



Department of
Scientific
Publications

Creating Effective Graphs

Sunita Patterson

THE UNIVERSITY OF TEXAS
MDAnderson
Cancer Center

Making Cancer History®

What we will cover

- When to use a graph
- Types of graphs
- General principles for effective graphs

What we will cover

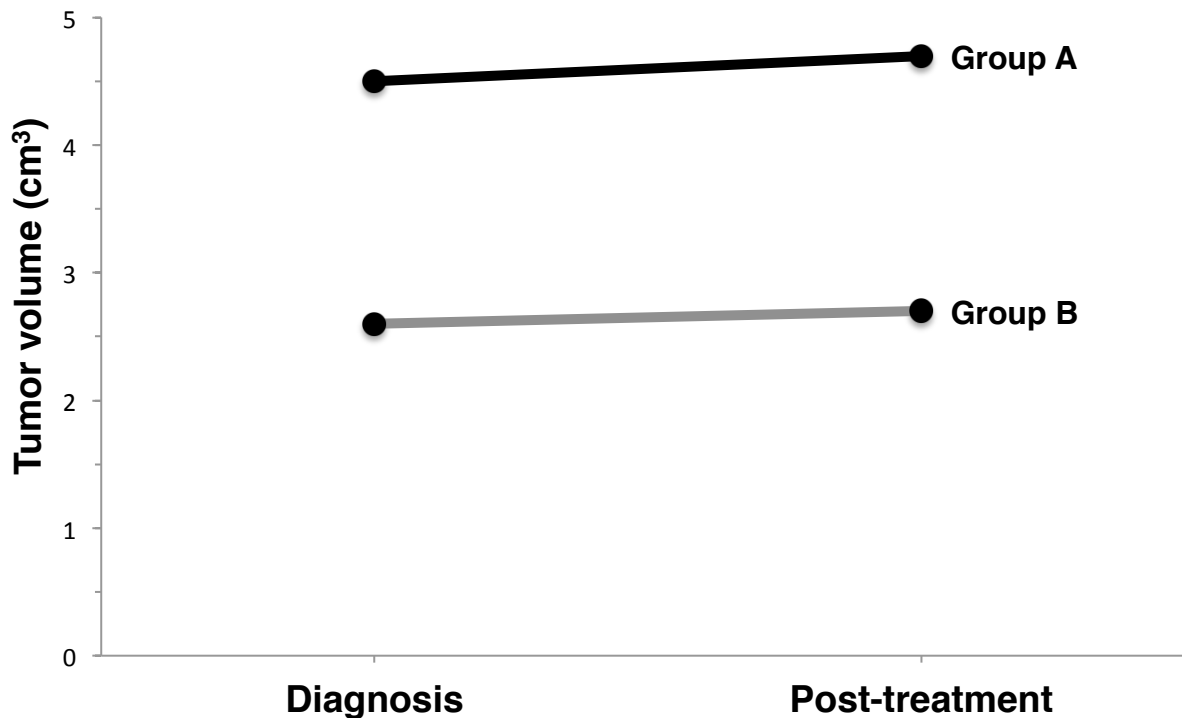
- **When to use a graph**
- Types of graphs
- General principles for effective graphs

When to use text

- Only a few data points
- No variation in the data

When to use text

At diagnosis, the mean tumor volumes were 4.5 cm³ for group A and 2.6 cm³ for group B. At the end of treatment, mean volumes had not significantly changed (4.7 cm³ for group A and 2.7 cm³ for group B).



When to use a table

- To present more than a few precise numeric values
- To show detailed item-to-item comparisons or complex relationships between data

When to use a table

Table 1. Adverse events observed in a randomized phase III clinical trial of UFT vs. S-1

	UFT (n = 254)				S-1 (n = 251)				p-value [†]
	All grade		Grade 3+4		All grade		Grade 3+4		
	n	(%)	n	(%)	n	(%)	n	(%)	
Leukopenia	40	(15.7)	2	(0.8)	91	(36.3)	13	(5.2)	0.004
Neutropenia	14	(5.5)	0	(0.0)	58	(23.1)	9	(3.6)	0.002
Thrombocytopenia	17	(6.7)	0	(0.0)	32	(12.7)	5	(2.0)	0.030
Anemia	33	(13.0)	1	(0.4)	70	(27.9)	5	(2.0)	0.121
Total bilirubin increase	17	(6.7)	1	(0.4)	45	(17.9)	2	(0.8)	0.622
AST increase	23	(9.1)	3	(1.2)	25	(10.0)	4	(1.6)	0.723
ALT increase	14	(5.5)	2	(0.8)	20	(8.0)	2	(0.8)	1.000
Fatigue	34	(13.4)	3	(1.2)	79	(31.5)	5	(2.0)	0.502
Anorexia	34	(13.4)	3	(1.2)	74	(29.5)	4	(1.6)	0.723
Weight loss	14	(5.5)	1	(0.4)	28	(11.2)	1	(0.4)	1.000
Rash/desquamation	6	(2.4)	1	(0.4)	40	(15.9)	5	(2.0)	0.121
Hyperpigmentation	3	(1.2)	0	(0.0)	48	(19.1)	9*	(3.6) *	0.002*
Diarrhea	7	(2.8)	1	(0.4)	24	(9.6)	2	(0.8)	0.622
Mucositis/Stomatitis	5	(2.0)	0	(0.0)	32	(12.7)	6	(2.4)	0.015
Nausea	8	(3.1)	0	(0.0)	27	(10.8)	0	(0.0)	-
Vomiting	1	(0.4)	0	(0.0)	10	(4.0)	0	(0.0)	-

