Scientific Writing Course
October 31- November 4, 2022

Tables

Lucas Othuon

Tuesday 1 November 2022

Objective

By the end of the presentation, workshop participants should be able to:

- Construct Tables that are appropriate for scientific research publications

Tables

Tables are lists of numbers or text in columns

Why use Tables? They

- Make a greater impact than just words
- Enable relationships to be seen easily
- Condense detailed information and thus avoid the necessity for complex and repetitive sentences.
- Act as a summary of detailed information.
- Allow side-by-side comparisons of facts.
- Present data that support results

Components of a table

- Title
- Row and column headings
- The rows themselves
- The data
- Footnotes (Legend)

Note that the **table** number and title should be placed ABOVE the **table**

Table components

Table 6.1 A Descriptive Title, Such as "Structure of a Typical Table"*

This Heading Describes the Rows	This Heading Labels the First Column	This Heading Labels the Second Column
What's in the first row (units)	Data	Data
What's in the second row (units)	Data	Data

*Not all tables follow this format

The table should make sense even without the text.

Suggestions for Effective Presentation of Tables

Refer to the table **BEFORE** it appears

e.g.
See Table 1 below.

Calculations are shown in Table 3.

Full details are given in Table 4.

Suggestions for Effective Presentation of Tables cont.

- Decide on the most appropriate size, according to the amount of information to be included
- Keep your table relatively simple. Keep such additions as lines, words and labels to a minimum
- Use a key if complex information needs to be presented

Suggestions for Effective Presentation of Tables cont.

- Place the table on the same page as your discussion about it; whenever possible
- Present the table in the normal ‘portrait’ orientation, rather than 'landscape‘ unless it is absolutely necessary to do landscape.

Suggestions for Effective Presentation of Tables cont.

- Integrate the table into your text by referring to particularly significant results
- Number the tables consecutively throughout the report by using Arabic numbers.
  e.g. Table 1
- Avoid using A,B,C or Roman numbers i, ii, iii as labels because such usage is quite clumsy.

Titles: Should be descriptive enough to tell reader what will appear in the table.

Table 6.2 Poor Titles and Better Alternatives

Poor Titles	Better Titles
Characteristics of subjects	Characteristics of the 54 men enrolled in the trial
Comparison of active treatment with diuretic therapy compared with placebo in 122 men	Effects of treatment of hypertension and placebo groups
Predictors of quality of life	Factors associated with differences in quality of life: multivariate models
Independent ($p<.05$) predictors of quality of life using logistic regression following step-wise selection procedures, using the criteria of reference 6	Factors associated with differences in quality of life: multivariate models

Headings

Table 6.3 Selected Hemodynamic Measurements (Mean +/- SD) at Baseline and During Follow-up in 58 Subjects with Hypertension

	Week of Treatment*		
Measurement	**Baseline**	**1**	**6**
Heart rate (per minute)	76 ± 12	68 ± 8	65 ± 7
Systolic blood pressure (mm Hg)	162 ± 21	142 ± 18	138 ± 14
Diastolic blood pressure (mm Hg)	96 ± 12	82 ± 10	80 ± 6

*All measures showed significant differences ($p < .01$) from baseline at weeks 1 and 6.

- The headings should be informative; don't make reader refer back to the text. Use a brief description.
- Column headings reflect the comparison of primary interest.
- Column headings should be distinctive; use *italics* or **bold**.
- Put units in parentheses immediately after row descriptions.

Table formatting

1. Rules for table details will be determined by the journal --- look at tables published in the journal you have chosen for examples and follow that format.
2. Keep footnotes to a minimum; use only for essential details and abbreviations.
3. Order or number your footnotes from top to bottom and within a line, from left to right. Use these symbols *, †, ‡, §, ‖,¶. Double these symbols if you need more **, ††, etc.

Table formatting, continued

4. Put the percentage symbol (%) right next to the number if space permits, e.g. 25%.
5. Align the numbers in each column by using a centering tab function or centering the cells in the table layout.
6. Center the column headings over the columns.
7. Cite all the tables in the manuscript text.

Tables of Subject Characteristics

Table 6.10 Characteristics of the Subjects

Male	594	(49.75%)
Female	600	(50.25%)
Age	64.47 ± 5.23	
History of diabetes	103	(8.63%)
History of CHD	56	(4.69%)
Body weight	74.1 ± 7.3	
Shoe size	9.2 ± 2.1	
Calories per month	62,125.4 ± 15,781.2	

Problems: title generic; stating both male & female not necessary; mean shoe size??? extraneous & distracting ; CHD undefined, no column labels, units not provided and meaning of +/- not specified.

