# Writing *Methods* & *Results* Sections

Darko Hren
University of Split Faculty of Philosophy

# Methods

- Concrete steps taken to gather data

# Methods

- Typically:
  - Data sources – studied individuals, objects...
  - Instrumentation – tools
  - Procedures

  - Use past tense for report of completed research

# Methods - guidelines

- Decide whether to use the term ***subjects*** **or** ***participants*** to refer to the individuals studied
- Describe the **informed consent** procedures, if any, as well as steps taken to maintain confidentiality.

# Methods - guidelines

- The participants should be described in **enough detail** for the reader to visualize them.
- When a sample is very small, consider providing a description of individual participants.

# Methods - guidelines

**Example 9.3.1**

Participants were 418 sixth and seventh graders ($M$ = 12.4 years) who attended an ethnically diverse middle school with a sixth-grade through eighth-grade configuration. Ethnicity of participants was determined by self-report. On the research questionnaire, students identified themselves as belonging to one of six groups that captured the range of ethnicities in the school. Twelve respondents did not answer the ethnicity question, resulting in a final sample of 406 students for this study. The ethnic breakdown of the study sample was as follows: 29.1% Latino ($n$ = 118; 59 girls, 59 boys), 28.6% African American….

At the principal's request, no data on individual student socioeconomic status (SES) were gathered. In terms of school-level indicators, 68% of the students were eligible for free or reduced-price lunch programs, and the school qualified for Title I compensatory education funds. Thus, by available indicators, the school population primarily was low SES.[3]

# Methods - guidelines

**Example 9.3.2**

Table 1

*Demographic Characteristics of the Participants*

| Characteristic | Number | Percent |
|---|---|---|
| Gender | | |
| Girl | 81 | 72 |
| Boy | 31 | 28 |
| Current grade level | | |
| Third | 16 | 14 |
| Fourth | 37 | 33 |
| Fifth | 55 | 49 |
| Sixth | 4 | 4 |
| Qualify for subsidized school lunch program? | | |
| Yes | 99 | 88 |
| No | 13 | 12 |
| Family status | | |
| Living with both parents | 72 | 64 |
| Living with one parent | 35 | 31 |
| Living with neither parent | 5 | 4 |

# Methods - guidelines

- A **population** should be named, and if only a sample was studied, the **method of sampling** should be described.
- Explicitly acknowledge **weaknesses** in sampling.

# Methods - guidelines

- Provide detailed information on **nonparticipants** when possible.
- If there was **attrition**, state the number who dropped out, the reasons for dropping out, if known, and information on the dropouts, if available.

# Methods - guidelines

- **Unpublished instruments** should be described in detail.
- For **published instruments** - describe the traits it was designed to measure, its format, and the possible range of score values.
- For both unpublished and published instruments, information on **reliability** and **validity**, when available, should be reported.

# Methods - guidelines

- Experimental procedures, equipment, and other mechanical matters should be described in **sufficient detail so that the study can be replicated**.

# Presenting numbers in text, tables and figures

# Summarizing data

- Numerical precision
- Percentages
- Categorical data
- Continuous data

# Numerical precision

- ...report all numbers with the appropriate degree of precision.
  - *The number of women physicians increased from 29 942 to 94 322, and that of men from 13 410 to 36 061.*
  - *The number of women physicians increased from 29 900 to 94 300, and that of men from 13 400 to 36 000.*
  - *The number of women physicians increased from 30 000 to 94 000, and that of men from 13 000 to 36 000.*

    - *!!! Reader can effectivelly deal with numbers which contain no more than two significant digits! Thus numbers should be rounded to two significant digits unless more precision is truly neccessary.*

# Reporting percentages

- ...always give the numerators and denominators of the calculations.
  - *25% (650/2598);*
  - *33% (30 of 90 patients);*
  - *12 of 16 participants (75%).*

# Reporting percentages

- ...when the sample size is greater than 100 report percentages to no more than one decimal place.
- ...when sample size is less than 100, report percentages in whole numbers.
- ...when sample size is small (less than, say, 20), consider reporting only the absolute numbers.

# Summarizing categorical data

- Sample presentation:
  - *Of 25 tumors, only 5 were malignant.*
    - *The ratio of malignant to non-malignant tumors was 1:4.*
    - *The proportion of malignant tumors was 5/25 or 0.2.*
    - *The percentage of malignant tumors was (5/25)x100%, or 20%*
    - *After 5 years follow-up, the tumor was malignant in 5 of the 25 patients, giving a 5-year recurrence rate of 20%.* (A rate is associated with time factor!)

# Summarizing categorical data

- ...specify the denominators of rates, ratios, proportions and percentages.
- ...if continuous data have been separated by “cutpoints” into ordinal categories, identify the cutpoints and the rationale for choosing them.

# Summarizing continuous data

- Sample presentation:
  - ...if the data are approximately normally distributed, they are appropriately described with mean and standard deviation:
  - *Antibody titers ranged from 25 to 347 ng/mL and had mean (SD) of 110 ng/mL (43 ng/mL)*
  - ...if data are markedly non-normally distributed, they are appropriately described with the median and interquartile range:
  - *Antibody titers ranged from 25 to 347 ng/mL, with a median (interquartile range) of 110 ng/mL (61 to 159 ng/mL).*

# Summarizing continuous data

- ...provide appropriate measures of central tendency and dispersion when summarizing data that have a continuous distribution.
  - ...measures of central tendency indicate where on the continuum data tend to cluster.
  - ...measures of dispersion indicate the spread of data over the continuum.
  - ...distributions that form a bell-shaped curve are said to be normally distributed and can correctly be described by mean and standard deviation; other distributions are better described with the median and the range or interquartile range.

# Summarizing continuous data

- ...do not summarize continuous data with the mean and the standard error of the mean (SEM).
  - ...the SEM is not a descriptive statistic and should not be used as such.
  - ...the SEM is correctly used only to indicate the precision of the estimated mean of population. Even then, however, the 95% confidence interval is preferred.

# Summarizing continuous data
# - normally distributed data

- ...use mean and standard deviation (SD) only when describing approximately normally distributed data.
  - ...data described with a standard deviation that exceeds ½ the mean are not normally distributed (assuming that negative values are not possible).
- ...report mean and standard deviation to no more than one decimal place more than the data they summarize.
- ...do not use the "±" symbol when presenting the mean and standard deviation.