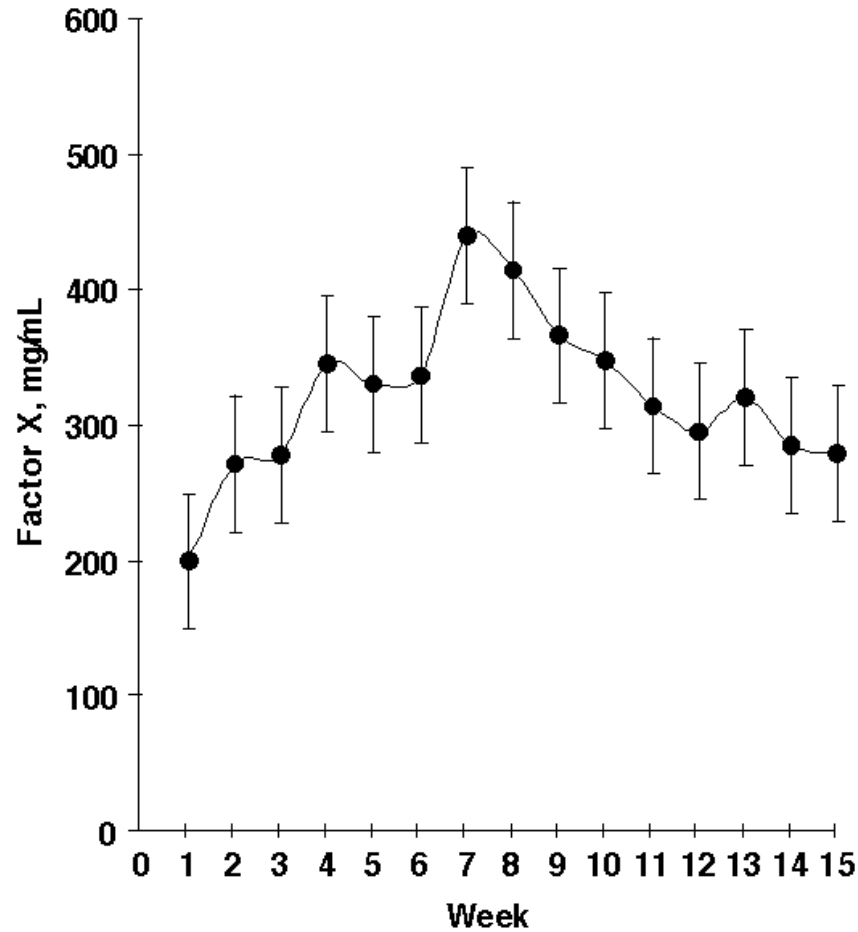
† Fisher's exact test

* Number of grade 2 events are indicated, and differences in the incidences of grade 2 are tested.

