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

January 2013

This document is designed to help you with two tasks: 1) organizing the reporting of the elements from your case studies and 2) writing clearly and succinctly about statistical results. Everything covered presumes, of course, that you have correctly and responsibly analyzed the data. This guide should be consulted while writing all case reports. 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 of these 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. In particular, please note that we do not expect your papers to have any reference to other literature on the subject matter, although this would be expected in a journal article or thesis.

The basic framework for reporting your results will include these elements:
- ❖ Title
- ❖ Introduction to the research question
- ❖ Methods
    - Subjects
    - Sample and study design
    - Measurements
    - Analysis
- ❖ Results
    - Text
    - Tables
    - Figures
- ❖ Discussion and conclusion
    - Summary of important findings
    - Limitations and strengths of the study

To give you a good idea of what we expect, at the end of each section we present an example from a student's paper that was particularly well-written. We thank this student, who prefers to remain anonymous. As a counterpoint, we occasionally present a few poor examples from other students who also graciously agreed to provide their work in the hope that it will help their peers.

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**".

Student's example:
Is reduced bone mass density related to arthropathy among persons with hemophilia?

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 isn't known, needs to be clarified, or needs more evidence).

- Plainly states the intent of the present research study – its objective and what you hope to add to the knowledge base on this topic. This is generally stated in the last paragraph of the introduction. For your case study, this will be the only introduction item you need to be concerned with.

Student's example:
Introduction: This study investigates whether reduced bone mass density (BMD) is associated with arthropathy among persons with hemophilia. It also considers associations between reduced BMD and age, daily alcohol consumption, body mass index (BMI), physical activity index, family history of osteoporosis and hip fracture, and HIV drug treatment use.

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.

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?

- Time frame 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.

Student's example (there was limited information provided by the instructors about the study design and sample, so this section was by necessity brief):

This was a cross-sectional study, with data obtained on all eligible and consenting patients with hemophilia selected between May 2010 and September 2011 from Emory University's Hemophilia Treatment Center. Eighty-eight subjects were selected in total..

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".

- 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>.

  > Examples: "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 isn't 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.

  > Example: "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.

  > Example: "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."

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

- Briefly mention the testing procedures used (though 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.

  Example: "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 variable names from SAS 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 case 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 this variable and what proportion of the sample these represent. Note this in data tables as well.

Student's example:

Data were collected using clinical reports, lab reports, self-administered questionnaires and interviewer-assisted questionnaires. The outcome variable was BMD (g/cm^3), recording the lowest measured value of left and right femoral neck BMD. The primary predictor variable was radiology summary score for arthropathy (hereinafter arthropathy) assessing six major joints (possible range: 0-76, with higher values indicating worse arthropathy). Other measured risk factors included age, BMI (kg/m^2), Framingham physical activity index , daily alcohol consumption (more than 2 drinks per day or otherwise), family histories of osteoporosis and hip fracture (defined as having a parent or sibling with the disease), and HIV drug treatment use (classified as yes, no, and not HIV positive). No missing data was observed prior to data cleaning.

All variables were examined for typos or values exceeding variable ranges. One observation (1.1%) had an arthropathy score outside the range and was set to missing. No other errors were observed.

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 2 variables, and finish with those analyses involving 3 or more variables.

- State the statistical tests used – 2 sample t-test; chi-square test of association; linear regression, etc. You don't 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 don't need to put that in a journal article. This is of interest to your teachers but not to the general reader.

> Example: "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."

- Briefly report verification of assumptions for higher level analyses.

    Example: "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;

    Example: "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 case reports unless specifically requested by your instructor.

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

Student example:

All data analysis was performed using SAS 9.3 (Cary, NC). Descriptive statistics for study variables were calculated using the total study population. Linear regression models were constructed to examine bivariate associations between each risk factor and BMD. Overall F-tests for linear associations were assessed at $\alpha=0.05$. Linearity assumptions were examined by reviewing residual plots for each model fitted. Independence was assumed based on sampling protocol.

 (The student went on to describe the multiple linear regression modeling strategy, criteria for variable selection, and the assumptions of the method and how they were confirmed.)

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. 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
- The results should always be presented 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 the reporting of the primary analysis. 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.

- If your document has more than 1 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 – THE FINDINGS MUST BE INTERPRETED 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.
- Supply the reader with:
  - the **direction** of the effect, if any
    - Examples:
      > "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) [See 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).