# Summarizing continuous data - non-normally distributed data

- ...describe markedly non-normally distributed data with the median and range (the minimum and maximum values) or the interquartile range (the values at the 25th and 75th percentiles).
- eg.
  - *Median weight was 72kg (25th percentile=60kg; 75th percentile=87kg).*
  - *Median weight was 72kg (interquartile range=60-87kg).*
  - *After 8 weeks, weight (median and interquartile range) was 72kg (60 to 87kg).*

# Tabular display of data and statistics

- Functions of tables:
  - ...**condense or summarize** large amounts of data, especially complex or detailled data.
  - ...**organize and display** data, especially exact numbers, more clearly and concisely than can be done with words.
  - ...**compare** individual values or groups of data.
  - ...improve the **ease and speed** with which specific information can be located and understood.
  - ...**facilitate calculations**.
- ...in general **tables** are better than figures for helping readers find **specific information** and for displaying **exact numbers**, whereas **figures** are better for showing **comparisons** and **overall patterns** in the data.

# Tabular display of data and statistics

- Components of tables :
  1. Table number.
  2. Table title.
  3. Column headings.
  4. Row headings.
  5. Data.
  6. Horizontal lines.
  7. Expanded forms of abbreviations used in the table.
  8. Footnotes referenced in the table, below the expanded abbreviations, usually used in the following order: *, †, ‡, §, ¶, ‖, **, ††, and so on.

[Table Number] **Table 20.1** [Title] **Components and Nomenclature of Tables**

| Row Head | Spanner Head | | | | Column Head |
|---|---|---|---|---|---|
| | Column subhead: Group size (units) | Column subhead: Group size (units) | Column subhead: Group size (units) | Column subhead: Group size (units) | Group size (units) |
| | **"Cut-in" Head** | | | | |
| **Row head** | | | | | |
| Row subhead | | | Data Field | | |
| Row subhead | | | | | |
| **Row head** | | | | | |
| | **"Cut-in" Head** | | | | |
| **Row head** | | | | | |
| Row subhead | | | Data Field | | |
| Row subhead | | | | | |
| **Total** | | | | | |

ABC = expanded abbreviation
* (asterisk)
† (dagger)
‡ (double dagger)
§ (section mark)
‖ (parallel mark)
¶ (paragraph mark)
** (double asterisk)
[a] superscript lowercase letter
[b] superscript lowercase letter
[c] superscript lowercase letter

# Tabular display of data and statistics

- Principles of table construction
  - ...tables should have a **purpose**; they should **contribute and be integrated** with the rest of the text.
  - ...the purpose of the table should determine its form.
  - ...tables should be organized and formatted to **assist readers** in finding, seeing, understanding, and remembering the information.
  - ...values to be compared should usually be placed side by side.
  - ...organize tables visually as well as functionally (but attend to guidelines for authors).
  - ...data presented in tables should not be duplicated elsewhere in the text.

- ...the purpose of the table should determine its form.

**Table 20.5** Leading Causes of Cancer Death in the United States, 1998, Organized to Show Patterns in the Data

| Type of Cancer | Number of Deaths |
|---|---|
| Lung | 160,000 |
| Colorectal | 57,000 |
| Breast | 44,000 |
| Prostate | 39,000 |
| Cervical | 5,000 |

**Table 20.6** Leading Causes of Cancer Death in the United States, 1998, Organized to Help Readers Find Information

| Type of Cancer | Number of Deaths |
|---|---|
| Breast | 44,000 |
| Cervical | 5,000 |
| Colorectal | 57,000 |
| Lung | 160,000 |
| Prostate | 39,000 |

- ...tables should be organized and formatted to assist readers in finding, seeing, understanding, and remembering the information.

**Table 20.7** Baseline Fitness Data for the Eight Patients Completing the Fitness Training Study, as Collected

| Characteristic | Patient Number | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| Age, years | 35 | 16 | 21 | 19 | 41 | 30 | 22 | 37 |
| Sex | M | M | F | F | M | F | M | F |
| Resting pulse, beats/min | xx | xx | xx | xx | xx | xx | xx | xx |
| 1.5-mile run, min | xx | xx | xx | xx | xx | xx | xx | xx |
| Hemoglobin, g/dL | xx | xx | xx | xx | xx | xx | xx | xx |
| Weight, kg | xx | xx | xx | xx | xx | xx | xx | xx |
| $10^6 \times$ RBCs/μL | xx | xx | xx | xx | xx | xx | xx | xx |

RBC = red blood cells.

**Table 20.8** Baseline Fitness Data for the Eight Patients Completing the Fitness Training Study, Revised for Publication

| | New Patient Number [Old patient number, for illustration only] | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Men | | | | Women | | | |
| | 1 [3] | 2 [1] | 3 [6] | 4 [5] | 5 [4] | 6 [7] | 7 [2] | 8 [8] |
| Age, years | 16 | 22 | 35 | 41 | 19 | 21 | 30 | 37 |
| Weight, kg | xx | xx | xx | xx | xx | xx | xx | xx |
| Blood values | | | | | | | | |
| Hemoglobin, g/dL | xx | xx | xx | xx | xx | xx | xx | xx |
| $10^6 \times$ RBCs/μL | xx | xx | xx | xx | xx | xx | xx | xx |
| Fitness results | | | | | | | | |
| 1.5-mile run, min | xx | xx | xx | xx | xx | xx | xx | xx |
| Resting pulse, beats/min | xx | xx | xx | xx | xx | xx | xx | xx |

RBC = red blood cells.

- ...values to be compared should usually be placed side by side.

**Table 20.9** Effect of Physician Experience on Surgical Simulation Scores

| Experience | Mean Surgical Simulation Scores* | | |
|---|---|---|---|
| | Cutting | Suturing | Ablating |
| Residents ($n$ = 12) | 79 | 63 | 80 |
| Fellows ($n$ = 8) | 88 | 87 | 91 |
| Staff Surgeons ($n$ = 15) | 96 | 92 | 97 |
| $P$ value† | 0.03 | 0.004 | 0.05 |

* Scores range from a low of 0 to a high of 100.

† ANOVA. Differences were significant only between residents and staff surgeons.

**Table 20.10** Effect of Physician Experience on Surgical Simulation Scores

| Surgical Simulation | Mean Simulation Scores (low = 0; high = 100) | | | $P$ Value* |
|---|---|---|---|---|
| | Residents ($n$ = 12) | Fellows ($n$ = 8) | Staff Surgeons ($n$ = 15) | |
| Cutting | 79 | 88 | 96 | 0.03 |
| Suturing | 63 | 87 | 92 | 0.004 |
| Ablating | 80 | 91 | 97 | 0.05 |

* ANOVA. Differences were significant only between residents and staff surgeons.

# Visual display of data and statistics

- Figures can:
  - ...**reveal underlying patterns** of data and deviations from these patterns that are not possible with text or tables.
  - ...**organize and display data**, especially patterns of data and group comparisons, more clearly and concisely than can be done in text or tables.
  - ...**condense or summarize** large amounts of data more effectively or efficiently than can be done in text or tables.
  - ...improve the **ease and speed** with which specific information can be located and understood.