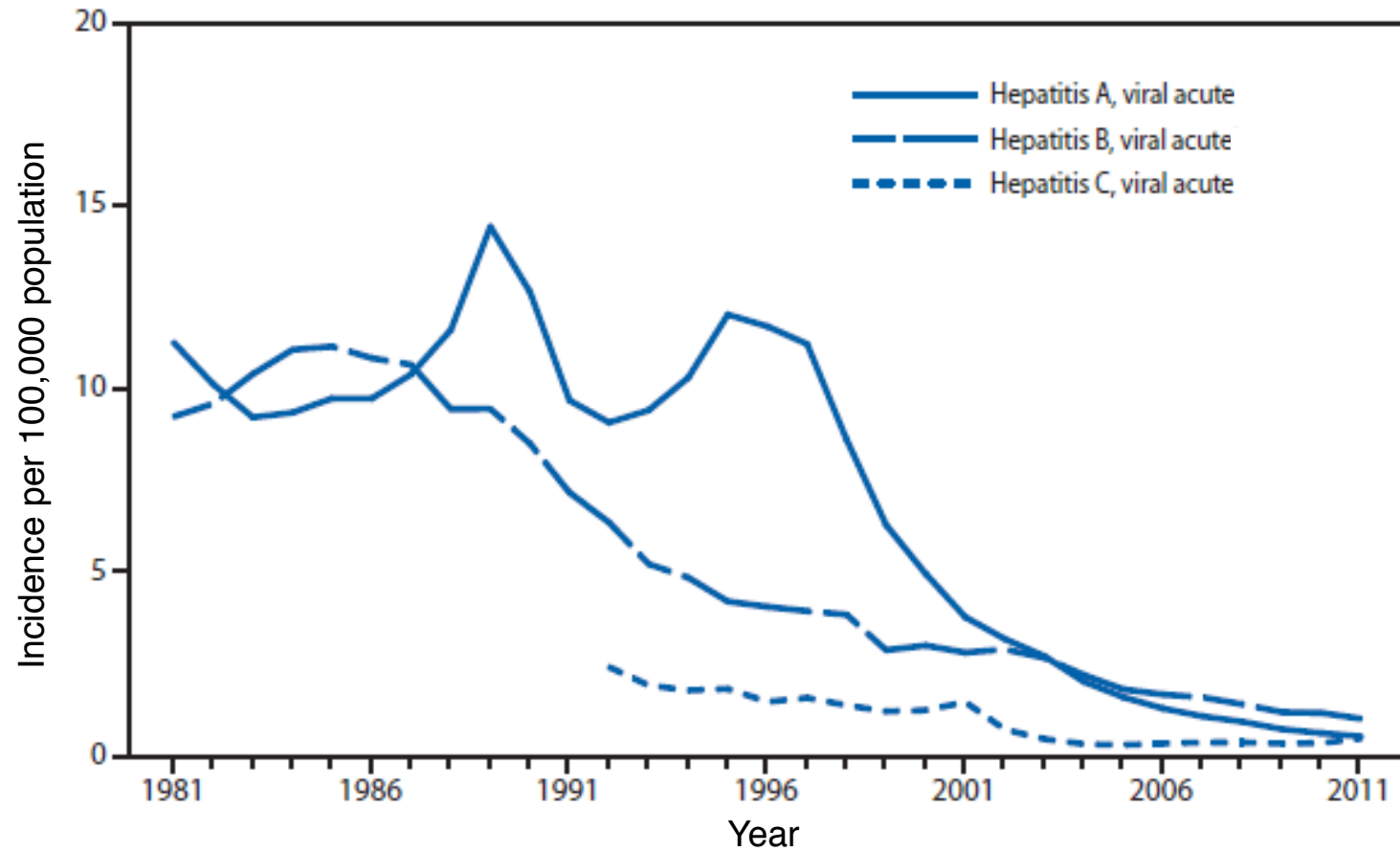
When to use a graph

- To show trends in data
- To show relationships between data

When to use a graph



When to use a graph



What we will cover

- When to use a graph
- **Types of graphs**
- General principles for effective graphs

