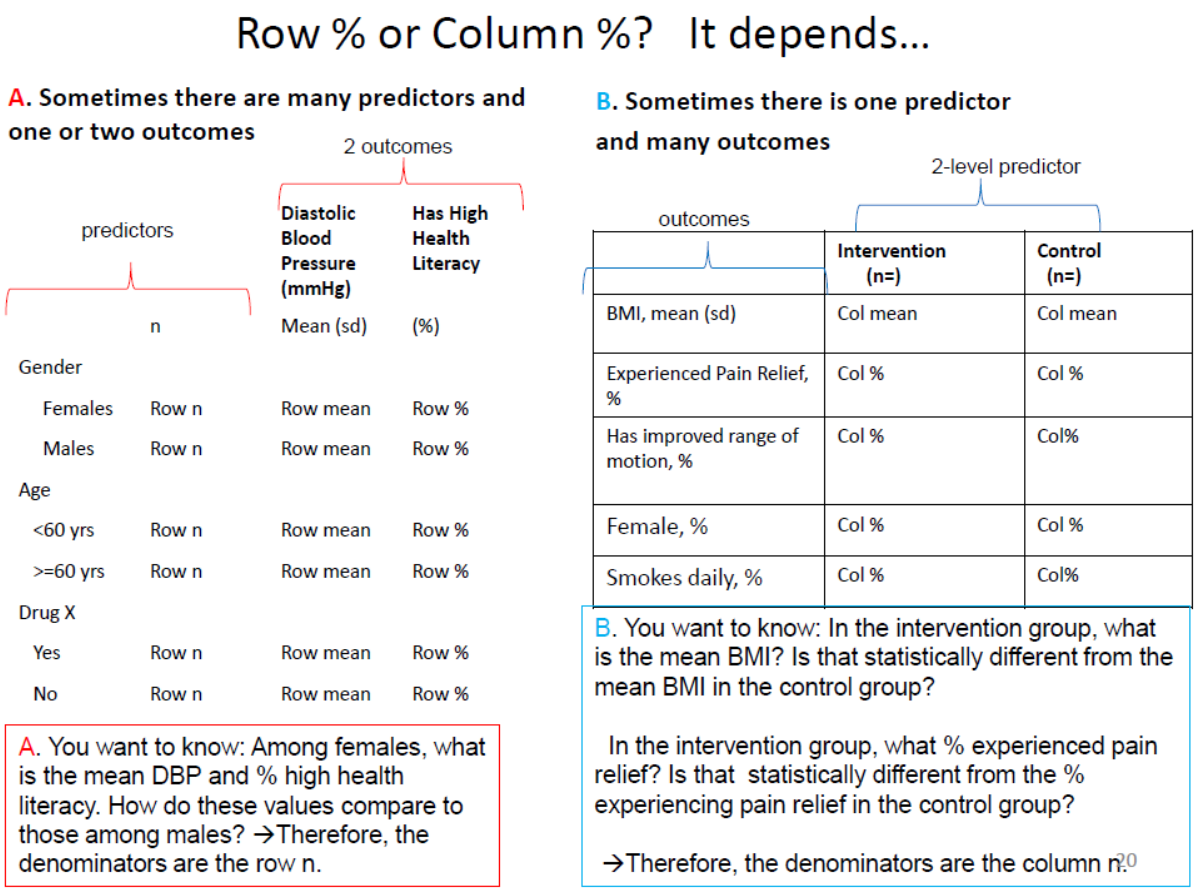
Example of a logistic regression results table with two models.

Table 1. The Association of Demographic and Health-Related Characteristics with Workplace Heat-Related Illness. The Solaris Study, N=325

Characteristics	Model 1[a] Odds Ratio (95%CI)	Model 2[b] Odds Ratio (95%CI)
Age, per 10 years	1.00 (0.96, 1.05)	1.00 (0.96, 1.03)
Gender (Female, ref = Male)	0.65 (0.31, 1.34)	0.76 (0.43, 1.35)
Nationality (non-US, ref = US)	0.93 (0.38, 2.30)	0.92 (0.49, 1.73)
Education, per 5 years	0.96 (0.90, 1.02)	0.93 (0.89, 1.00)
Body Mass Index	1.07 (0.99, 1.17)	---
Blood Pressure (ref = <120/80)		
Prehypertension (120-139/80-89)	0.95 (0.43, 2.08)	---
Hypertension (<140/90)	0.93 (0.35, 2.48)	---
History of Diabetes (Yes, ref = No)	0.80 (0.19, 3.33)	---

[a]Full logistic regression model with demographic and health-related covariates
[b]Reduced logistic regression model with demographic covariates

- The title gives enough information that it can stand on its own without necessarily reading the whole report.
- Characteristics have a logical order (demographics, then health), are clearly defined, and the referent group is specified.
- The sample size is given.
- The estimates and 95% confidence intervals are presented.
- Odds ratios and confidence intervals have a consistent number of decimal places.
- The number of decimal places is not so many that the reader is in shock.
- The type of modelling method is mentioned.