# Visual display of data and statistics

- Components of figures:
  1. Figure number.
  2. Figure caption.
  3. Data field.
  4. Vertical scale.
  5. Horizontal scale.
  6. Labels for each scale.
  7. Data.
  8. Reference lines.
  9. Keys and legends.

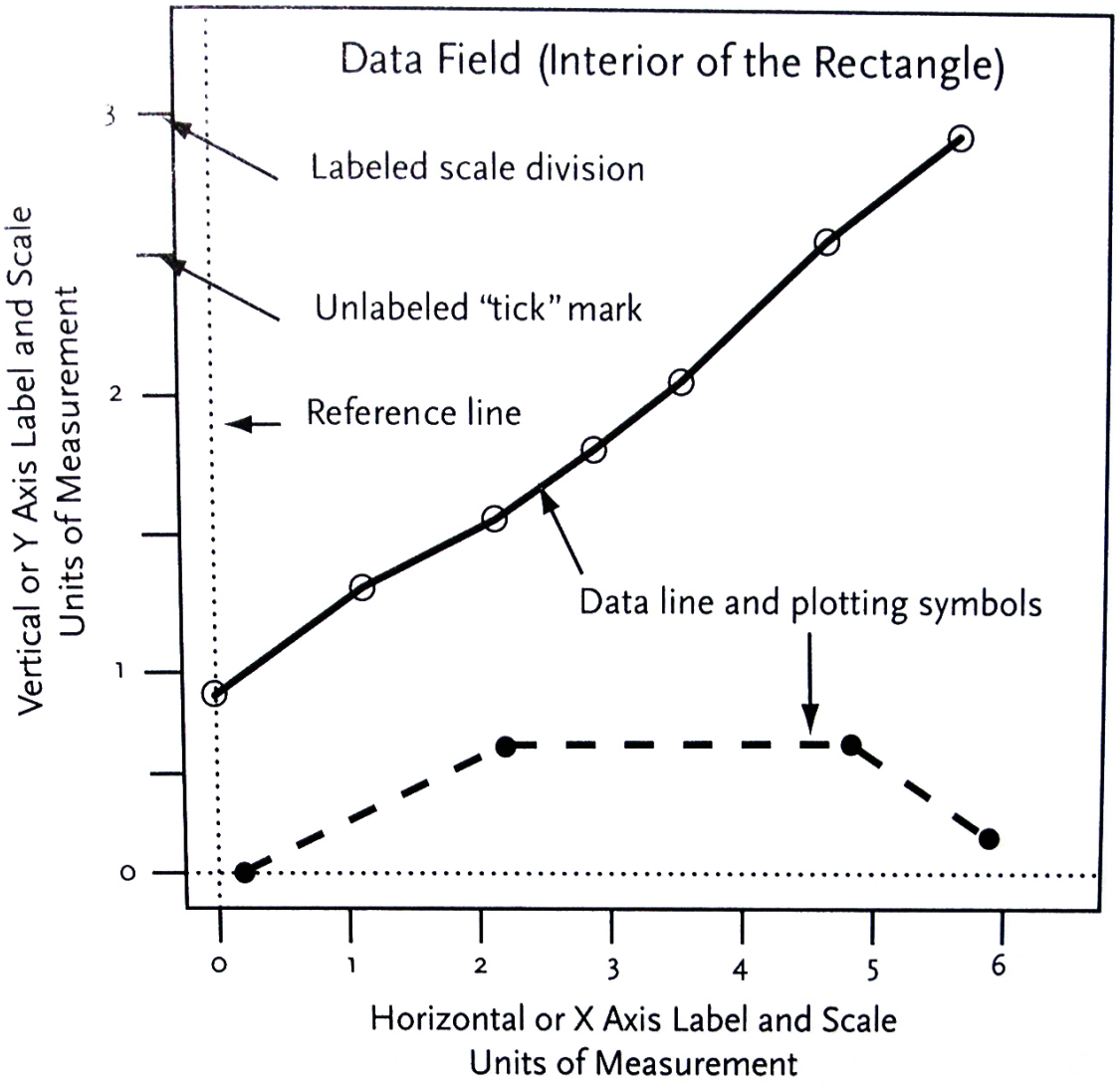



**Figure 21.1**

Components of a typical scientific figure. The data field is enclosed by a rectangle and is offset so that the zero-zero axis lines are distinct. Scale divisions and unlabeled "tick" marks are outside of the field.

# Visual display of data and statistics

- Types of figures:
  - **Publication figures** – usually in black and white.
  - **Presentation graphics** – usually in color, to be seen at a distance.
  - **Electronic figures** – may have color, motion, links...

  - ...**figures generated by statistical programs are rarely suitable** for commuinicating results of a study in any medium, let alone publication.
  - ...as a rule, figures need to be redrawn if they are to be viewed in a different medium.

# Visual display of data and statistics

- Principles of figure construction:
  - ...figures should have a **purpose**; they should **contribute to and be integrated** with the rest of the text.
  - ...figures should be designed to **assist readers** in finding, seeing, understanding, and remembering the information.
  - ...figures should contain **only those elements that are necessary** to fulfill their purpose.
  - ...the **data should be emphasized** over other elements in the figure.
  - ...figures should be **consistent with the principles of perceptual psychology**.
  - ...data presented in figures should not be duplicated in the text.

# Writing captions

- ...the caption should identify the data in the data field.
  - *Figure 8. Total manuscript submissions.*
    (Too general)
  - *Figure 8. Number of manuscripts received annually by Croatian and international scientific journals between 2000 and 2010, by country of origin.*
    (More specific)

# Writing captions

- ...the caption should allow the figure to be understood without reference to the text.
  - *Figure 8. Number of manuscripts received annually by Croatian and international scientific journals between 2000 and 2009, by country of origin.*
    (missing important qualifying information)

  - *Figure 8. Number of manuscripts received annually by Croatian and international scientific journals between 2000 and 2010, by country of origin* ***(47 journals responding to the 2009 survey of 55 journals with IF between 1 and 5)****.*
    (qualifying information included)