Pie chart

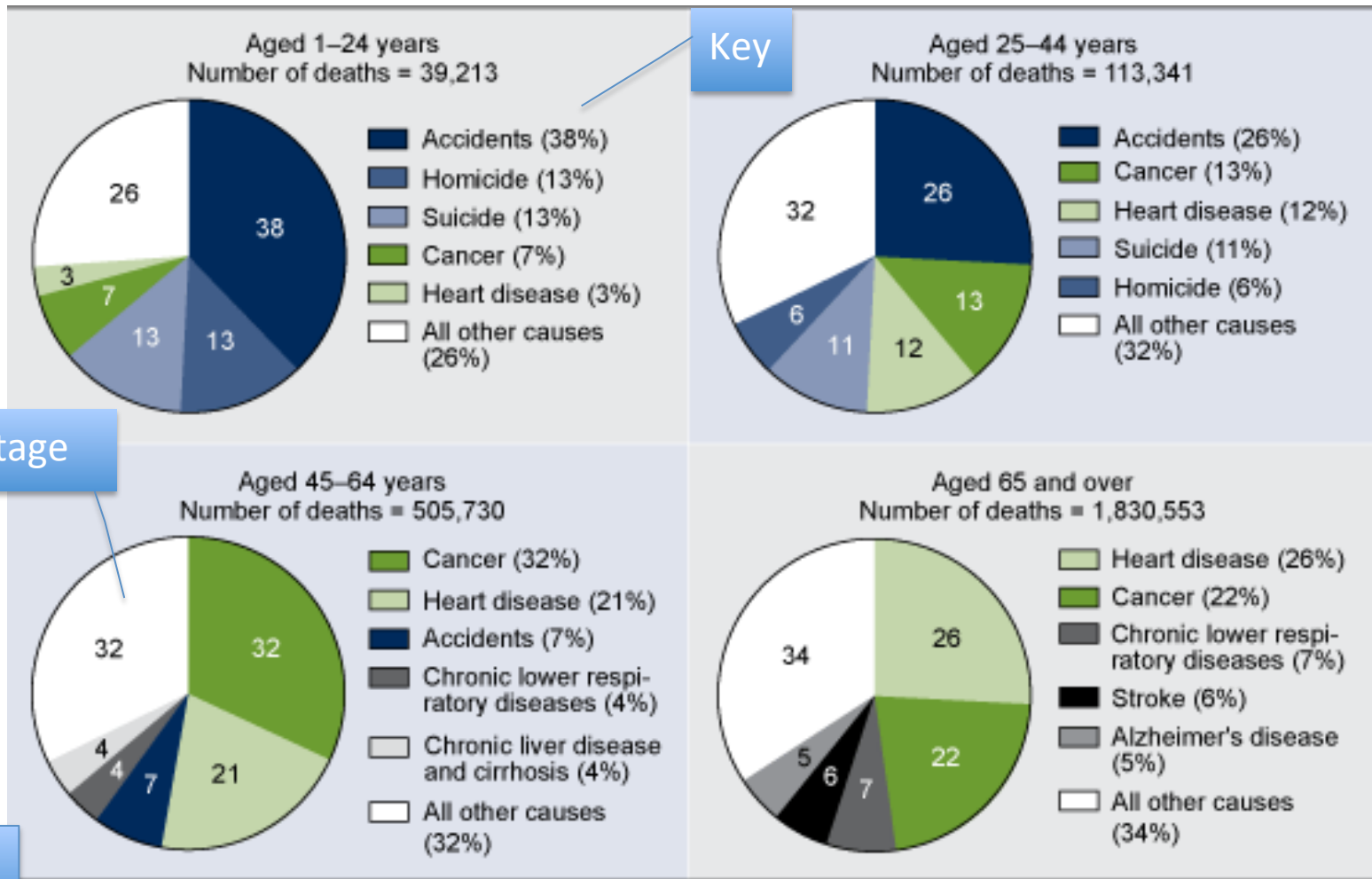
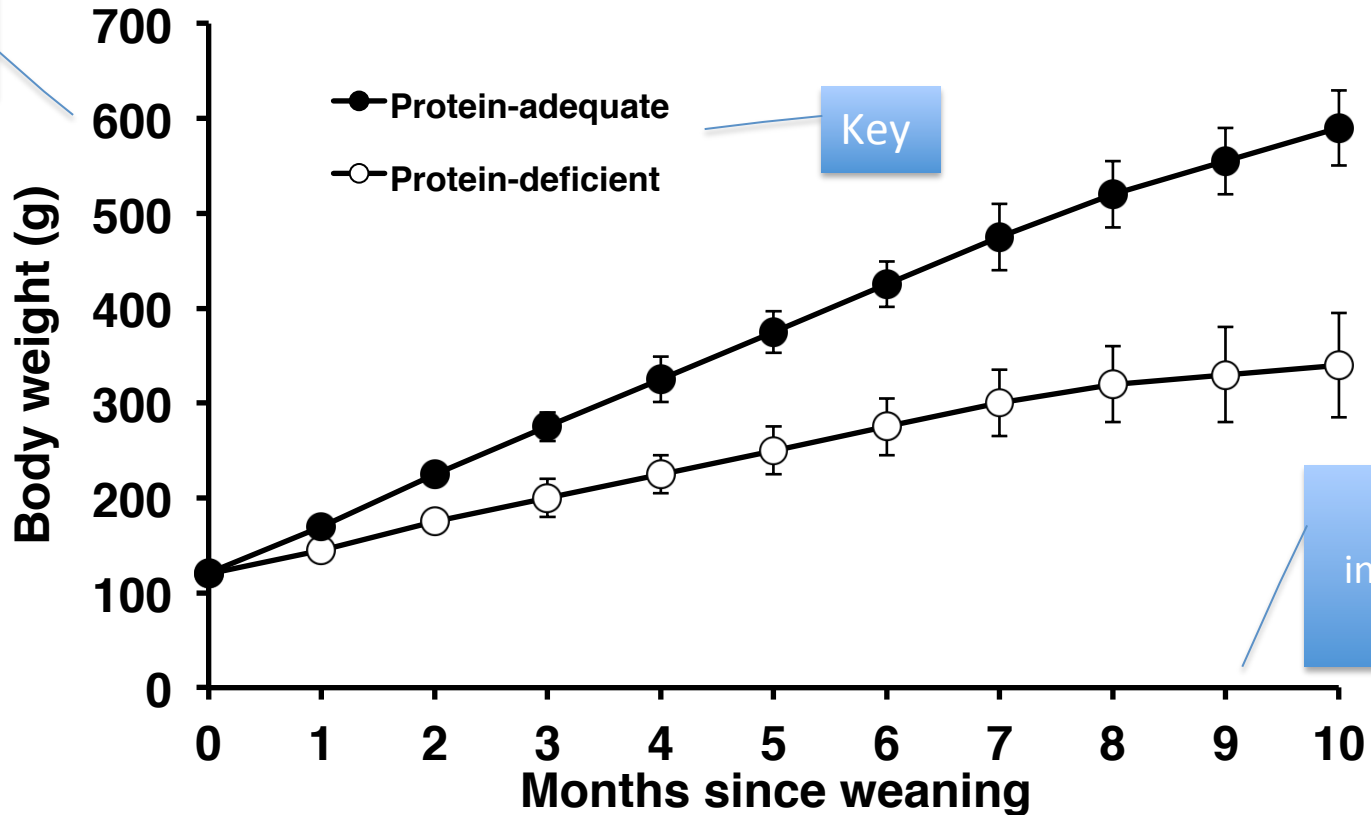


Figure 4. Percentage distribution of five leading causes of death in the U.S. in 2011 by age group.