A long table is generally easier to read than a wide one. Choose the format you feel best conveys the information.

How to report dichotomous variables. Notice that in example **B** there is one line each for the dichotomous variables gender (M/F) and smokes daily (yes/no). If reporting column percentages, choose one level of the dichotomous variable to report. For example, in the intervention column, the numbers should answer *either* 'Among the intervention group, what % smokes daily?' *or* 'Among the intervention group, what % does not smoke daily?', *but not both*.

Reporting both gives redundant information since the two values add to 100%.

Contrast this to example **A**. There, the questions are 'Among females, what is the average DBP?' and 'Among males, what is the average DBP?'. Clearly, in Table A you need an entry for both males and females.

Graphics:

Article readers often go straight to the figures, so they should be self-explanatory. If a figure is taken out of its original context, it should still convey all the information necessary for the reader to understand the data.

- Place the title under the figure
- Label both axes
- Label with clear names. Include the units (if not obvious). Again, refer to the variables using English terms, not SAS variable names.
- As with a table title, a figure title should meaningfully describe the content of the graphic.
- Do not use SAS to apply a title. If you need to change the title, it is much easier to do it in Word than rerunning the SAS code. You may also use the figure caption to explain the symbols and provide some interpretation (see example 2, below).
- For most plots and graphs, the outcome (dependent variable) should be on the y-axis.
- Use clear, bold plotting symbols (but not so bold to be overwhelming)
- The default resolution for many programs is 100 dpi. Plots with this low dpi will be faint and fuzzy, especially if the reader enlarges it. Set to atleast 400 dpi.
- Color-blind friendly? About 12% of males and 0.5% of females are color blind to some degree. See what your plot looks like to a color-blind reader by using one of these simulators:
  https://www.color-blindness.com/coblis-color-blindness-simulator/
  https://pilestone.com/pages/color-blindness-simulator-1

  An example of a good color scheme in R that works for three types of colorblindness, using shades of blue: azure2 skyblue2 royalblue1 blue midnightblue

Examples Graphics

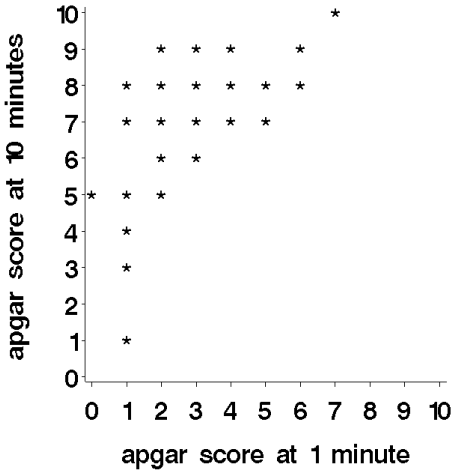


Figure 1. The relationship between 1- and 10-minute apgar scores among premature infants; Atlanta Neonate Study.

Comments:

1. Descriptive title (what group is represented, what measurements, what study)
2. Both axes are labeled with a complete descriptor of the measurement (what and when, in this case).
3. 'at 1 minute' score is on the X axis (predictor), 'at 10 minutes' score is on the Y axis (outcome).
4. The axes have just the right number of tick marks for these measurements (apgar score is measured in whole numbers, from 0 to 10).
5. Each star is distinct and bold, but not overwhelming.