# Constructing the data field

- ...minimize the number of non-data elements in the data field

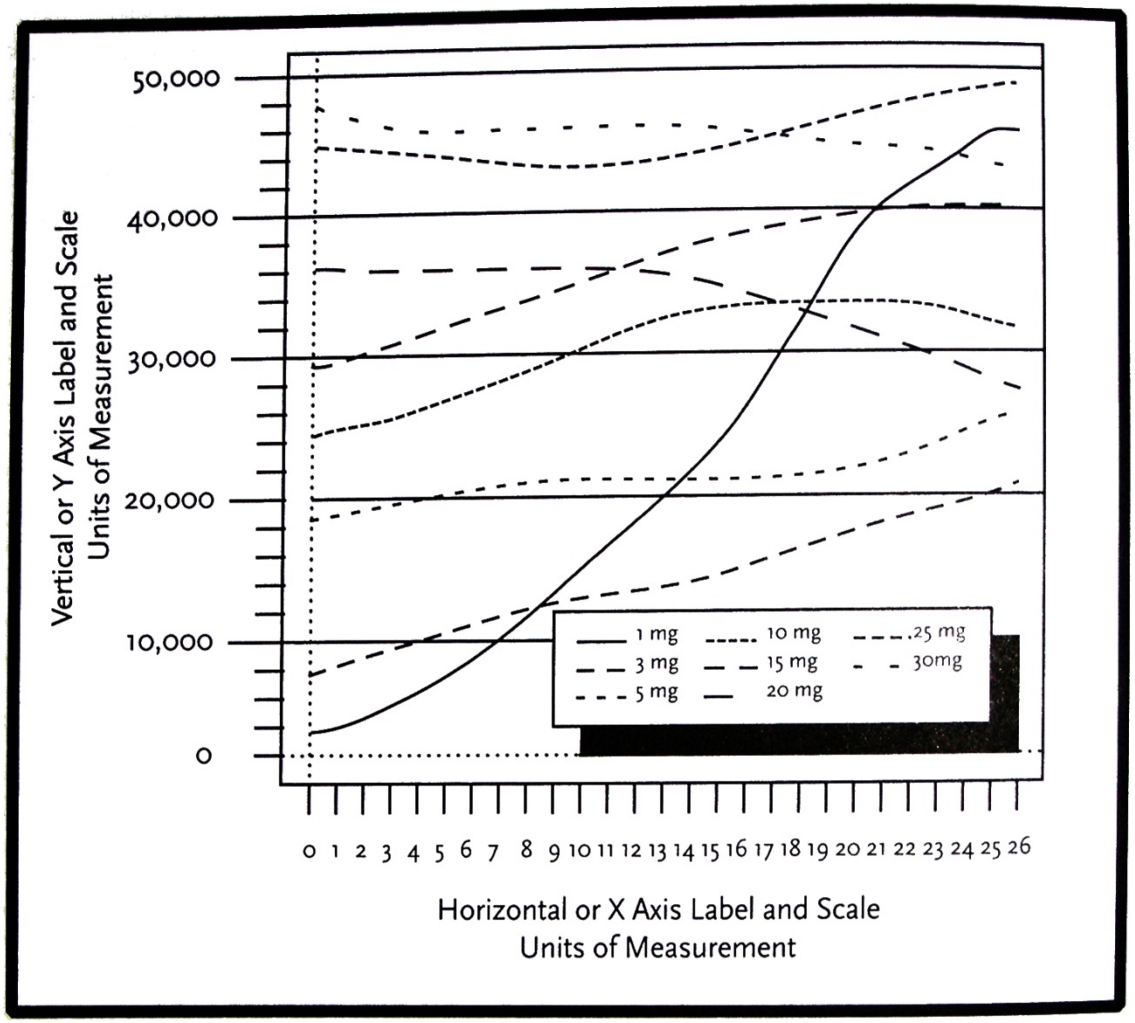



**Figure 21.2**

A figure containing several unnecessary elements. The Y scale shows too many zeros; a unit multiplier in the scale label is preferred. Too many values are labeled in the X scale, and both scales have too many divisions and tick marks. The number of data lines is excessive, and they are labeled with a legend, rather than directly. The horizontal lines intended to aid in abstracting values more accurately are ill-advised; accuracy is better achieved with a table. The third dimension in the legend (the "drop shadow") adds no information, only visual confusion. Finally, all the lines are the same weight, so readers are not encouraged to focus on the data.

# Constructing the data field

- ...identify all the elements in the data field.
  - ...place **labels close to the element they label**. When direct labeling is not possible, a key or legend can be included in an empty area of the data field or in the caption (also, consult guidelines for authors).
  - ...always **identify values indicated by "error bars"**.

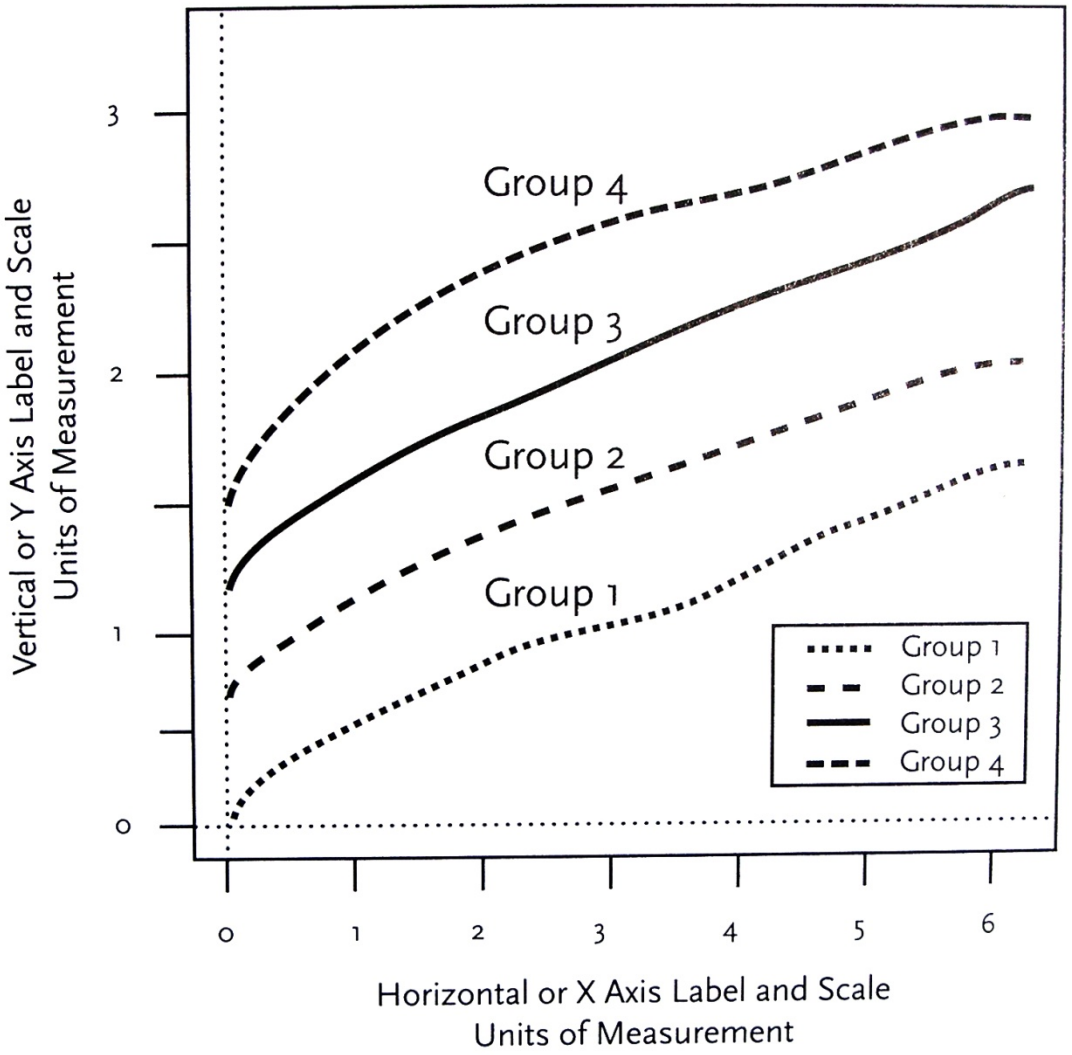



Figure 1. Changes in Y over time. (Group 1 = ..... ; Group 2 = - - -
Group 3 = ——— ; Group 4 = - - - )

**Figure 21.3**

Place labels as close as possible to the elements they name. If direct labeling is not possible, use a key or legend. Place the key or legend in the data field if it will not detract from the data and in the caption if it will.