- o **estimates and their units** (e.g., "a mean of 10 minutes of exercise per day"). Add interpretation when it might not be clear. Readers will understand a report of means, but will need the interpretation of regression estimates. (e.g., "Activity increased by 10 minutes for every 1 hour decline in TV viewing.")

- o measures of **precision,** such as standard deviation for descriptive statistics, standard error or confidence intervals for inferential statistics (e.g., "The average activity level per week was 35 minutes (*SD 15*)").

- o 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)".

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

- 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.

- Numerical values may be reported parenthetically, especially when they are reported elsewhere in tables.

  Example, "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.

- 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.*

- Put some thought into the degree of number rounding that is appropriate. For most variables, rounding to 1 digit is appropriate. Sometimes, for finer gradations, more digits are necessary. A rule of thumb is that If the original measurement was recorded to 1 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. Certainly don't use decimal places in percentages unless the denominator in the upper hundreds.

- P-values less 0.05 are typically rounded to the third digit; therefore, the smallest p-value you would report would be denoted as p < 0.001.

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 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 difference in 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 # immediately so the reader can find the data. There were 5 variables in the table. The 2 significant items are reported in detail first, while the 3 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.*}

- 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.

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. You will need to thoughtfully design a table. All aspects of the format and content should intentionally 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, describing distributions of risk factors in the primary predictor groups.

- Put some thought into the order of the variables—what order will best tell the 'story' of the data? Some possibilities are: 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 wouldn't 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.

- Clearly label the statistics—mean, std, %, 95%CI, etc.

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

- 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.

Right justified		centered		left justified	
105	15.5%	105	15.5%	105	15.5%
87	13.0%	87	13.0%	87	13.0%
51	7.6%	51	7.6%	51	7.6%

- Use <u>footnotes</u> for a) information on the # of observations with 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 in size 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.

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

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

- Nonsignificant results are also important (really!). If you tested the relationship of Y with 5 variables and found 4 with p>0.05, do not snub these poor slackers—this is useful information. Briefly report that they were not related to the outcome. (e.g., "Popularity was not significantly related to hair length, clothes brand, gpa, nor geekyness (all p>0.05)."

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. Please do not intersperse them with the text.**

Student's example

Table 1: Demographic and Clinical Characteristics of 88 individuals[*] with hemophilia from Emory University Hemophilia Treatment Center, 2009-2010.

	Mean (SD) or n (%)
Demographic	
Age, yrs	43.0 (13.5)
Consumes >2 drinks per day	9 (10%)
Physical Activity Index	32.5 (6.5)
Clinical	
Bone Mineral Density, g/cm^2	0.93 (0.19)
Radiology Summary Score for Arthropathy	15.5 (13.8)
Body Mass Index, kg/m^2	27.0 (5.9)
Family History of Osteoporosis	10 (11%)
Family History of Hip Fracture	7 (8%)
Patient has taken HIV drug	
Yes	16 (18%)
No	10 (11%)
Not HIV Positive	62 (70%)

[*] Radiology summary score for arthropathy for one observation was set to missing due to recorded value being outside the specified variable range.

Note that the # of decimal places depends on the variable.

Also, observe that %s are reported in whole numbers – as stated in the title, the denominator is 88 individuals; based on this denominator, decimal places would not be appropriate.

Accompanying text:

Eighty eight individuals with hemophilia were examined (Table 1). Mean subject age (± standard deviation) was 43.0 ± 13.5 years. Approximately 10% of subjects consumed 2 or more drinks daily and averaged a al activity index of 32.5. Regarding clinical outcomes, mean BMD was 0.93 ± 0.19 kg/m^2 and mean arthropathy was 15.46 ± 13.77. 11% and 8% of subjects had family histories of osteoporosis and hip fracture, respectively. 70% of subjects were HIV negative. Of the 26 who were HIV+, 16 were receiving HIV drug treatment.

Other Examples:

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.