Line graph

Y-axis:
dependent
variable



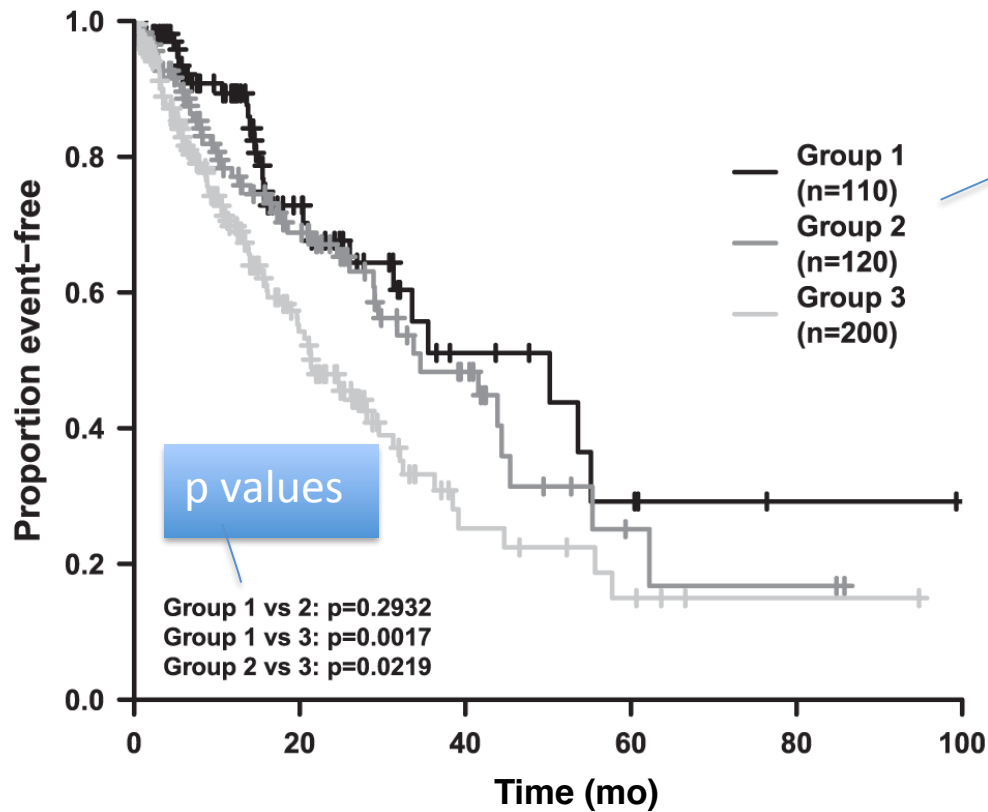
Key

X-axis:
independent
variable

Legend

Figure 5. Mean body weights of weanling guinea pigs fed protein-adequate (n=10) and protein-deficient (n=10) diets for 10 months. Error bars, standard error of the mean.

Line graph (survival curve)



No. at risk

Under risk					
Group 1	30	9	4	1	0
Group 2	45	16	3	2	0
Group 3	53	9	4	1	0

Figure 6.
 Sample
 Kaplan-Meier
 survival plot for
 three patient
 groups.

Scatterplot

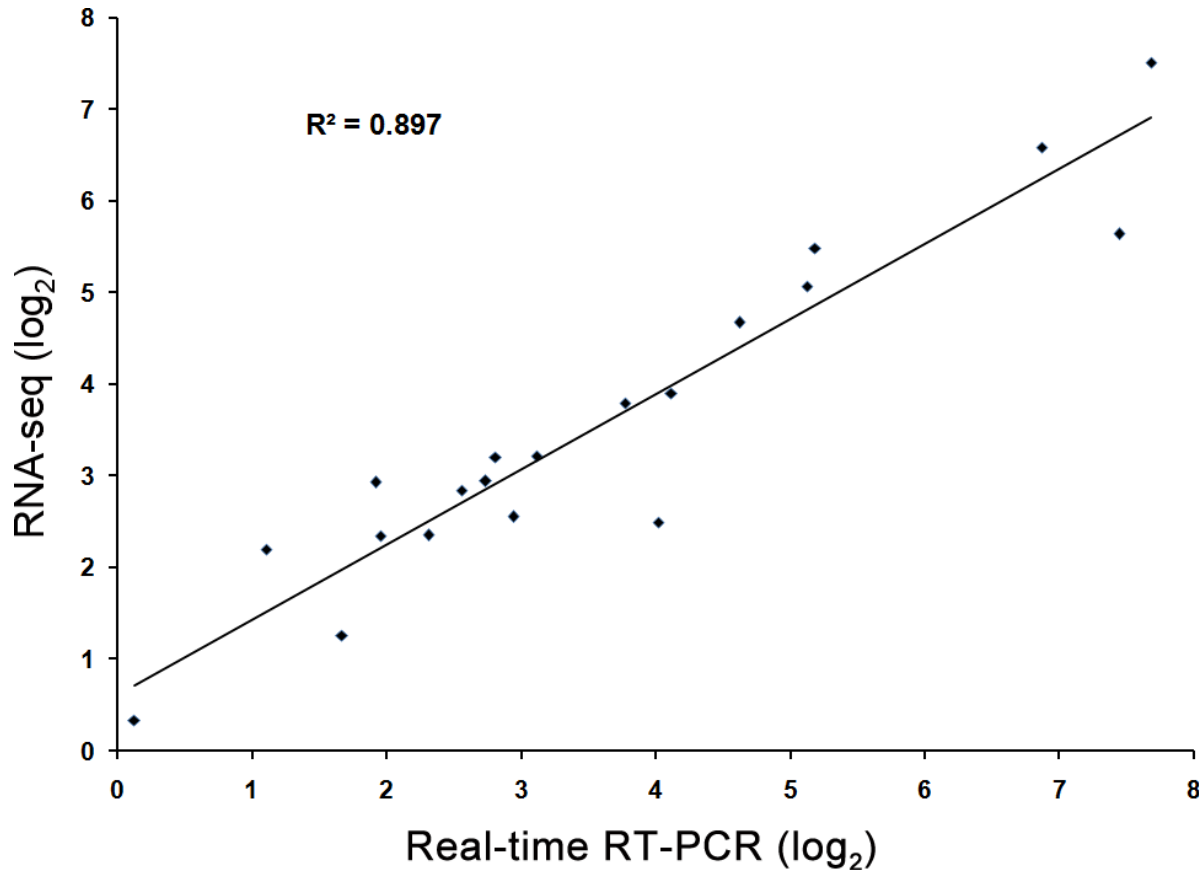


Figure 7. Validation of RNA-seq expression data via quantitative real-time RT-PCR for 20 genes. Average RNA-seq values were plotted against average RT-PCR values and fit into a linear regression.

