Effective Figures and Tables

Overview

Making figures and tables can be the first step in writing your paper. You may make your figures and tables before you start to write—for instance, when you plan talks about your work. In fact, making figures and tables can be thought of as part of the scientific process; it can help you analyze and organize your data to write the paper.

The purpose of figures and tables is to best present your data so that their meaning is easily understood by readers. Figures usually show results, although they can show methods and conclusions (such as a model you propose from your data).

Although the text of your paper should not repeat information in the figures and tables, the text should introduce or refer to all your figures and tables at least once each. The order of your figures and tables will be determined by the order in which they are mentioned in the text.

**Republishing figures and tables:** If you wish to reuse or adapt a figure or table that was previously published, you may need to obtain permission. Check the publication for its copyright and licensing information. Under the traditional publishing model, if the publisher holds the copyright, you will need to get permission to reuse a figure or table. Permission is needed even if you were the author of the previous publication; after you transfer copyright, you no longer “own” your previous work and usually cannot reuse it without the permission of the current copyright holder. Publishers are increasingly using a copyright model that allows reuse (such as a Creative Commons license). In such cases, you may be able to reuse or adapt the material without permission as long as the previous publication is cited in the figure legend or in a credit line beneath the table.

**Please note:** All data presented in the figures and tables in this chapter are fictitious unless a source is cited.
Deciding Whether to Use Text, a Figure, or a Table

Use text, a figure, or a table to show a set of data—but not 2 or all 3 to show the same data. Journals have space constraints—they want to publish as much as possible, so they want you to report your work as concisely as possible. Part of that is not repeating data in text, figures, and tables.


When to Use Text

Sometimes you should describe data in the text instead of using a figure or table. In general, use text rather than a figure or table to

- Present quantitative data that can be given concisely and clearly.
- Describe simple relationships among qualitative data.

For example, a graph with 2 bars in which the control value is 100% and the test value is 60% (relative to the control) could easily be replaced with text: “The test value was 60% of the control value.”

![Figure 1. Test value relative to control value.](image-url)
When to Use Figures

In general, use a figure to

- Highlight patterns or trends in data.
- Demonstrate changes or differences over time.
- Display complex relationships among quantitative variables.
- Clarify or explain methods.
- Provide information to enhance understanding of complex concepts.
- Provide visual data to illustrate findings (for example, slides, photographs, maps).
- Illustrate scientific or clinical concepts, mechanisms, or pathophysiology.

For example, the following figure shows changes over time:

![Figure 2. Effects of the 3 treatments on tumor growth over time.](image-url)
The following figure illustrates a complex concept:

![Diagram](image)

**Figure 3.** A model for exon junction complex (EJC) assembly. EJC assembly is initiated as early as the H complex, when REF and SRp20 become associated with the 5' exon. Additional EJC components join the complex before exon ligation. As the splicing reaction proceeds through exon ligation and mRNP release, at least 3 additional proteins become closely associated with the RNA, while other factors may remain with the spliceosome. Association of some EJC proteins becomes less stable after mRNP release. Proteins identified by cross-linking are named according to their apparent molecular weights, whereas p32/REF and p25/SRp20 were identified by denaturing immunoprecipitation. Proteins lacking a molecular-mass designation (bottom section) were shown to be present in the C complex or spliced mRNP by mass spectrometry or Western blotting.

(Reprinted with permission from Reichert VL et al. 5' exon interactions within the human spliceosome establish a framework for exon junction complex structure and assembly. Genes Dev 16:2778–2791, 2002; Cold Spring Harbor Laboratory Press © 2002.)
When to Use Tables

In general, use a table to

- Present large amounts of detailed quantitative information in a smaller space than would be required in the text.
- Demonstrate detailed item-to-item comparisons.
- Display many quantitative values simultaneously.
- Display individual data values precisely.
- Demonstrate complex relationships in data.

For example, the following table clearly presents a large amount of data and many comparisons between groups:

*Table 4. Rad expression status relative to the clinicopathologic features of the cohort of invasive breast tumors studied*

<table>
<thead>
<tr>
<th>Clinicopathologic feature</th>
<th>Total (n = 48)</th>
<th>Rad-negative (n = 34)</th>
<th>Rad-positive (n = 14)</th>
<th>P value&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estrogen receptor status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>37 (77)</td>
<td>26 (54)</td>
<td>11 (23)</td>
<td>NS</td>
</tr>
<tr>
<td>Negative</td>
<td>11 (23)</td>
<td>8 (17)</td>
<td>3 (6)</td>
<td></td>
</tr>
<tr>
<td>Progesterone receptor status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>30 (63)</td>
<td>20 (42)</td>
<td>10 (21)</td>
<td>NS</td>
</tr>
<tr>
<td>Negative</td>
<td>18 (37)</td>
<td>14 (29)</td>
<td>4 (8)</td>
<td></td>
</tr>
<tr>
<td>Tumor size</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;2 cm&lt;sup&gt;2&lt;/sup&gt;</td>
<td>12 (25)</td>
<td>11 (23)</td>
<td>1 (2)</td>
<td>0.022</td>
</tr>
<tr>
<td>2–5 cm&lt;sup&gt;2&lt;/sup&gt;</td>
<td>22 (46)</td>
<td>17 (36)</td>
<td>5 (11)</td>
<td></td>
</tr>
<tr>
<td>&gt;5 cm&lt;sup&gt;2&lt;/sup&gt;</td>
<td>7 (15)</td>
<td>3 (6)</td>
<td>4 (9)</td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td>7 (15)</td>
<td>4 (9)</td>
<td>3 (6)</td>
<td></td>
</tr>
<tr>
<td>Nodal status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>18 (38)</td>
<td>15 (31)</td>
<td>3 (6)</td>
<td>0.016</td>
</tr>
<tr>
<td>1–2 nodes positive</td>
<td>9 (19)</td>
<td>8 (17)</td>
<td>1 (2)</td>
<td></td>
</tr>
<tr>
<td>≥3 nodes positive</td>
<td>14 (29)</td>
<td>6 (12)</td>
<td>8 (17)</td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td>7 (15)</td>
<td>5 (11)</td>
<td>2 (4)</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> NS, not significant.