# Constructing scales (axes)

- ...**label each scale clearly** with the name of the variable, the units in which the variable is graphed, and any multipliers associated with the units.
- ...indicate the **zero-zero point** of the graph, especially if one or both scales do not begin at zero.
- ...**adjust the scales** so that data fill as much of the data field as possible.
- ...scale divisions should mark **only major, logical, and (usually) equal intervals** on the scale.
- ...depict **three dimensions only when the data are three dimensional** and require a third scale (a Z axis).

...indicate the zero-zero point of the graph, especially if one or both scales do not begin at zero.

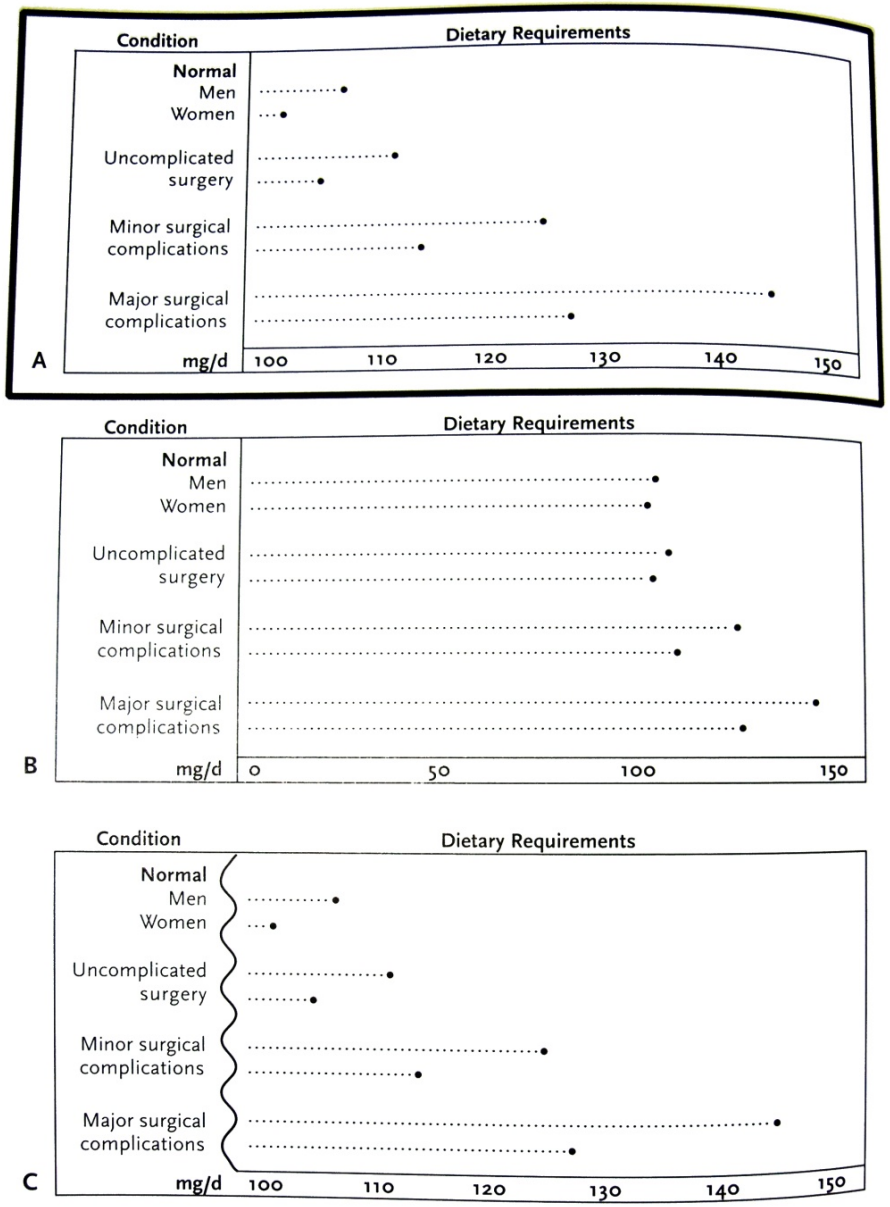



**Figure 21.6**

Visual distortions: the "suppressed zero" problem. In **A,** the visual impression is that the values for women are about half those of men. This impression is created because the baseline is not zero, as expected, but rather at a value of 100. In **B,** the zero baseline has been preserved, giving an accurate visual representation of group sizes. In **C,** the typical solution for preventing the suppressed zero problem is shown. The wavy line is more effective at indicating a scale break than is a straight line because is it less likely to seen as a typical figure element.

...adjust the scales so that data fill as much of the data field as possible.