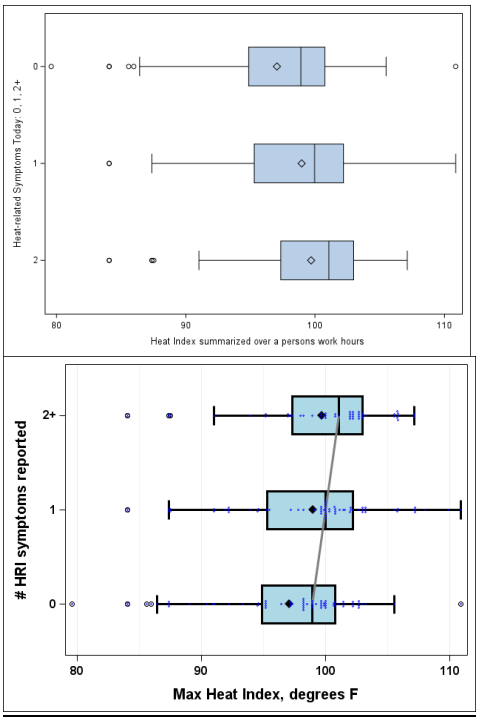


Figure 1. The relationship between max heat index experienced during the workday and the number of heat-related illness symptoms (HRI) reported by agricultural workers in the Girasoles Study.

The top graph was produced with default settings. Several things to note:

1. The resolution is low (dpi=100)
2. The text font is small and hard to read.
3. The axis labels are the assigned the variable labels, which may not be the best for an axis label.
4. In this graph the y-axis ticks represent categories. Categories may not be ordered intuitively. Here the default put the highest category at the bottom. If someone were to casually read the graph, they might get the impression that there are fewer heat-related symptoms when the temperature is higher.

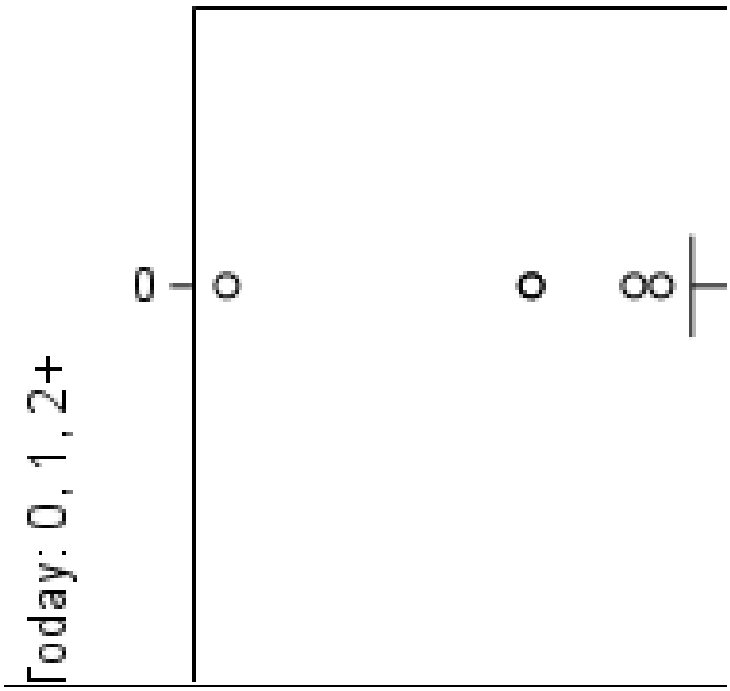
The lower graph was produced using an enhanced SAS template, 400 dpi, and short but clear labelling. (see appendix)

1. The text is bolder and easier to read.
2. The graphing elements no longer look faded.
3. Datapoints were added in smaller symbols in the background as bonus information.