Bar graph

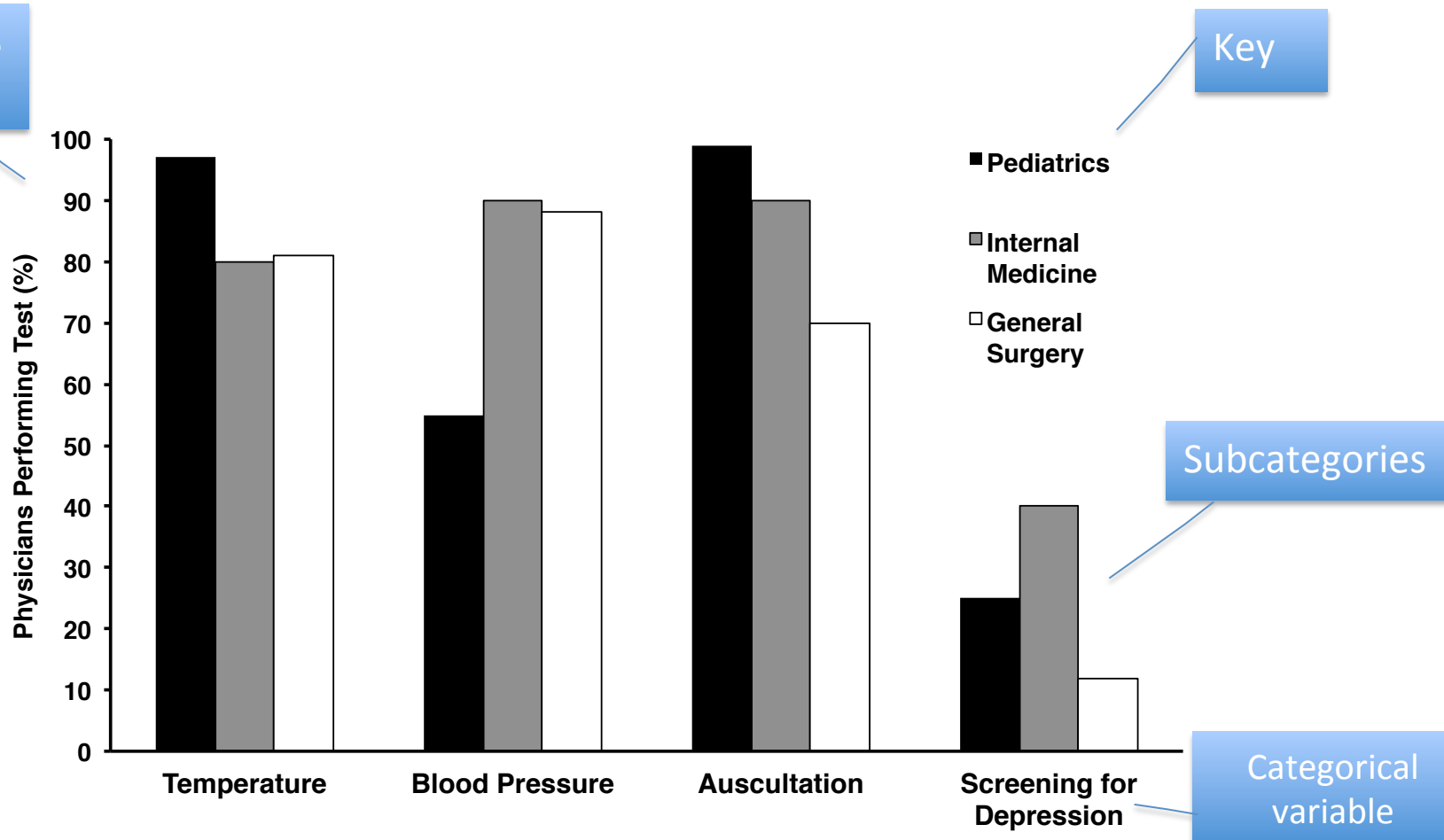