Table or Text?

Table 6.11 Characteristics of the 1194 Subjects Enrolled in the Better Eating Trial (BET)

Characteristic	N (%)
Male	594 (50)
History of diabetes	103 (9)
History of coronary heart disease	56 (5)
Age (yr)	64 ± 5
Body weight (kg)	74 ± 7
Calories per day	2,070 ± 530

*Plus-minus values are means ±SD

Text could read: Similar numbers of men and women were in the study; 33% of subjects were over 65 years old; 25% were more than 10 kg above ideal body weight; most were free of chronic medical problems.

Table 6.12 Characteristics of the 1194 Subjects Enrolled in the Better Eating Trial (BET)

Characteristic	Percentage or Mean ± SD
Female	50%
History of diabetes	9%
History of coronary heart disease	5%
Age (yr)	64 ± 5
Body weight (kg)	74 ± 7
Calories per day	2,070 ± 530

If actual numbers really don't matter, an acceptable alternative is to show only the percentages and the means.

Stratify the subjects into groups if there are important differences between the groups

Table 6.13 Characteristics of the 1194 Subjects Enrolled in the the Better Eating Trial (BET), By Gender

	Men (n=594)	Women (n=600)
Age (yr)	62 ± 5	66 ± 6
Body weight (kg)	80 ± 6	68 ± 8
History of diabetes	40 (7)	63(10)
History of coronary heart disease	38(7)	18(3)

*Plus-minus values are means ±SD

Differences should be pointed out in the text: Men were more than twice as likely to have a history of heart disease and diabetes was 40% more common among women.

Results from a randomized trial – stratify by study groups

Table 6.14 Characteristics of the 1194 Subjects Enrolled in the the Better Eating Trial (BET), By Randomization Status

	Special Diet (n=797)	Control (n=397)	p
Age (yr)	64 ± 5	65 ± 6	0.35
Body weight (kg)	74 ± 6	73 ± 6	0.42
History of diabetes	8%	9%	0.26
History of coronary heart disease	5%	4%	0.64

*Plus-minus values are means ±SD

- Percentages may be easier to follow especially if the numbers in each study group vary a lot.
- P-values are traditionally presented to show that the randomization worked.

Tables that tell what happened

Table 6.15 Risk of Death During 3.5 Years of Follow-up in 682 Subjects Between the Ages of 50 and 64 Years

Cause of Death	N	(%)
Cardiovascular disease	60	(8.8)
Myocardial infarction	34	(5.0)
Anterior	18	(2.6)
Inferior	12	(1.8)
Stroke	17	(2.5)
Cancer	41	(6.0)
Lung	12	(1.8)
Colon	10	(1.5)
Breast	9	(1)
Other	15	(2.2)
Total	116	(17.0)

Percentages should refer to the same denominator, e.g. # of study participants.
Additional info should go in text: “Stroke was responsible for 28% of cardiovascular disease deaths.”

Emphasize proportion of deaths due to each cause rather than the absolute risk of each cause

Table 6.16 Causes of 116 Deaths During 3.5 Years of Follow-up in 682 Subjects Between the Ages of 50 & 64 Years

Cause of Death	N	(%)
Cardiovascular disease	60	(52)
Myocardial infarction	34	(29)
Anterior	18	(17)
Inferior	12	(12)
Stroke	17	(15)
Cancer	41	(35)
Lung	12	(10)
Colon	10	(9)
Breast	9	(8)
Other	15	(13)

Tables that compare groups

- When you compare groups you are presenting either of 2 types of information

    1. The measurements or characteristics of the groups
    2. The differences between the groups

- You need to decide which is more important because it will determine how you design your table

Emphasis on the characteristics themselves; not the difference in the groups

Table 6.17 Characteristics (Mean +/- SD) of 112 Subjects Enrolled in Vacuum Away Dust (VAD) Study by Type of Pulmonary Disease

Characteristic (unit)	Asthma (n=51)	COPD* (n=66)	p
Age (yr)	32 ± 8	66 ± 6	<.001
Forced expiratory volume, 1 sec. (L)	2.5 ± 0.6	1.2 ±0. 8	<.001
Peak expiratory flow (L/min)	320 ± 110	203 ± 90	<.001
Prednisone dose (mg/day)	15 ± 20	12 ± 18	>0.25

*COPD = Chronic obstructive pulmonary disease. (This is an abbreviation that is actually helpful, because it is widely recognized and unambiguous.)

It's clear that the 2 types of subjects, asthma/COPD, are different. Only need a p-value to show differences statistically significant.