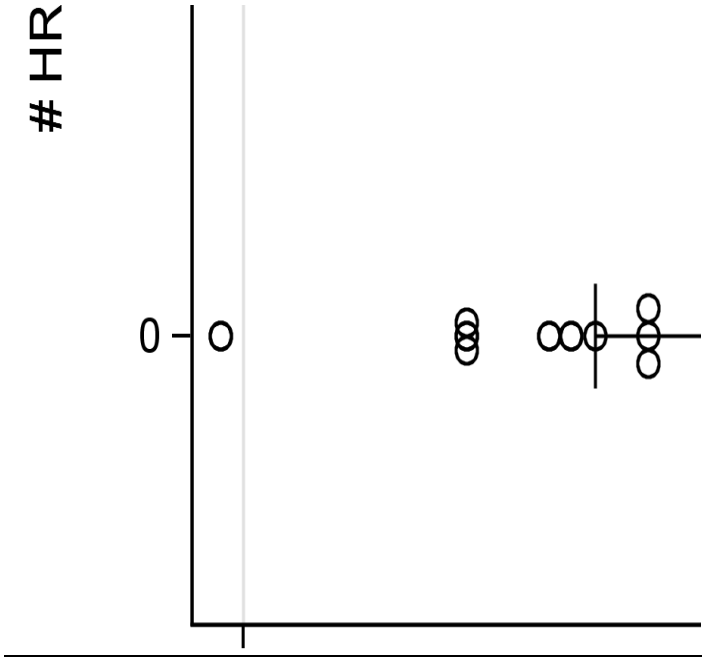
Image Resolution

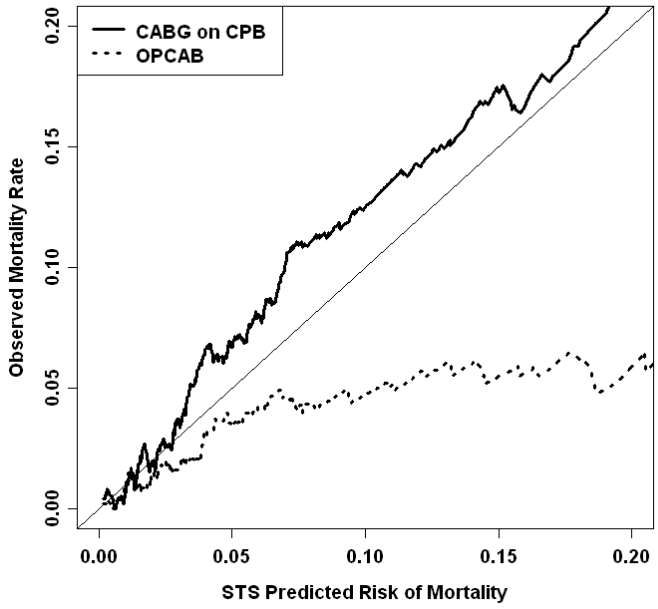
Sharper image resolution is especially important if your graph will be viewed on a computer screen and likely to be enlarged.

100 dpi



400 dpi





Note: the two surgery types use obviously different line patterns. This is helpful even if you chose to use different colors.

Figure 1: The relationship between predicted mortality risk and observed mortality stratified by surgery type. Each jagged line represents a smoothed curve for its respective surgery type.

Section 5: Discussion/Conclusion

The purpose of the Discussion section is to describe the importance/impact of your study. In this section, the authors can deal more in their own opinions than in the other sections.

- **The first paragraph should summarize the most important findings for the reader**. For example, if the purpose of your study was to test a relationship between two variables, state clearly whether there was or was not evidence of such a relationship

  "Anxiety was significantly reduced among those who received the intervention. However, there was no association between anxiety and age."

- **Next, compare/interpret present research to previous research** (consistent with previous research, inconsistent, novel, etc.).

- **Limitations and Strengths**. **Do not underestimate how important this section is**. Readers and reviewers (and people in general) are much more likely to embrace the strengths of your study if you own up to its weaknesses. This should be one of your longest Discussion section paragraphs.

  Limitations might include:

  - Sources of bias: Design limitations, including lack of experimental control over variables in retrospective and cross-sectional studies. Other sources of bias: recall, confounding, interviewer bias, observer bias, sampling method, lack of follow-up, differential dropout, differential missingness, poor wording of questions, etc.

  - Non-validated questions.

  - Categorized data that could have been originally recorded as continuous data. When categorized, detail was lost and the power to detect differences was reduced. Categorization may also be associated with misclassification bias.

  - Shortcomings in statistical Power: It is important in studies designed to test hypotheses that adequate numbers of subjects are recruited to be able to detect an association, if one exists. In other types of studies, it is important to estimate population parameters with a certain level of precision. When recruitment did not meet sample size goals, we need to be concerned that an association was missed because we lacked the

statistical power to detect it –what statisticians call a Type II error. The fact that a Type II error might have been made should be mentioned in the Discussion section as a limitation.

- Inference may not apply to a wider population (i.e., *generalizability* may be an issue)
- Anything else you can think of.
- Please note that lack of a significant effect is not a limitation, it is a perfectly acceptable finding.

Strengths might include:

- a design that permits a wide inference (generalizability)
- large sample size, adequate power to detect difference
- study design and procedures that controlled various biases
- used statistical methods that controlled for confounding factors

- **Conclude with comments on how this research advances the field and why this new information is important.**
  - Implications and recommendations for public policy and future research. You may not need to include this in your report.
  - Recommended extensions of the present study.

Some Thoughts on Effective Scientific Writing

Readers expect a statistical report to have a certain structure and scientific tone. This style of writing may be uncomfortable at first to many of you, especially to those of you from humanities backgrounds. However, effective and efficient communication demands that you structure the paper so as not to distract your reader. Consider the tips below:

When people are reading for comprehension (as opposed to reading for pleasure) they respond better to simple, direct statements than to poetic prose or literary discourse. Use simple, jargon-free language.