Figures

Figures can show data visually or show complex relationships between values. They also can serve to enhance your readers’ understanding of complicated information. Types of figures include photographs, line drawings, flow charts, schematics, and graphs (graphs will be discussed separately).

Check the journal’s author instructions for specific guidelines on electronic figure preparation, including file format and resolution.

Figures Other than Graphs

Figure Guidelines

The variety of types of figures is huge, and many books have been written about creating figures. Providing advice on creating effective photographs, line drawings, and other figures is outside the scope of this chapter. However, the following points are applicable to most types of figures.

Make it easy for your readers to find the most important points in your figures and to understand what your figures show. Use arrows or other indicators to point out important items in a figure so that readers do not need to search for them. And be sure to label the items required to understand the figure so that readers can identify them easily.

Whenever possible, put information on the figure itself rather than in the figure legend. All figures need legends, but it is hard on readers if they have to switch back and forth between a figure and its legend to understand what is being illustrated. For example, label lanes on photographs of gels and autoradiographs.

Most journals will print your figures much smaller than the originals. Therefore, make labeling on figures large. Otherwise, when the figures are reduced by the publisher, the labeling will not be readable. Also, make sure that letters and symbols can be distinguished from the images on the figure (for example, do not place a black letter on a dark image or a white symbol on a very light image).

Be sure to include any relevant information that does not appear on the figure itself in the accompanying legend. General guidelines for legends include the following:

- Give enough information that the reader can understand the figure without referring to the text of the article.
▪ State briefly the message you wish the reader to receive from the figure or the most important finding evident in the figure.
▪ State original magnification and stain, if applicable.
▪ Define abbreviations and explain symbols used in the figure.
▪ Name the method used; describe the method in detail only if that is the journal’s style.
▪ Keep the legend as short as possible.
▪ Put the legends for all the figures on a separate page, not on the figures themselves.

Figures often have 2 or more parts, labeled A, B, etc. If you are considering a multipart figure, be sure to consider whether the readers need to see the parts of the figure together on the same page, either side by side or stacked vertically. Photographs showing a patient before and after a procedure, for example, might need to be printed together to facilitate comparison. If the readers do not need to see the images together, it is probably better to make them separate figures (which are easier for the journal’s production staff to work with). For example, a drawing of a surgical procedure and a flow chart showing the steps in deciding whether to perform surgery do not need to be printed on the same page because they deal with concepts that are not closely related.

Publishers vary in their policies regarding color figures. Color is best used when it aids understanding of a figure (for example, to distinguish subgroups). Color can also make a figure more attractive. However, some journals require the author to pay for color printing, which can run as much as $1,000 per figure. Check your target journal’s author instructions for its policies.

**Examples of Common Types of Figures**

Following are examples of common types of figures used in medical and basic science articles.

The first example includes photographs and a line drawing. The photographs of tissue sections from transgenic mice show that the head muscles are altered when certain genes are not expressed. Points to note include the following:

▪ In this multipanel figure, each panel is labeled with a letter, and the figure legend describes each panel by referring to the letter.
▪ Arrows indicate items of interest in the figure so that readers do not need to search for them.
- Some items in the figure are identified by labels like “ma,” which stands for “masseter muscle.” The abbreviation is defined in the legend. (Shorter words could be spelled out on the figure itself.)

- The figure includes a line drawing of mouse head muscles for readers who are not familiar with them.

- The figure was printed in the journal in color, to show the staining of the various anatomic parts.

- The legend is long because it has to state all of the following:
  - What the figure shows (the first phrase of the legend).
  - What is in each panel (A–F).
  - What each arrow or arrowhead in each panel indicates.
  - The definitions of any abbreviations used in the figure.
  - The scale or magnification of the magnified images in the figure.