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 2006.

| | 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;
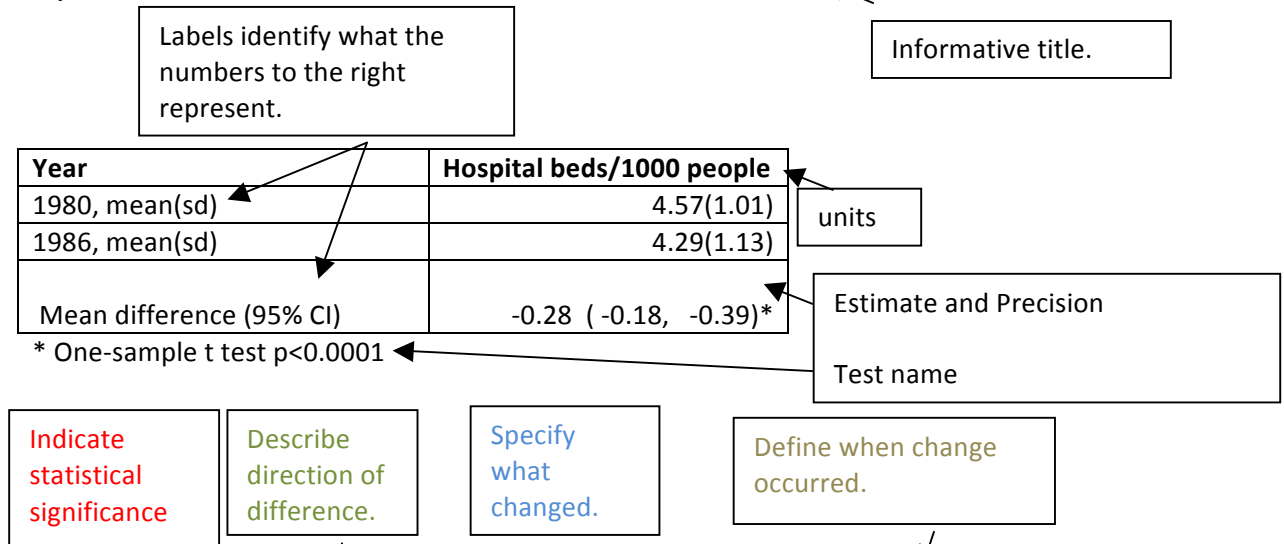**Diastolic blood pressure

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 reported in whole units.
10. Whole numbers for % are appropriate for this sample size.

**<u>Here is an annotated example of a simple table with its accompanying text.</u>**

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

Labels identify what the numbers to the right represent.

Informative title.

Year	Hospital beds/1000 people
1980, mean(sd)	4.57(1.01)
1986, mean(sd)	4.29(1.13)
Mean difference (95% CI)	-0.28 (-0.18, -0.39)*

units

* One-sample t test p<0.0001

| Estimate and Precision

Test name |
| --- |

Indicate statistical significance	Describe direction of difference.	Specify what changed.	Define when change occurred.

There were **significantly** fewer **US** hospital beds per 1000 people in 1986 than in 1980. The average decrease was 0.28 *beds per 1000 people* (95% CI -0.18 to -0.39) [Table 1].

Provide estimate and *units*.

Describe precision of estimate.	Refers to table.
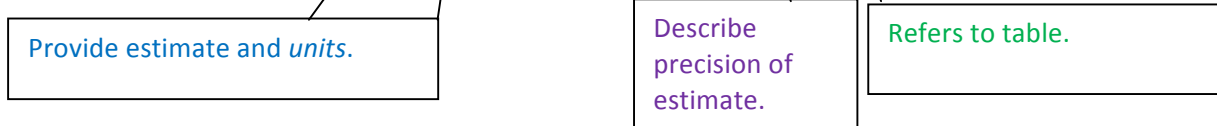
Notice that the table also provides the number of hospital beds/1000 in each of the two years (4.57 and 4.29). Although the focus of the study was on whether there was a change between 1980 and 1986, some readers would be interested in the actual values and not just the change. You might mention these values in the text, or you might just leave it to the interested reader to find them in the table.

Reporting simple linear regression results

Table 2. The association of factor Y with several patient characteristics, assessed with simple linear regression; the Zip Study, 2012.

Characteristic	Intercept estimate (se)	Slope estimate (se)	p-value for test of slope	R^2
Urban Area (ref=NYC)	1.042 (0.046)			0.18
Atlanta		-0.109 (0.047)	0.023	
Chicago		-0.219 (0.052)	<0.001	
Drinks any Alcohol (ref=no)	1.000 (0.050)	0.350 (0.043)	<0.001	0.35

Notice that the reference level is specified (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 a few 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): -0.109 (0.047) and -0.219 (0.052), respectively) [Table 1]. Urban differences explained 18% of the variability in Y.

2. The average factor Y of Atlanta residents (0.93 units) was lower than that of New Yorkers by 0.11 units (standard error for the difference = 0.05 units, p-value=0.02). Similarly, the average Y of Chicagoans (0.82 units) was lower than that of New Yorkers by 0.22 units (standard error for the difference = 0.05 units, p-value<0.001). [Editor's Comment: Decimal places in the table should probably be revised to match this reporting approach]. 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, 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 CI, 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.