Emphasis on the comparison between groups (e.g. in a randomized trial)

Table 6.18 Effect of Intensive Vacuuming on Pulmonary Function at 6 Months in the Vacuum Away Dust (VAD) Study

Measurement (unit)	Vacuum (n=60)	Control (n=57)	Vacuum-Control Difference (95% CI)*	p
Forced expiratory volume, 1 sec. (L)	2.0 ± 0.6	1.6 ±0. 8	0.4(0.1, 0.7)	<.01
Peak expiratory flow (L/min)	290 ± 80	260 ± 120	30(5, 55)	<.02
Prednisone dose (mg/day)	10 ± 15	14 ± 12	4(-2, 6)	>0.15

*CI = confidence interval.

When emphasis is on the differences between the groups, also include whether the difference is significant, as well as a measure of the effect size, and how precise it is.

Comparing group differences in a case-control or cohort study

Table 6.19 Characteristics of the 1346 Subjects by Outcome

Characteristic (unit)	Stroke (n=122)	Controls (n=1224)
Age (yr)	72 ± 5	66 ± 6
History of diabetes	40(33%)	63(5%)
Previous MI		
None	70(57%)	1103(90%)
1	32(26%)	105(9%)
2	20(16%)	16(1%)

*Plus-minus values are means ±SD

This format doesn't provide the desired information – what predicted stroke in this study? Table doesn't show that 33% of those with a history of diabetes had strokes, but that 33% of those with stroke had diabetes.

Comparing group differences in a case-control or cohort study

Table 6.20 Incidence of Stroke by Selected Characteristics of the 1346 Subjects

Incidence of Stroke Characteristic	Stroke in Those With Characteristic	Stroke in Those Without Characteristic	RR (CI)*
Age ≥ 70 years	12% (80/660)	6% (42/686)	2.0 (1.2-3.2)
Diabetes	38% (40/103)	7% (82/1243)	5.8 (2.2-16)
Previous myocardial infarction	30% (52/173)	6% (70/1173)	5.0 (3.2-8.0)

*Indicates relative risk (95% confidence interval).

Formatting the table this way shows the desired information: which characteristics are associated with stroke. Putting the numbers in parentheses keeps reader from being distracted from the purpose of the table.

Use column subheadings for nested comparisons

Table 6.21 Association Between Smoking Status & Selected Characteristics (Mean +/- SD) in Men & Women Between the Ages of 20 and 39 Years*

	Men		Women	
Measurement	**Smokers (n=51)**	**Nonsmokers (n=62)**	**Smokers (n=33)**	**Nonsmokers (n=35)**
Weight (kg)	68 ± 8	72 ± 9	55 ± 6	66 ± 7
Hemoglobin (g/dL)	14.5 ± 2.0	13.3 ±1.6	12.2 ± 1.8	11.3 ±1.5
Leukocytes (1000 per uL)	10.3 ± 2.4	9.1± 1.4	10.9 ± 2.1	9.2 ± 1.7

*All differences between smokers and nonsmokers are significant at p <.05.

The nested comparison should be used for the most important comparison because it will be side-by-side for easier comparison. Emphasis here on smokers vs. nonsmokers. In both men & women, smokers & nonsmokers weigh less and have higher hemoglobin levels.

Tables with Many Rows and Columns

Table 6.24 Choice of Postgraduate Training Among 1567 Fourth Year Medical Students by Selected Characteristics

Characteristics	Medicine (n=219)	Psychiatry (n=407)	Pediatrics (n=125)	Surgery (n=470)	FP* (n=470)	Other (n=126)	Total
Women, %	10	40	54	45	38	23	46
Nonwhite, %	8	12	6	11	18	5	12
Varsity athlete, %	24	4	2	3	3	5	4
History of Psychotherapy, %	8	12	63	23	32	9	28
Total choosing discipline, %	14	26	8	14	30	8	100

*FP indicates family practice.

Need to orient yourself to what is being shown – in this case it's column percents. 46% of the total 4th year students are women. 14% of the students choose medicine. Of those choosing pediatrics, 54% were women. Can't tell of the women, how many chose pediatrics. If need that, switch column percents for row percents.

Presenting Results from Multivariate Analyses

Table 6.25 Independent Predictors of Coronary Heart Disease Among 2124 Middle Aged Subjects Using Logistic Regression Models

Predictor	Regression Coefficient	Standard Error	p
Sex	.51	.22	.01
Age	.05	.01	<.0001
Serum cholesterol	.3	.15	.05
Systolic blood pressure	.7	.3	.02
Smoking	1.1	.3	<.0001

Problems:

- Will the reader know what a regression coefficient is?
- What is the unit of change in the predictors --
  - does sex imply difference between men & women or vice-versa
  - is the age per year or per multiple years?
  - what is smoking; lifetime, current smoker?