---

Figure 5. Deficiency of head skeletal muscles and diaphragmatic hernia in MyoR−/capsulin−/− neonates. Shown are coronal sections from MyoR−/capsulin+/− mice (A and B) and MyoR−/− capsulin−/− mice (C and D); panels B and D show higher magnifications of A and C, respectively. gl, glands; m, mandible; ma, masseter muscle; p, palate; pt, pterygoid muscles; t, tongue; te, temporalis muscle. Asterisks in panel C denote missing muscles. (E) Diagrams of muscle groups missing from the double mutant. (F) Sagittal section of a MyoR−/capsulin−/− mouse with a diaphragmatic hernia. Arrowheads mark the boundaries of the defect. The arrow indicates the diaphragm. d, diaphragm; g, gut; l, liver; p, pancreas. Scale bars, 200 µm. (Figure reprinted with permission from Lu J et al. Control of facial muscle development by MyoR and capsulin. Science 298:2378–2381, 2002. © 2002 AAAS; legend adapted.)
The next example is a flow chart for a prospective clinical trial. Such a figure can make it easier for reviewers and readers to understand the inclusion/exclusion and stratification processes and to see what happened to all the patients.

- This figure shows the steps in the processes and the numbers of patients at each step.
- The legend provides additional information about the patients who withdrew after stratification. It would have been difficult to include such information directly on this particular figure.

Figure 6. Design of the trial. Three patients withdrew during follow-up because they moved to other regions: the 2 patients in the conventional-therapy group withdrew after 0.4 and 4.7 years of follow-up, respectively, and the patient in the intensive-therapy group withdrew after 3.2 years. CVD denotes cardiovascular disease. (Reprinted with permission from Gaede P et al. Multifactorial intervention and cardiovascular disease in patients with type 2 diabetes. N Engl J Med 348:383–393, 2003. Copyright © 2003 Massachusetts Medical Society.)
The following example models the decision-making process when screening for tuberculosis. Figures such as this not only facilitate readers’ understanding but also can be valuable aids in the clinic. These figures are sometimes referred to as “decision trees.”

- The figure shows each point in the process of deciding what tests to run and what interventions to consider. At each decision point, a choice between 2 responses leads readers to the correct next step.
- The original legend was very long (it has been truncated here). Whenever possible, information essential to the decision-making process should be included on the figure itself. Doing so was not feasible in this particular case.

Figure 7. Tuberculosis screening flowchart. . . . A chest radiograph is considered abnormal if it reveals parenchymal abnormalities; radiographs showing pleural thickening or isolated calcified granulomas are not considered to be suggestive of previous tuberculosis. . . . (Reprinted with permission from Jasmer RM et al. Latent tuberculosis infection. N Engl J Med 347:1860–1866, 2002. Copyright © 2002 Massachusetts Medical Society.)
The next example is a diagram of the effect of budesonide on methylation of the \textit{Igf-II} gene. Note the following:

- This figure summarizes a lot of data.
- The gene map shows the reader the section of the gene studied.
- The numbers across the top indicate the cytosines studied. The open circles indicate unmethylated cytosines, and the filled circles indicate methylated cytosines. The numbers to the left of each row are the identification numbers of the animals studied.
- The figure legend explains all of this; there is no good way to indicate it on the figure itself.
- The pattern is the important thing: the many filled circles in the top section indicate lots of methylation in normal lung samples, the few filled circles in the middle sections indicate little methylation, and the many filled circles in the bottom section indicate remethylation in the treated samples. It is easy to see this pattern.
Figure 8. Effect of budesonide on the methylation of the Igf-II gene. The methylation status of the CpG sites in the DMR2 of the Igf-II gene is depicted. The numbers above the circles give the location of cytosines in the CpG sites. Open circles indicate unmethylated sites, and filled circles indicate methylated sites. The numbers at the start of the strings of circles are the identification numbers of the mice. The numbers in parentheses after animal number 8-245 identify the 2 tumors from that animal. (Figure from Tao L et al. Effect of budesonide on the methylation and mRNA expression of the insulin-like growth factor 2 and c-myc genes in mouse lung tumors. Mol Carcinog 35:93–102, 2002. Copyright © 2002 John Wiley & Sons, Inc. Reprinted by permission of Wiley-Liss, Inc., a subsidiary of John Wiley & Sons, Inc.; legend adapted. Only panel A is shown; panel B and its accompanying legend have been omitted.)
This example shows electrophoretic gels. The contents of each lane are indicated above the lane. Such labeling is very helpful for the reader and should be done when there is room.

Figure 9. Detection of Bcr1-type transcripts in the 6 cell lines. (A) PML/RARA. (B) RARA/PML. (C) PML. (D) RARA. Water was used as a negative control, and NB4 Marsaille diluted 10:1 (10\(^{-1}\)), 100:1 (10\(^{-2}\)), and 1000:1 (10\(^{-3}\)) was added to HL-60 RNA for semiquantitative evaluation. The marker was the biomarker low biotin (BioVentures, Murfreesboro, TN). (Figure from Mozziconacci M-J et al. Molecular cytogenetics of the acute promyelocytic leukemia–derived cell line NB4 and of 4 all-trans retinoic acid–resistant subclones. Genes Chromosomes Cancer 35:261–270, 2002. Copyright © 2002 John Wiley & Sons, Inc. Reprinted by permission of Wiley-Liss, Inc., a subsidiary of John Wiley & Sons, Inc.; legend adapted.)
Graphs

Graphs are figures that show relationships between data or trends in data. Following are best practices for some common types of graphs: line graphs, bar graphs, and pie graphs.