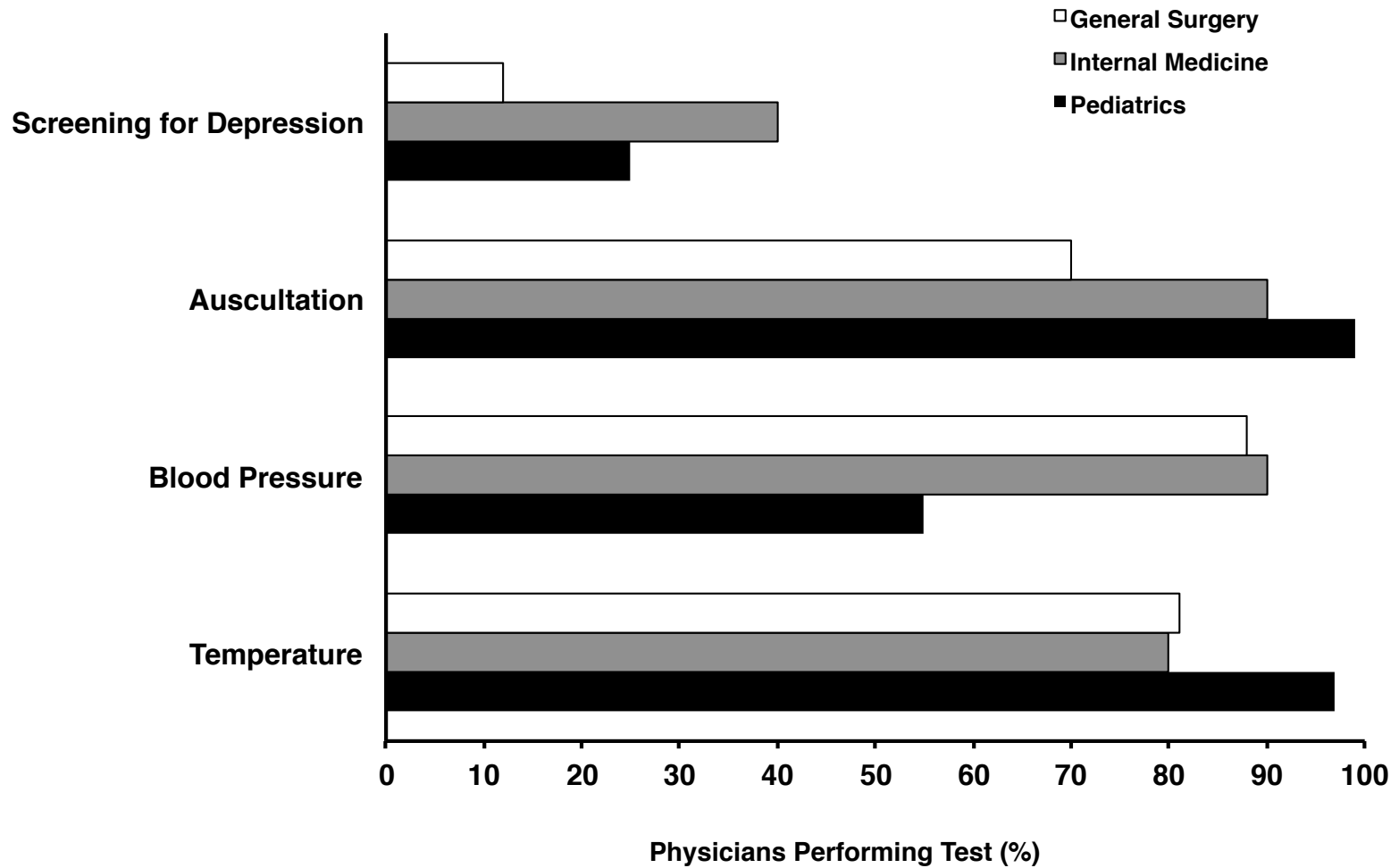
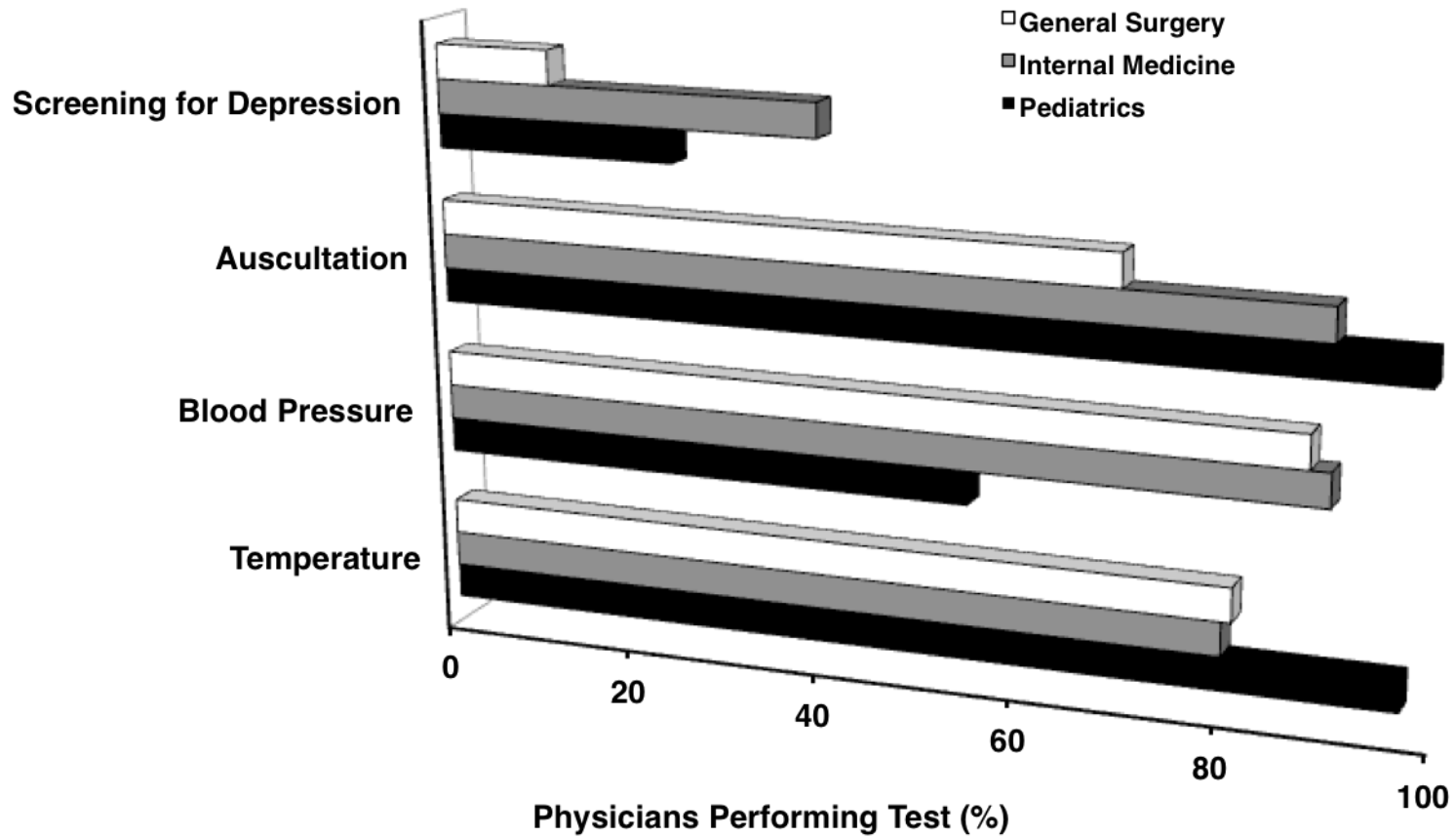