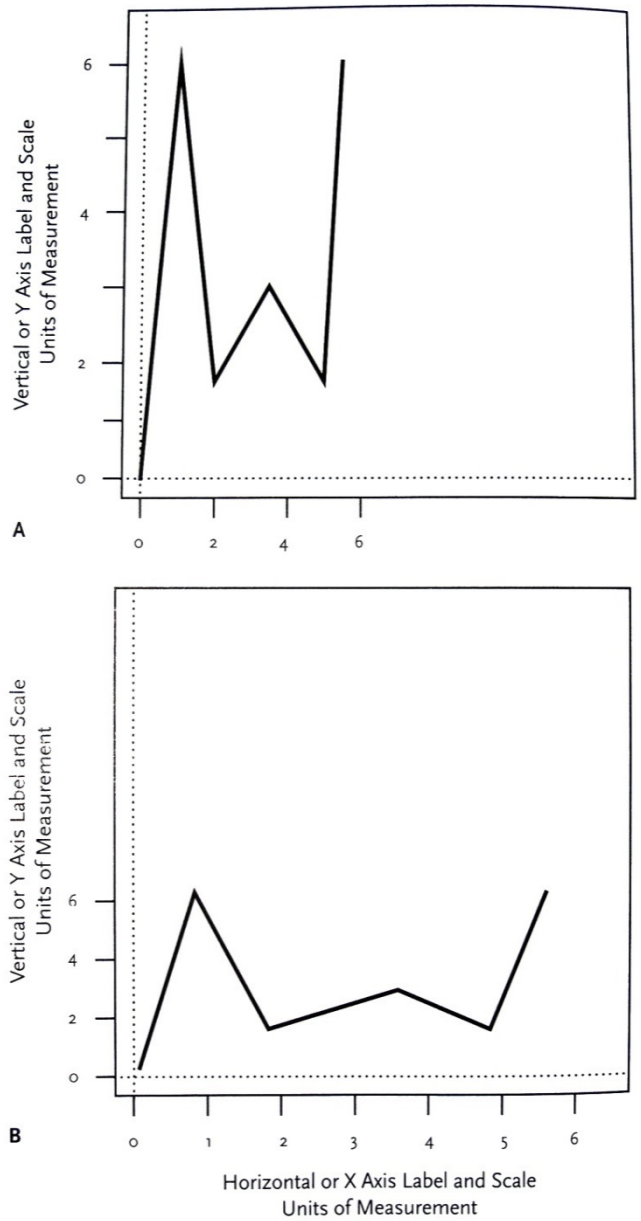



**Figure 21.7**

Visual distortions: the "elastic scale" problem. Here, the same data provide different impressions, depending on the relative length of each scale. In general, the scales should be constructed to display the data over the largest space practical, to increase their resolution. Data falling more or less along the diagonals have the best resolution for both axes.

...scale divisions should mark only major, logical, and (usually) equal intervals on the scale.

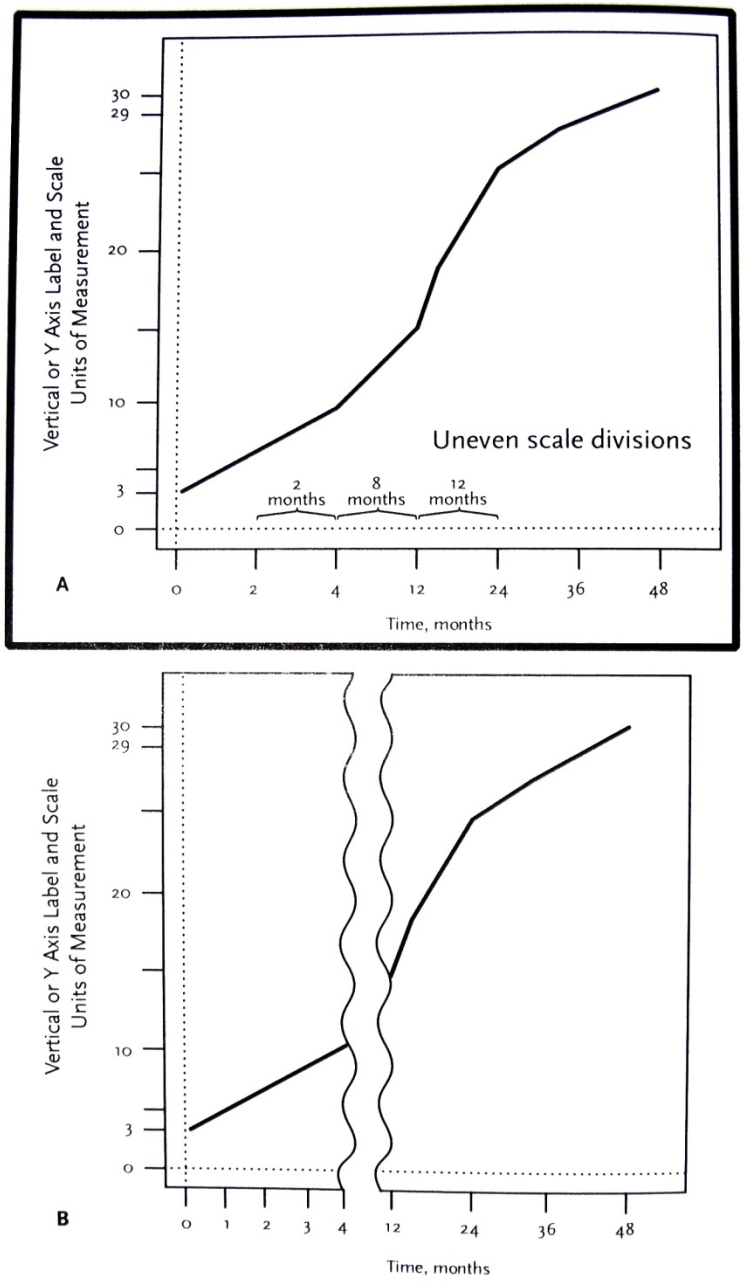
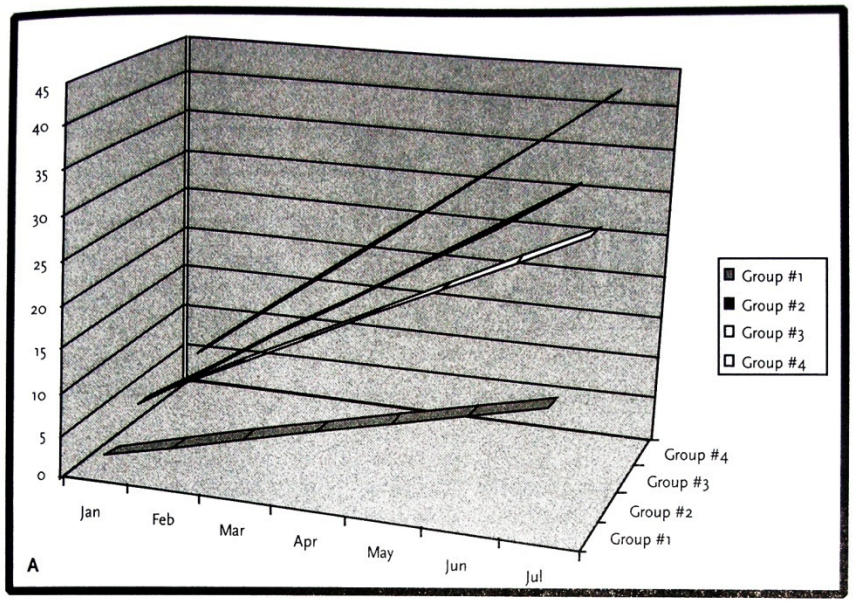



**Figure 21.9**

Equal scale intervals should represent equal units. **A,** The divisions of the X axis here do not represent equal amounts of time. **B,** The same figure drawn correctly by showing a scale break.

...depict three dimensior only when the data are, fact, three dimensional and require a third scale (a Z axis).