**Line Graphs**

Line graphs are very commonly used in scientific articles. Line graphs are 2-axis graphs that show the relationship between 2 quantitative variables. Line graphs depict trends, such as changes over time. They are usually designed with the dependent variable (the outcome variable of interest) on the vertical axis (y-axis) and the independent variable (the variable that is thought to influence the dependent variable) on the horizontal axis (x-axis). The axes are labeled with both the variable and the unit of measurement; on linear scales, tick marks representing equal intervals must be equally spaced. Axes and their labeling should not extend far beyond the data shown. The symbols used for the data points must be defined and easily distinguishable. Axis labels should be short and printed so that they are easily readable.

![Figure 10](image_url)

Figure 10. Mean body weights (± SEM) of weanling guinea pigs fed protein-adequate (n = 10) and protein-deficient diets (n = 10) for 10 months. The guinea pigs were weighed monthly. Body weights for the 2 groups were significantly different ($P < 0.05$, Student’s $t$ test) beginning at month 2 and continuing until the end of the study.

The legend here specifies the species studied, experimental groups, number of animals used, duration of study, statistical test and significance level, and other information. Here, the symbols used to distinguish results for the 2 groups are defined in the figure itself. In some figures, that information may be found in the legend.
Bar Graphs

Bar graphs are 1-axis graphs used to compare amounts, frequencies, or magnitudes for categories of discontinuous data. Bar graphs may be vertical (a column graph) as in the example shown or horizontal (horizontal bar graph). The bars should all be the same width, and the space between bars should be less than the bar width. Each bar must be appropriately labeled.

![Bar Graph Example](image)

Figure 11. Variations by physician specialty in performance of tests during typical office visits.

The legend for this bar graph is brief because all the information necessary for interpretation is found within the graph itself.

Pie Graphs

Pie graphs can be useful when you need to show the proportions of the components of a “whole.” A pie graph shows components as wedges of a circle. When the values of the individual segments are given in percentages, they should add up to 100%. Limit the segments to those large enough to be seen and labeled. Components too small to be shown individually can be grouped into 1 wedge labeled “other.”

Each segment of the pie should be labeled. Because it is difficult to estimate the sizes of angular areas, include the percentage for each segment as part of the label. It is common but not necessary to distinguish the individual wedges with shading or color.
When there is a need to compare 2 or more wholes, you should use bar graphs with segments (corresponding to the wedges of a pie graph) rather than pie graphs because the angular areas of the wedges of the pies may make comparisons difficult. Additionally, some journals may discourage or not allow the use of pie graphs, although such figures are common in nonscientific publications.

Figure 12. Composition of gas used in study.

The legend for this pie graph is brief because the figure is largely self-explanatory.

**Graph Guidelines**

In addition to the advice given above for each type of graph, please consider the following when preparing graphs.

Keep your graphs simple. Be selective about which PowerPoint or Excel features you use. The primary goal of a graph is easy understandability. Therefore, avoid cluttering a graph with more than 3 or 4 curves. And use 3 dimensions only if you are plotting 3 variables. When the bars in a graph are 3-dimensional, readers are unable to figure out where to read the y-axis values.

Label the axes on your graphs so that readers can tell what variable is shown on each axis and what the units of measure are.

Consider your target journal’s color guidelines in deciding how to differentiate between multiple bars or lines on graphs. Use colors, shades of gray, or black and white patterns that are easy to tell apart and, ideally, that reproduce well if printed in black and white.
Make sure the symbols that represent data points are big enough to read and are easily distinguishable from each other (and if possible define them on the graph rather than in the figure legend).

Also make sure that the increments on the axes are appropriate for the data shown. For example, a small difference in values can be made to look very large or a large difference made to look very small if inappropriate increments are used. The scale of the axes should also be uniform. The values on an axis usually extend from 0 to the highest number of any data point on the graph. If all values are very high, it may be appropriate to put a break in the axis (and the bars, on a bar graph) to indicate to the reader that a set of values has been skipped on the axis.

As for all figures, provide legends for your graphs. In general, the legends should

- Give enough information that the reader can understand the graph and its intended message without reading the text.
- Define abbreviations and explain symbols used in the graph.
- Name the method used; describe the method in detail if that is the journal’s style.
- Be as short as possible.
Common Graph Problems and Proposed Solutions

Problem
The data in this graph are discontinuous, and the line graph format is thus inappropriate.

Solution
The graph has been converted to a bar graph, the preferred format for discontinuous data.

Figure 13. Number of admissions by department for March 2003.
Problems

- The symbols on the graph are too small and not easily distinguishable.
- The y-axis intervals are not uniform.
- There are no units of measure on the y-axis, and the abbreviation “conc.” is not defined.
- The x-axis extends far beyond the data set.
- The symbols are defined in the legend, not on the figure.
- The number of animals is not indicated.