Presenting Multivariate Results

Table 6.26 Independent Predictors of Coronary Heart Disease Among 2124 Middle Aged Subjects

Predictor	Relative Risk*	95% Confidence Interval	p
Male	1.7	1.1-2.6	.01
Age (per 10 yr)	1.6	1.4-2.0	<.0001
Serum cholesterol (per 20 mg/dL)	1.3	1.0-1.8	.05
Systolic blood pressure (per 10 mm Hg)	2.0	1.1-3.6	.02
Current Smoker (vs. never smoked)	3.0	1.7-5.4	<.0001

*Relative risks approximated with odds ratios from logistic regression model.

Simple fix: Use meaningful terms such as relative risk and provide units for the predictor values. Units sometimes need to be spelled out, e.g., current vs never smoked, & sometimes can be implied, e.g., men compared to women.

Presenting Univariate and Multivariate Results

Table 6.27 Univariate Predictors That Were No Longer Associated with Lung Cancer After Adjustment for Other Factors in Multivariate Models

Predictor	Univariate Relative Risk (95% CI)*	Multivariate Relative Risk (95% CI)*	Removed by
Thinness (<90% IBW)	2.1 (1.3-3.1)	1.4 (0.8-2.5)	Subject's smoking
Income (per $10,000)	0.8 (0.6-1.0)	1.0 (0.8-1.2)	age
Spouse's smoking (yes/no)	3.1 (1.5-6.2)	1.3 (0.7-2.2)	Subject's smoking
Body weight (per 5 kg)	0.6 (0.4-0.8)	0.9 (0.7-1.2)	disease stage

*Relative risks approximated with odds ratios. CI denotes confidence intervals.
†IBW=ideal body weight

It may be worth showing a table that indicates which variables were associated with outcome in univariate models, but not in multivariate models, and why. Here, thinness may have been associated with development of lung cancer in a univariate model, but may no longer be associated in a model that takes smoking into account, since people who smoke tend to be thin.

Checklist for tables

1. Is the title sufficiently descriptive without being too much/too long?
2. Do the rows and columns line up neatly? Is each column centered under its heading? Are there denominators for the column headings? Do the row characteristics have units?
3. Are there any unneeded data, repeated *N*'s, excessive precision, or ambiguous abbreviations? Ask yourself: Do I need it? Do I need it in so much detail? Do I need to abbreviate it?

Checklist for tables

4. Is the meaning of every item obvious without referring to the text?
5. After you have completed all of your tables, ask yourself: Can two or more of them be combined?
6. Are all the tables cited in the text? Are they cited in order?

What should be left out of a table

1. Don’t include everything that was measured. Pick out the important items and make your point.
2. However, don’t make this determination just by what was statistically significant. This is misleading.
3. To avoid accusations of multiple-hypothesis testing, have a few pre-specified hypotheses and indicate what they are. Report on these.
4. If you find interesting but unanticipated results, clearly state that they were unexpected.

An example of a poor Table

Table 5: Pearson's Chi-Square and Cramer's V Test for Hypothesis H_{02}

Question	df	χ^2	p-value	γ
Do at least one of your three best friends smoke at least one cigarette a day?	1	2.307	0.7	0.03

- If a table has only a few rows and columns, try stating the findings in a few sentences. Information in small tables can often be presented better in the text.
- The title should be meaningful

Points to remember

- Each journal has its own style guidelines, so always consult the publisher's Guide for Authors, also for the References list and citations format, and for the requested set-up, resolution, etc. for illustrations.

Points to remember

- Remember that the visual tools of your paper are the first visible and the most efficient way to present your results.
- How do you decide between illustrating your data with Figures or Tables? Generally, Tables give the actual experimental results, while Figures are often used for comparisons of experimental results with those of previous works, or with calculated/theoretical values.
- No illustrations should duplicate the information described elsewhere in the manuscript and remember that the legends have to be self-explanatory.

Points to remember

- Relate the tables and figures to the text. The point illustrated in the table or figure must be the point stated in the text.
- Use the fewest tables and figures needed to tell the story.
- Do not present the same data in both a figure and a table.

Acknowledgments

- For didactic material & examples:
  - Warren S. Browner MD, *Publishing and Presenting Clinical Research*, 1999
  - George Rutherford MD, UCSF
  - Mimi Zeiger, UCSF

More homework

- **Draft (revise):**
  - 3 to 4 tables
  - 1 figure (if any)
  - New analyses?
  - Revise other sections, as needed