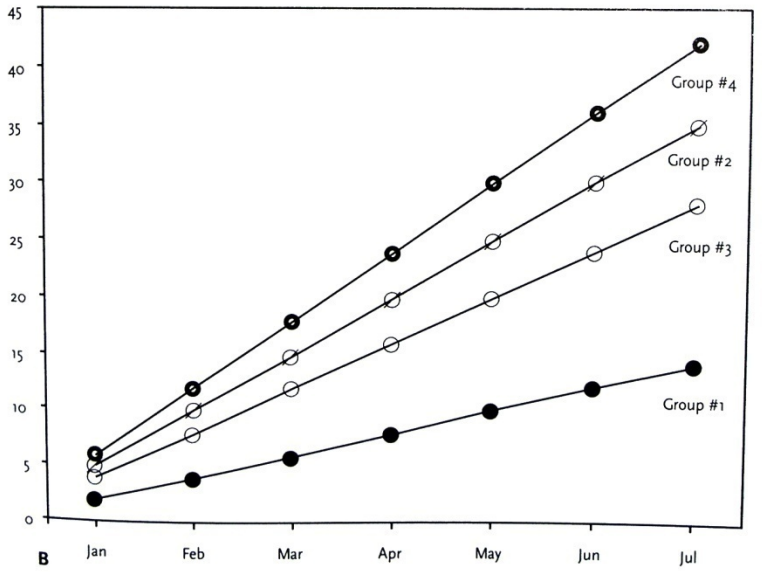




**Figure 21.10**

**A,** Adding a third dimension to a figure for artistic reasons adds no useful information and can even reduce the ability of readers to compare the data. **B,** The corresponding two-dimensional figure is less cluttered and the data more easily compared.

# Presenting individual values – categorical data

- ...when possible, use a *dot chart*.
- ...if you *must* use column or bar chart:
  - ...use it only as presentation graphics.
  - ...make sure that the space between the columns differs from the width of the columns so that columns are not confused with background.
  - ...fill all columns with contrasting shades of gray to keep them distinct from each other and from the background.
  - ...remember that extreme contrasts of gray or color can exagerate or diminish the visual importance of some columns.
  - ...do not fill columns with lined or striped fill patterns because they can create optical illusions that detract from data.

...when possible, use a *dot chart*

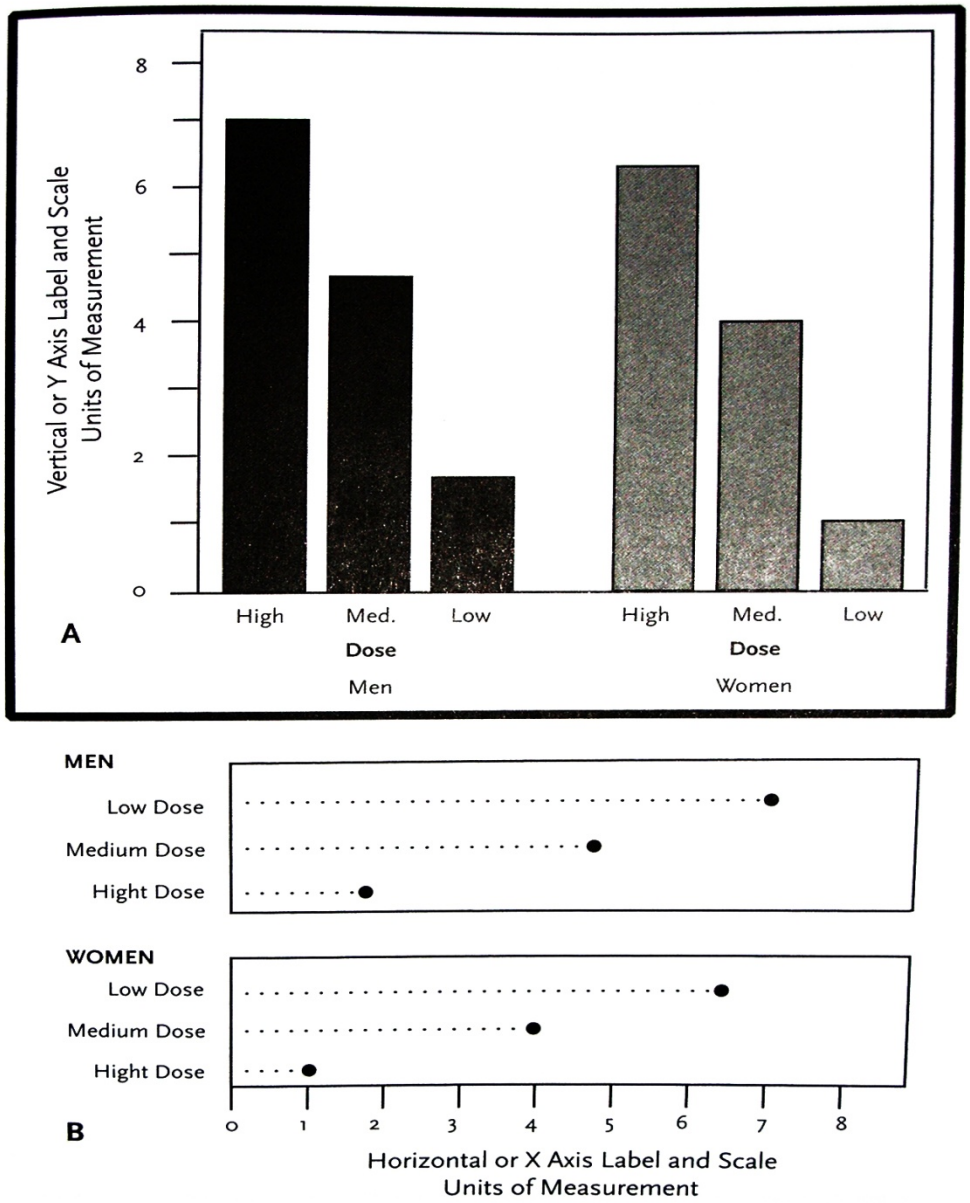



**Figure 21.13**

Vertical column charts (or horizontal bar charts), **A,** require more space then the more efficient dot chart, **B**. The information imparted by columns is at the end of the column, so the column width, color, or shading detract from the information. The dot chart focuses attention on the data, not on the columns. In addition, most dot charts can present text horizontally, making it easier to read.

...do not fill columns with lined or striped fill patterns because they can create optical illusions that detract from data.