Figure 14. Serum phlogiston concentrations in salamanders fed the control diet (σ), diet A (λ), and diet B (ν) for 8 weeks. Salamanders were weighed weekly. Values are means.
Solutions

- The symbol size has been enlarged.
- The y-axis intervals are now uniform.
- Units of measure have been added to the y-axis, and the abbreviated term was spelled out.
- Intervals on the x-axis now match the data set.
- A symbol key has been added to the figure (instead of being in the figure legend).
- The number of animals has been added to the legend.

Figure 14. Serum phlogiston concentrations in salamanders fed the control diet, diet A, or diet B for 8 weeks. Salamanders were weighed weekly. Values are means; n = 9 per group.
Problems

- The 3-dimensional effect is not needed and adds confusion (where do you read the y-axis value?).
- The x-axis labels (key) are off to the side, not below the axis, and are hard to distinguish.
- Abbreviations are not defined.
- There is excessive labeling on the y-axis.
- The original number of mice in each group is not indicated.

![Figure 15. Numbers of mice surviving at the end of 4 weeks of treatment.](image)

Solutions

- The 3-dimensional effect has been removed.
- The bar labels appear below the bars on the x-axis.
- Abbreviations are spelled out.
- The number of labels along the y-axis has been reduced.
- The legend includes the original number of mice in each group.

![Figure 15. Numbers of mice surviving at the end of 4 weeks of treatment for each diet group. Each group started with 225 mice.](image)
Problems

- This graph is cluttered and unreadable. Information is obscured by the amount of raw data.
- The data-point symbols (squares, circles, etc.) are not defined.
- The y-axis is not labeled.

![Graph](image)

Figure 16. Serum concentrations of X in seven men treated with Y for 15 weeks.

Solutions

- The graph’s message has been clarified by presenting the data as means with standard deviations, which also solved the problem of the undefined data-point symbols. The legend explains the error bars.
- A y-axis label has been added.

![Graph](image)

Figure 16. Mean serum concentrations of X in seven men treated with Y for 15 weeks. Error bars represent standard deviations.
Problems

- The dependent variable is plotted on the x-axis.
- The axes are not labeled with variables or units.
- The symbols are too small and virtually indistinguishable.
- The symbol key is in the legend, not within the graph, and does not identify which curve represents controls.

Figure 17. Body weights of rats treated for 10 weeks with drug A (v) or drug B (σ). Control rats received vehicle. Treatment began when rats were 2 weeks old and weighed approximately 50 grams. Each group consisted of 10 rats. Values shown are means.
Solutions

- The dependent variable is now plotted on the y-axis.
- Both axes are labeled with a variable and units.
- The symbols are larger and easier to distinguish.
- A symbol key has been added to the graph. The key identifies the control curve.
- The legend has been revised to remove the symbol key and to simplify the expression of the idea that the data are mean values.

Figure 17. Mean body weights of rats treated for 10 weeks with drug A or B. Control rats received vehicle. Treatments began when rats were 2 weeks old and weighed approximately 50 grams. Each group consisted of 10 rats.
**Tables**

Tables enable the concise presentation of large quantities of data. Tables also allow for clear comparisons of large quantities of data, including non-numeric data.

A table will typically have a number, a title, column headings and columns (at least 2 and preferably 3 or more), and row headings (called “row stubs”) and rows (again, at least 2 and preferably 3 or more). Many tables also have explanatory footnotes.

Each column should have a heading, which describes the data in the column and usually includes the units of measure (if any) for numbers in the column. Tables comparing subject (patient) or disease characteristics between 2 or more groups often have the units of measure in the row stubs instead of in the column headings. (This concept is illustrated in the section “Subject Characteristic Tables.”)

Material that is to be directly compared should be adjacent in the table. Research has shown that readers usually prefer to make comparisons horizontally (across rows) rather than vertically (down columns).

The best order of rows is the order that is most useful to readers for the type of data being presented. For example, if the rows show progression over time, then chronologic order will probably work best. If the rows show group characteristics, then ordering them by values, from largest to smallest, might work well. Rows showing related information (for example, tumors < 5 cm and tumors ≥ 5 cm) should be adjacent. Alphabetical order might be used if no other ordering principles are applicable.

Spanning column headings are used to avoid repeating the same information in multiple column headings. Indenting row stub subcategories under main categories achieves a similar goal in the row stubs. (Both concepts are illustrated in the section “Model of a Good Table.”) It is not wrong to have repeated information in column headings or row stubs. Just be aware that sometimes tables become clearer if the repeated information is separated.

The intersection of a column and a row is commonly referred to as a “cell.” Cells should be left blank only if no value is possible. If a value is possible but not known or otherwise lacking, use an em-dash (2 hyphens) or another designator that is defined in the footnotes (for example, “ND” for “not done” or “NA” for “not available”). If the value for that cell is 0 (for example, if no subjects in 1 group displayed a characteristic seen in other groups), use a 0.
Footnotes are the place to define abbreviations and special symbols used in the table and to document the source of the data in the table (if not from the current article) or indicate that the table has been reprinted with permission from another source. Be sure that the footnote symbols used in the body of the table match up with the footnotes listed below the table. Journals commonly use 1 of 2 footnote systems in tables: superscript symbols or superscript lowercase letters (a to z). If symbols are used, the typical ones are *, †, ‡, §, ¶, ||, and # [and then double symbols (**, ††, etc.) and even triple symbols if additional designators are needed].