Row % or Column %? It depends...

A. Sometimes there are many predictors and one or two outcomes

predictors

2 outcomes

	n	Diastolic Blood Pressure (mmHg) Mean (sd)	Has High Health Literacy (%)
Gender			
Females	Row n	Row mean	Row %
Males	Row n	Row mean	Row %
Age			
<60 yrs	Row n	Row mean	Row %
>=60 yrs	Row n	Row mean	Row %
Drug X			
Yes	Row n	Row mean	Row %
No	Row n	Row mean	Row %

A. You want to know: Among females, what is the mean DBP and % high health literacy. How do these values compare to those among males? →Therefore, the denominators are the row n.

B. Sometimes there is one predictor and many outcomes

outcomes

2-level predictor

	Intervention (n=)	Control (n=)
BMI, mean (sd)	Col mean	Col mean
Experienced Pain Relief, %	Col %	Col %
Has improved range of motion, %	Col %	Col%
Female, %	Col %	Col %
Smokes daily, %	Col %	Col%

B. You want to know: In the intervention group, what is the mean BMI? Is that statistically different from the mean BMI in the control group?

In the intervention group, what % experienced pain relief? Is that statistically different from the % experiencing pain relief in the control group?

→Therefore, the denominators are the column n.

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, 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. Both axes should be labeled, including the units (if not obvious). Again, refer to the variables using English terms, not SAS variable names.

- The figure title should meaningfully describe the content of the graphic.

- It is best to title a graph in the figure caption and not on the graph itself; if you need to change the title, it is much easier to do it in Word than rerunning the SAS code. You may 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)

Example 1:
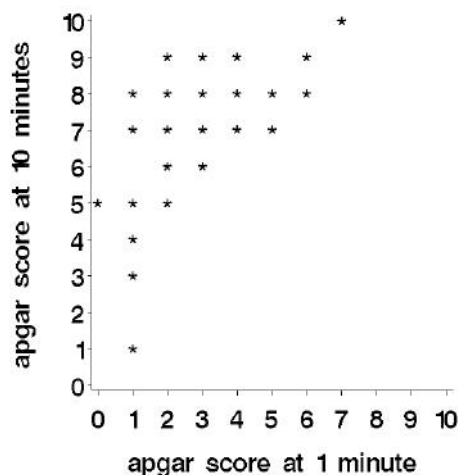


Figure 1. Among premature infants, the relationship of apgar scores at 1 and 10 minutes; Atlanta Neonate Study, 2010.

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 minute' score is on the Y axis (outcome).
4. 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.

23

Notice that this SAS option-enhanced graph can be miniaturized and still be quite easy to read; the SAS default-style graph on the right has a very poor quality when reduced.
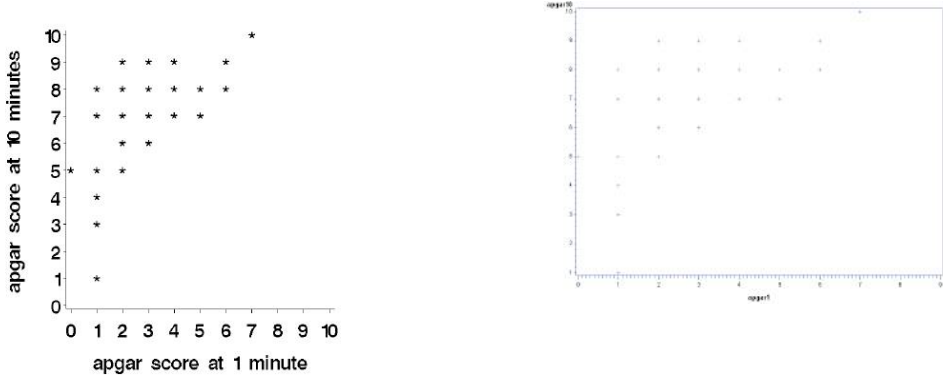
Example 2.
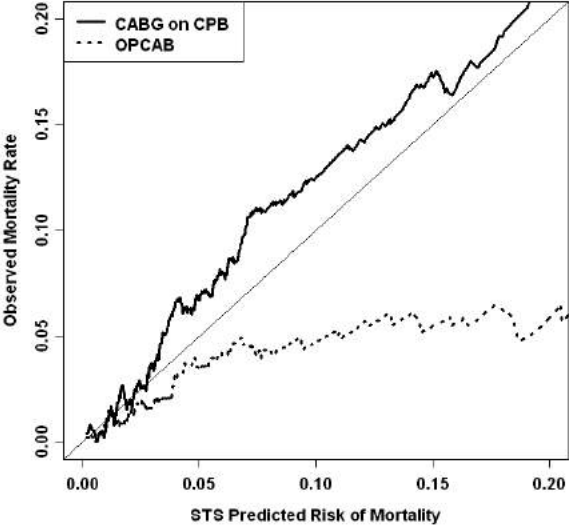
Figure 1: The relationship between predicted mortality risk and observed mortality stratified by surgery type. Each jagged line represents a smoothed curve for its respective surgery type.