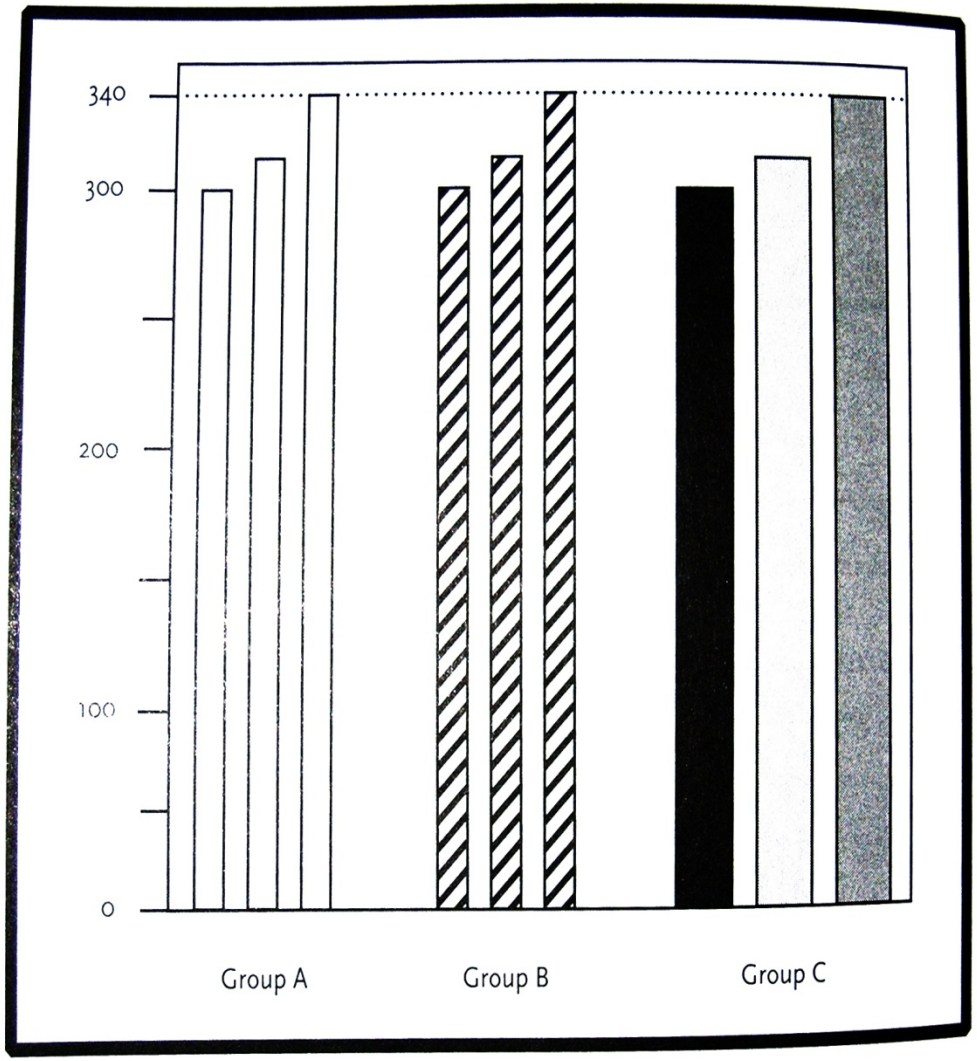



**Figure 21.14**

Common problems with column charts. Columns should always be shaded and distinct from the background, and the space between columns should be different than the width of the columns (Group A). Optical effects should be avoided by using uniform shading, rather than patterned fills (Group B). The contrast between columns should not be so severe that it calls undue attention to one column (Group C).

# Presenting individual values – categorical data

- ...as a general rule – do not use pie charts.
- ...if you *must* use it:
  - ...use it only as a presentation graphic.
  - ...make it only 2-dimensional; 3D pie charts can distort perceptions of what is important.
  - ...use it only to show percentages that total to 100%; do not give absolute numbers because readers will automatically add them, expecting summ to be 100.
  - ...limit the number of slices to no more than 5, if possible.
  - ...start the largest slice at 12 o'clockposition, and fill the rest of the pie going clockwise so that the smallest slice is the last in the circle.
  - ...no slice should be less than 5% of the total (an angle of 18 degrees).

# Presenting continuous data

- ...box plots are most informative; do not use bar and column charts.

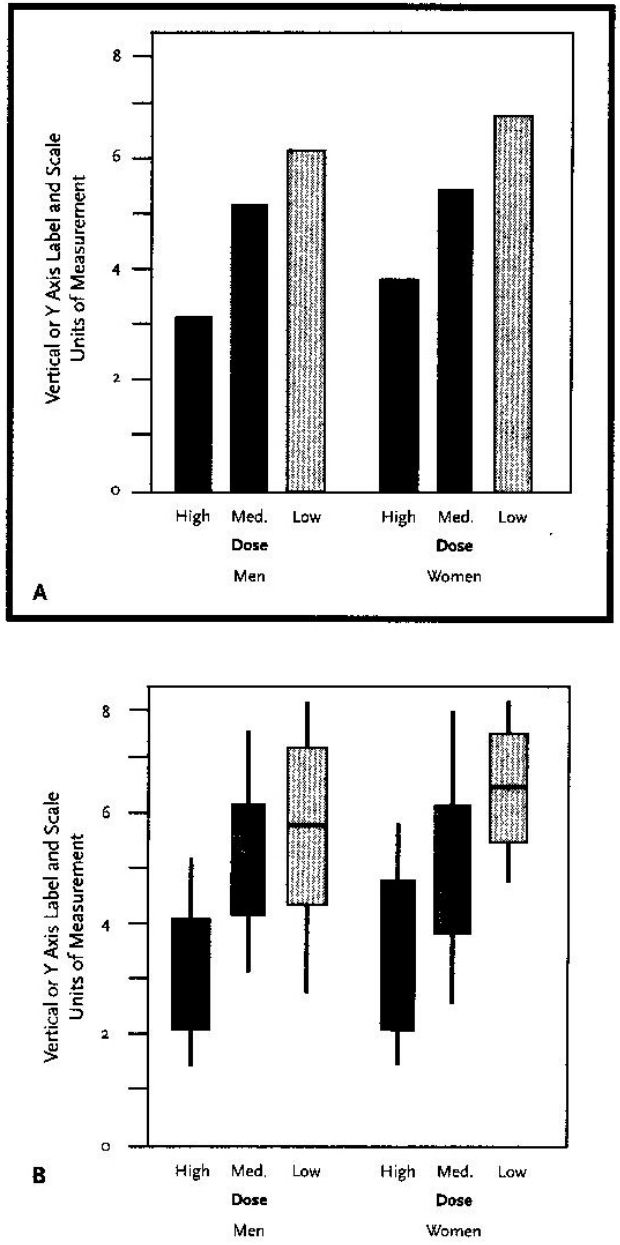



**Figure 21.15**

**A,** Column charts are often used to report continuous data, usually by plotting the mean and sometimes the standard deviation of the distribution with error bars. **B,** However, Tukey's box plot presents more information about the distribution and does so more efficiently.

- Lang, T. & Secic, M. (2006). Reporting Statistics in Medicine. Philadelphia: American College of Physicians