In most journals, vertical lines are not used to separate columns in tables. The careful alignment of data in the columns makes vertical lines unnecessary. By convention, the only horizontal lines used are

- Lines above and below the column headings,
- A line below the last row of data, and
- Lines below spanning column headings.

In tables showing comparisons or subsets, it is helpful to readers if percentages are given in addition to actual values. It is also helpful to provide totals for columns, when the totals for the columns are relevant. If a column shows many sets of data for the same group, the total number might go best in the column heading.

If all the values in a column would be the same, the information should be presented in a footnote to the table or given in the text instead.

Be sure that numeric categories do not overlap or leave out possible values. For example, if 1 row stub is “tumors ≤ 5 cm” and the next is “tumors ≥ 5 cm,” then tumors of exactly 5 cm could be listed in either row or in both rows. Conversely, if 1 row stub is “tumors < 5 cm” and the next is “tumors > 5 cm,” then tumors of exactly 5 cm would not be included in either row. Therefore, the row stubs should be “tumors < 5 cm” and “tumors ≥ 5 cm” or “tumors ≤ 5 cm” and “tumors > 5 cm.”

In columns, align numbers on the decimal point (real or imaginary). If columns contain text, align the first line in each cell on the left; all lines in a cell after the first line should be indented slightly.

It is easy to make errors in numbers when you are revising tables and text to match new data. Be sure to check the numbers in all tables before you submit your manuscript.
If a table contains reference citation numbers, those references should be numbered as though they were cited in the text where the table is cited. For example, if reference 19 is the last new reference cited in the text before Table 1 is cited, then the first new reference in Table 1 would be reference 20.

A table that is really a list—in other words, a table that has only 1 column—should generally be presented in text form, which requires less space. However, a list in table format might be appropriate if the goal is to highlight or emphasize the contents.

**Model of a Good Table**

*Table 18. Frequency of types of mutations in 3 cell lines*

<table>
<thead>
<tr>
<th>Mutation type</th>
<th>No. mutations (%)</th>
<th>ABC cells</th>
<th>DEF cells</th>
<th>GHI cells</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base substitution</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G:C to T:A</td>
<td>29 (45)</td>
<td>32 (51)</td>
<td>18 (26)</td>
<td></td>
</tr>
<tr>
<td>G:C to A:T</td>
<td>24 (37)</td>
<td>11 (17)</td>
<td>20 (29)</td>
<td></td>
</tr>
<tr>
<td>G:C to C:G</td>
<td>0 (0)</td>
<td>9 (14)</td>
<td>12 (18)</td>
<td></td>
</tr>
<tr>
<td>Single-base-pair deletion</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G:C</td>
<td>4 (6)</td>
<td>6 (10)</td>
<td>9 (13)</td>
<td></td>
</tr>
<tr>
<td>T:A</td>
<td>3 (5)</td>
<td>4 (6)</td>
<td>1 (1)</td>
<td></td>
</tr>
<tr>
<td>Other mutations*</td>
<td>5 (8)</td>
<td>1 (2)</td>
<td>8 (12)</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>65 (100)</td>
<td>63 (100)</td>
<td>68 (100)</td>
<td></td>
</tr>
</tbody>
</table>

* Includes deletions of 2 or more base pairs, insertions, and other complex mutations.

- The spanning column heading [“No. mutations (%)”] is used to avoid repeating the same information in each column heading. Indented row stubs (for example, “G:C to T:A” under “Base substitution”) eliminate the need for repetition of words in the row stubs.
- Numbers in columns are aligned on the decimal point (the numbers in parentheses are aligned on the left parenthesis; this is customary in a column containing numbers outside parentheses and numbers inside parentheses).
- The table has no unnecessary lines.
- The footnote symbol in the table has a matching footnote below the table.
Subject Characteristic Tables

Clinical papers often include a table showing characteristics of study groups. Following is an example of such a table. Note that the units (in bold) appear in the row stubs.

*Table 19. Baseline characteristics of patients in the 2 study subgroups*

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Group A (n = 48)</th>
<th>Group B (n = 52)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men, no. (%)</td>
<td>24 (50)</td>
<td>30 (58)</td>
</tr>
<tr>
<td>Mean age (SD), years</td>
<td>47.8 (17.7)</td>
<td>50.8 (16.6)</td>
</tr>
<tr>
<td>Tumor classification, no. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>20 (42)</td>
<td>18 (35)</td>
</tr>
<tr>
<td>T2</td>
<td>15 (31)</td>
<td>24 (46)</td>
</tr>
<tr>
<td>T3</td>
<td>13 (27)</td>
<td>10 (19)</td>
</tr>
<tr>
<td>Histologic subtype, no. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adenocarcinoma</td>
<td>10 (21)</td>
<td>12 (23)</td>
</tr>
<tr>
<td>Squamous cell carcinoma</td>
<td>20 (42)</td>
<td>26 (50)</td>
</tr>
<tr>
<td>Other</td>
<td>18 (38)</td>
<td>14 (27)</td>
</tr>
</tbody>
</table>