All of the written statements should generally map back to the testable hypotheses of the study. You need to "sell" the question to the readers, that is, make them believe it is important and worth asking. After this, you need to "sell" yourself to the readers, that is, make them believe that you have competently addressed the issue and provided a good solution.

Do not try to pack too many ideas into a single sentence. Enough said.

Avoid the use of nebulous words like *sometimes* or *really* or *much* or *very*. These words have no quantifiable meaning and do not belong in scientific papers.

Assume that your reader has an above average grasp of the scientific discipline under discussion. It is not your job to educate a reader comprehensibly about the background of a discipline – that is the purpose of references. Your job is to describe the current study in its scientific context and interpret your results in light of the larger body of research.

Do not "over-describe" methods. Most readers will be familiar with the statistical methods you are using. For instance, you may simply state that you used a chi-square test without describing its assumptions, how it works, when it should be used, etc. Neither should you under-describe when that might prevent accurate understanding of your methods.

Your writing should be sequential, ferrying the reader toward the solution or conclusion. Deliberately guide the reader through the material so that results and conclusions are clear, unambiguous, and memorable.

Check Lists

General

Re-read your paper. Try to see it through the eyes of a person who has is brand new to your data. They do not know the study's purpose, your methods, or your variables. They do not want to suffer to understand it. Organize, explain, clarify.

Tables

- ✓ Titles clearly describe the content, orienting the reader to what will be included in the table (who/what/where/when/how, as appropriate).
- ✓ Use real variable names, not the SAS names
- ✓ Visual guideposts—use structural elements such as bolding, white space, line weight, subheadings to define different elements of the table. Use shading thoughtfully, if at all.
- ✓ Numbers within a row can be easily distinguished, one from one another.
- ✓ Columns of numbers can be scanned easily (use spacing, alignment, etc).
- ✓ Decimal places: appropriate number
- ✓ Decimal places: consistent (i.e., X always to 1, Y always to 3, Z always none)
- ✓ Thoughtfully organized (by subject matter, by time, by importance, etc).
- ✓ Default settings—are they the best choice?
- ✓ Check for typos and data entry errors.
- ✓ Units are specified.
- ✓ Variable definitions are clear (possibilities: a commonly understood variable such as gender, a short phrase label such as ‘Eats >5 fruits/wk’, or footnoted to provide a longer definition).
- ✓ Statistical test names are provided, either in the p-value header or in a footnote.
- ✓ Explain any codes, abbreviations, or symbols in a footnote.

Graphs

- ✓ High resolutions graphics are used.
- ✓ Label each axis with real variable descriptor, not the R or SAS variable name.
- ✓ Label with units, if applicable.
- ✓ Thoughtfully choose axis scales and ranges.
- ✓ Removing extraneous tick marks on the axes.
- ✓ Choose bold, easily distinguishable plotting symbols.
- ✓ Axis values and lines are easy to see even if the graph is miniaturized. [try it out]
- ✓ In scatter plots, use the horizontal axis for X (predictor, independent), the vertical axis for Y (outcome, dependent).
- ✓ If graphs are in color, can color-blind folks interpret them?

Text of Results section

- ✓ In the text, address tables and graphs in logical order. Reference tables/figures at first mention, not at the end of the paragraph.
- ✓ Write about items within tables in a logical order.
- ✓ The structure (formatting) of the text eases the reader through the material.
- ✓ Provides enough detail so that the reader can form their own opinion on the quality of the results.
- ✓ Observations but no opinions (keep those for the discussion section).
- ✓ Sentences are well-crafted, fairly short, and avoid jargon. Remove redundant information. EDIT, EDIT, EDIT
- ✓ Shorter paragraphs are preferable, all else being equal.