<mark>Section 5: Discussion/Conclusion</mark>

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 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.").

- **The next paragraph(s) should compare/interpret present research to previous research** (consistent, inconsistent, novel, etc.).

- **Limitations and Strengths**. <u>**Do not underestimate how important this section is**</u>. 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

    - Power shortcoming. It is important in studies designed to test hypotheses to recruit enough subjects 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 didn't 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.

    - Anything else you can think of

    - Inference may not apply to a wider population (i.e., generalizability may be an issue)

    - <u>Please note that lack of a significant effect is not a limitation, it is a perfectly acceptable finding.</u>

Strengths might include:

- o a design that permits a wide inference (generalizability)

- o large sample size, adequate power to detect difference

- o study design and procedures that controlled various biases

- o 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.**

   - o Implications and recommendations for public policy and future research. You may not need to include this in your case study.

   - o Recommended extensions of the present study.

Student's example:

A multiple linear regression model for the association between arthropathy and BMD, adjusting for age, BMI and HIV drug treatment use, explained 30.6% of the total observed variation in BMD; however, in this model there was no significant independent association between arthropathy and BMD. This contrasted with the results of an unadjusted simple linear regression and challenged our initial research hypothesis. Confirming previous research, increasing age and BMI was associated with lower BMD.

Regarding study limitations, this was a cross-sectional study; therefore associative but not causal relationships can be inferred from our results. Furthermore, this was one-site sample of male hemophiliacs, which limits generalization to the population of those with hemophilia. The small sample size meant that there was limited power to detect all but the largest effects. Lastly, our model selection process did not consider interactions between covariates.

[Editor's comment: this student forgot to mention study strengths, of which there were several.]

Some Thoughts on Effective Scientific Writing

Statistical reports are expected to be written in a formal, 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 communication demands that you structure the paper so as not to distract your reader. Consider the tips below:

When writing scientific manuscripts your primary goal is to structure the paper in such a way that a reader can easily decipher the approach you took and reproduce your results if he/she desires. **When people are reading for comprehension (as opposed to reading for pleasure) they respond much better to simple, direct statements than they do to poetic prose or literary discourse.** .

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.**

Let each sentence be concise. **Don't try to pack too many ideas into a single sentence.** Some of you can't sleep at night unless you have emptied your vast vocabulary onto the pages of your reports. This only serves to irritate the reader and unnecessarily lengthens papers.

Avoid the overuse of adjectives. They seldom have a place in the scientific treatment of an issue. They can be appropriate in the Discussion section of a paper but are not typically appropriate when discussing methods or interpreting results.

Avoid the use of nebulous words like *sometimes* or *really* or *much* or *very*. These words have no quantifiable meaning and don't 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 comprehensively educate a reader about the background of a discipline – that's what references are for. Your job is to describe the current study in its scientific context and interpret your results in light of the larger body of research.

Don't "over-describe" scientific 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.

Your writing should be sequential, that is, you should start with the basics and **build the reader sequentially towards your solution.** For instance, in a Methods section first describe the subjects, then the samples, followed by measurements and finally statistical procedures.

An excellent paper is one that is well-written, organized thoughtfully, and deliberately guides the reader through the material in a way that makes results clear, unambiguous, and memorable.

As You Write Your Reports Use These Check Lists

Tables

✓ Titles are informative, 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 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.

✓ The definitions of variables are clear (either they are a commonly understood variable such as 'gender', the label is a short phrase such as 'Eats >=5 fruits/wk', or a footnote provides 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 name, not 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).

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.

✓ Items within tables are addressed 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. Remove redundant information. EDIT, EDIT, EDIT

✓ Shorter paragraphs are preferable, all else being equal.