As in the above example, sometimes the units for almost all the rows in a table are “no. (%).” When this is the case, you can use a footnote to indicate this and thus streamline the “Characteristic” column. When you use this footnote option, it is often helpful to place the other units of measure in the table cells rather than in the row stubs so that readers can easily distinguish between the cells indicating “no. (%)” and the cells with different units of measure. For example, the above table could be reformatted as follows (changes are in bold):

*Table 20. Baseline characteristics of patients in the 2 study subgroups*

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Group A (n = 48)</th>
<th>Group B (n = 52)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men</td>
<td>24 (50)</td>
<td>30 (58)</td>
</tr>
<tr>
<td>Mean age (SD)</td>
<td>47.8 yr (17.7 yr)</td>
<td>50.8 yr (16.6 yr)</td>
</tr>
<tr>
<td>Tumor classification</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>20 (42)</td>
<td>18 (35)</td>
</tr>
<tr>
<td>T2</td>
<td>15 (31)</td>
<td>24 (46)</td>
</tr>
<tr>
<td>T3</td>
<td>13 (27)</td>
<td>10 (19)</td>
</tr>
<tr>
<td>Histologic subtype</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adenocarcinoma</td>
<td>10 (21)</td>
<td>12 (23)</td>
</tr>
<tr>
<td>Squamous cell carcinoma</td>
<td>20 (42)</td>
<td>26 (50)</td>
</tr>
<tr>
<td>Other</td>
<td>18 (38)</td>
<td>14 (27)</td>
</tr>
</tbody>
</table>

* Values are numbers of patients (percentages) unless otherwise indicated.
However, if the units of measure for *all* the rows in a table are “no. (%),” the units are best given in a spanning column heading (shown below in bold), as follows:

*Table 21. Baseline characteristics of patients in the 2 study subgroups*

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Group A (n = 48)</th>
<th>Group B (n = 52)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men</td>
<td>24 (50)</td>
<td>30 (58)</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;50 yr</td>
<td>16 (33)</td>
<td>19 (37)</td>
</tr>
<tr>
<td>≥50 yr</td>
<td>32 (67)</td>
<td>33 (64)</td>
</tr>
<tr>
<td>Tumor classification</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>20 (42)</td>
<td>18 (35)</td>
</tr>
<tr>
<td>T2</td>
<td>15 (31)</td>
<td>24 (46)</td>
</tr>
<tr>
<td>T3</td>
<td>13 (27)</td>
<td>10 (19)</td>
</tr>
<tr>
<td>Histologic subtype</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adenocarcinoma</td>
<td>10 (21)</td>
<td>12 (23)</td>
</tr>
<tr>
<td>Squamous cell carcinoma</td>
<td>20 (42)</td>
<td>26 (50)</td>
</tr>
<tr>
<td>Other</td>
<td>18 (38)</td>
<td>14 (27)</td>
</tr>
</tbody>
</table>
**Common Table Problems and Proposed Solutions**

**Problem**

The goal of this table is to allow readers to compare the effects of vitamin E supplementation between 3 strains of rats. Research shows that readers usually prefer to make comparisons horizontally rather than vertically. In this table, the between-strain comparisons must be made vertically. Thus, a new layout is needed.

*Table 2. Effect of vitamin E supplementation on body weights of 3 rat strains*

<table>
<thead>
<tr>
<th>Strain</th>
<th>0.1</th>
<th>1.0</th>
<th>5.0</th>
<th>10.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>150.1</td>
<td>160.3</td>
<td>211.7</td>
<td>209.3</td>
</tr>
<tr>
<td>2</td>
<td>143.7</td>
<td>209.1</td>
<td>243.4</td>
<td>199.1</td>
</tr>
<tr>
<td>3</td>
<td>183.4</td>
<td>190.1</td>
<td>214.0</td>
<td>225.7</td>
</tr>
</tbody>
</table>

*Rats were fed the experimental diet for 6 weeks and were weighed at the end of the 6-week period. Values in table are mean body weights in grams.

**Solution**

The layout has been changed so that the between-group comparisons are made horizontally.

*Table 2. Effect of vitamin E supplementation on body weights of 3 rat strains*

<table>
<thead>
<tr>
<th>Vitamin E, μmol/kg diet</th>
<th>Mean body weight, g</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>150.1</td>
</tr>
<tr>
<td>1.0</td>
<td>160.3</td>
</tr>
<tr>
<td>5.0</td>
<td>211.7</td>
</tr>
<tr>
<td>10.0</td>
<td>209.3</td>
</tr>
</tbody>
</table>

*Rats were fed the experimental diet for 6 weeks and were weighed at the end of the 6-week period.*
Problems

- A description of the scores is missing.
- The row for totals is unlabeled, and the total for each column is the same (repeated information).
- There are no units for the “protein” columns.
- The 2 columns of “Protein A” data are separated by a column of “Protein B” data.
- The numbers in parentheses are not defined.
- The table has too many lines.