APPENDIX

A. Example SAS template to enhance graphics

```
proc template;
   define style Styles.Scatterstyle;
      parent = styles.default;
      style GraphValueText /
         color = black
         fontsize = 12pt
         fontfamily = "Arial"
         fontweight = bold;
      style GraphLabelText /
         color = black
         fontsize = 14pt
         fontfamily = "Arial"
         fontweight = bold;
      style header /
         fontfamily = "Arial"
         fontsize = 12pt
         fontweight = bold;
      style GraphData1 /
         color = black
         contrastcolor = black
         markersymbol = "Trianglefilled"
         linestyle = 1;
      style GraphData2 /
         color = grey
         contrastcolor = grey
         markersymbol = "Circle"
         linestyle = 2;
      style GraphData3 /
         color = blue
         contrastcolor = blue
         markersymbol = "Square"
         linestyle = 3;
      style GraphData4 /
         color = green
         contrastcolor = green
         markersymbol = "Star"
         linestyle = 4;
      style GraphBackground
         "Graph backgroundcolor attributes" /
         backgroundcolor = white;
   end;
        run;

        ods html style=scatterstyle;
```

SAS code to increase graphics resolution

```
ods graphics on / width=6in height=5in  /*set the size you want, if it is too big a high
dpi might fail (uses up too much memory and there will be a warning in the log) */
                  outputfmt=png      /*png is a good format */
                  imagename="try_it"  /*name the image*/        ;


ods listing gpath="C:\SAS JUNK"  /* set the path where the image will be stored */
            image_dpi=400        /* 96 is the default, use at least 400 for high quality
reporting */;
```

B. Good coding and foolproof variable creation

a) Your code must be human readable and capable of being understood by someone else or the future you. Therefore:
   - Include a Header [purpose of code, your name, date],
   - Use many explanatory comments,
   - Use spacing and formatting so coding errors can be caught easily.

b) If you create new variables, have code that confirms they were made correctly.
Proc Print is NOT enough to check new variables.

For example, in SAS, use

```
PROC FREQ;
TABLES New_var * Old_var1 * Old_var2/ LIST MISSING;
**example of a new variable based on values of two other variables. The missing option shows you what happened to the missing values—they should be missing in the new variable, too;
RUN;
```

Something like this in R:

```
mytable <- xtabs(~newVar + oldVar1 + oldVar2 + oldVar3, data=mydata)
ftable(mytable)
```

Data will surprise you—there may be values that you did not expect, your code might have typos or just doesn't work. You have to use a fool-proof method to check; otherwise, sooner or later you will be sorry and embarrassed.
If you merge datasets, same story. Double-check the merge.

C. A short guide to naming variables and datasets

Naming Variables

Most studies have multiple sources of data and 100's or 1000's of variables. For sanity's sake, it is essential to use a clear and consistent naming convention. We suggest basing names on the system of ***Tree_branch_leaf*.** This system prevents confusion of similar variables, allows variables from different sources to stay together in a list of contents, and also aids macro processing because the formatting is predictable.

For example, imagine you have data from a form called PPP and variables recording usual sleep start time, sleep end time, sleep quality, leisure time exercise days, and exercise minutes per day. Another form, DIA, also collects data on physical activity, but uses a daily diary. In this example, the form name is the Tree, the variable topic is the Branch, and the sub-question in the topic is the Leaf.

Variable Name	**Explicit Variable Label**
PPP_sleep_begin	PPP: time falls asleep
PPP_sleep_end	PPP: time wakes up
PPP_sleep_quality	PPP: sleep quality (poor/good/excellent)
PPP_exercise_days	PPP: days per week exercises (mod/vig)
PPP_exercise_minperday	PPP: usual minutes per day of mod/vig exercise, on days with any exercise
DIA_exercise_days	DIARY: days per week exercises (mod/vig)
DIA_exercise_minperday	DIARY: avg minutes per day of mod/vig exercise on days with any exercise

Naming datasets

You will nearly certainly be making many versions of your datasets. You will also be sharing with current and future collaborators. How do you keep everything straight, avoid using an out of date version, or mixing up similar data from a different study?

To clearly identify datasets, use the naming convention ***Study_DataType_date.***

Write dates in YYYYMMDD format so that files sort in temporal order.

Example:
Two related studies, MIMS2 and MIMS3, collect the same exercise data. Datasets are updated as data is collected. Here is the sorted list of datasets:

MIMS2_exercise_20201009
MIMS2_exercise_20210308 ← most recent version for MIMS2 study
MIMS3_exercise_20220409

What the sort looks like when using ddmmmyyyy formatting and no study name:

Exercise_08mar2021
Exercise_09apr2022 (oops, mims3 data accidentally put in the mims2 folder)
Exercise_09oct2020

For more information on good practices, read:

Broman KW, Woo KH. 2018. Data organization in spreadsheets. PeerJ Preprints 6:e3183v2 https://doi.org/10.7287/peerj.preprints.3183v2

Ellis SE, Leek JT. 2017. How to share data for collaboration. *PeerJ Preprints* 5:e3139v5 https://doi.org/10.7287/peerj.preprints.3139v5