*Table 23. Results of immunohistochemical analysis in 84 tumor specimens*

<table>
<thead>
<tr>
<th>Score</th>
<th>Protein A (membranous)</th>
<th>Protein B (cytoplasmic)</th>
<th>Protein A (nuclear)</th>
<th>Protein C (membranous)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>84 (100)</td>
<td>84 (100)</td>
<td>84 (100)</td>
<td>84 (100)</td>
</tr>
<tr>
<td>1</td>
<td>24 (29)</td>
<td>35 (42)</td>
<td>65 (77)</td>
<td>70 (83)</td>
</tr>
<tr>
<td>2</td>
<td>25 (30)</td>
<td>34 (40)</td>
<td>10 (12)</td>
<td>10 (12)</td>
</tr>
</tbody>
</table>

Solutions

- “Score” is defined in a footnote.
- The row of totals has been deleted because the title gives the total and the columns are short.
- A spanning column heading has been added to provide units.
- The “Protein A” columns are now side by side.
- The numbers in parentheses are defined in the spanning column heading.
- The vertical lines and unnecessary horizontal lines have been removed.

*Table 23. Results of immunohistochemical analysis in 84 tumor specimens*

<table>
<thead>
<tr>
<th>Score*</th>
<th>Protein A (membranous)</th>
<th>Protein A (nuclear)</th>
<th>Protein B (cytoplasmic)</th>
<th>Protein C (membranous)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>24 (29)</td>
<td>65 (77)</td>
<td>35 (42)</td>
<td>70 (83)</td>
</tr>
<tr>
<td>1</td>
<td>25 (30)</td>
<td>10 (12)</td>
<td>34 (40)</td>
<td>10 (12)</td>
</tr>
<tr>
<td>2</td>
<td>35 (42)</td>
<td>9 (11)</td>
<td>15 (18)</td>
<td>4 (5)</td>
</tr>
</tbody>
</table>

* 0, no staining; 1, moderate staining; 2, pronounced staining.
Problems

- The table arranges data in a way that does not allow easy comparison of the “<50 Gy” data with the “≥50 Gy” data.
- The table has many empty cells, a sign of poor table design.

Table 24. Recurrences of solid tumors stratified by total radiation dose

<table>
<thead>
<tr>
<th>Radiation dose</th>
<th>Overall</th>
<th>Recurrent tumor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. patients</td>
<td>%</td>
</tr>
<tr>
<td>&lt;50 Gy</td>
<td>107</td>
<td>43.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥50 Gy</td>
<td>139</td>
<td>56.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Solutions

The information from column 1 has been moved to the column headings, along with the totals from column 2. The information from column 4 now appears in column 1. The new design takes less space and provides easier comparisons for readers by putting data to be compared in adjacent columns.

Table 24. Recurrences of solid tumors stratified by total radiation dose

<table>
<thead>
<tr>
<th>No. recurrences</th>
<th>No. patients (%) by radiation dose</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt;50 Gy</td>
</tr>
<tr>
<td>N=107</td>
<td>73 (68.2)</td>
</tr>
<tr>
<td>N=139</td>
<td>34 (31.8)</td>
</tr>
<tr>
<td>≥1</td>
<td>21 (19.6)</td>
</tr>
<tr>
<td>2</td>
<td>9 (8.4)</td>
</tr>
<tr>
<td>≥3</td>
<td>4 (3.7)</td>
</tr>
</tbody>
</table>
Activity 1

Choosing the Best Format to Present Data

Here is an exercise in deciding how to display data in papers. You will work in groups of 5; each group will get a “bag of data.” The data are represented by chips. The chips in your bag may differ in color or pattern or may be organized by the day on which they were collected.

Decide how to best depict these data—that is, in a figure (a graph or other kind of figure), a table, or text. Then draw or write your depiction on an overhead transparency and be prepared to discuss the reasons for your choice.
Possible Solutions to Activity 1, Choosing the Best Format to Present Data

Bag 1.

Contains 4 smaller bags, labeled as follows:
- Day 1 (contains 2 white chips)
- Day 5 (contains 1 red and 5 white chips)
- Day 10 (contains 3 red and 10 white chips)
- Day 15 (contains 7 red and 17 white chips)

![Figure 1. Number of chips by day of assessment.](image)

Bag 2.

Contains 2 white chips with elaborate patterns on their fronts and backs.

Photographs of the fronts and backs of the chips.

![Figure 2. Photographs showing fronts and backs of chips.](image)
**Bag 3.**

Contains 10 blue chips (1 with dots on 1 side and 1 with dots on both sides) and 11 white chips (6 with dots on 1 side and 1 with dots on both sides).

*Table 3. Distribution of dots on chips*

<table>
<thead>
<tr>
<th>Dotted sides</th>
<th>Blue</th>
<th>White</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>8 (80)</td>
<td>4 (36)</td>
</tr>
<tr>
<td>1</td>
<td>1 (10)</td>
<td>6 (55)</td>
</tr>
<tr>
<td>2</td>
<td>1 (10)</td>
<td>1 (9)</td>
</tr>
<tr>
<td>Total</td>
<td>10</td>
<td>11</td>
</tr>
</tbody>
</table>

**Bag 4.**

Contains 6 red chips and 6 blue chips.

Text: “There were 6 red chips and 6 blue chips.”