Figure 8. Variations by physician specialty in performance of tests during typical office visits.

Horizontal bar graph



3-D bar graph (not recommended)



Stacked bar graph

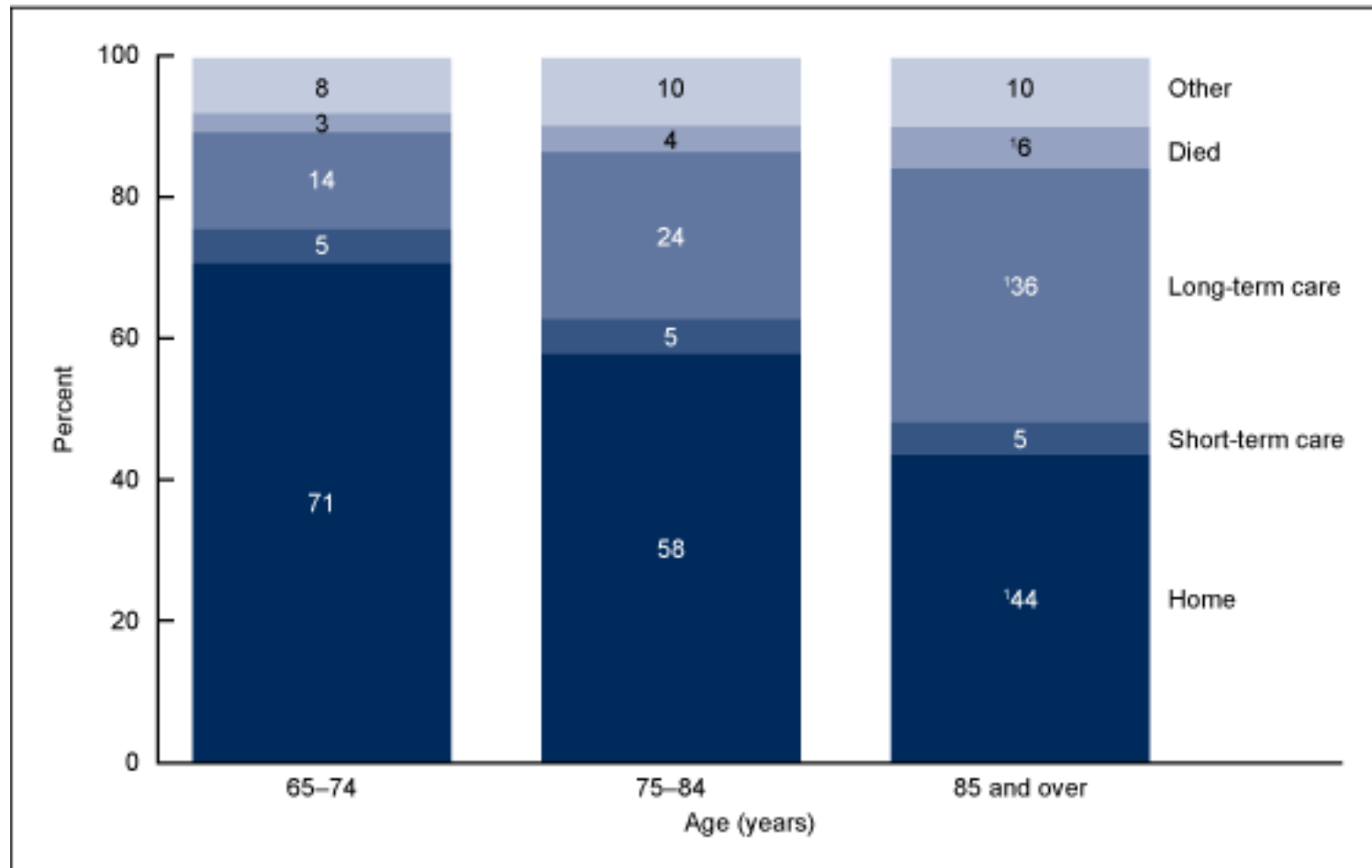


Figure 11. Discharge status, by age group: United States, 2010.

Dot plot

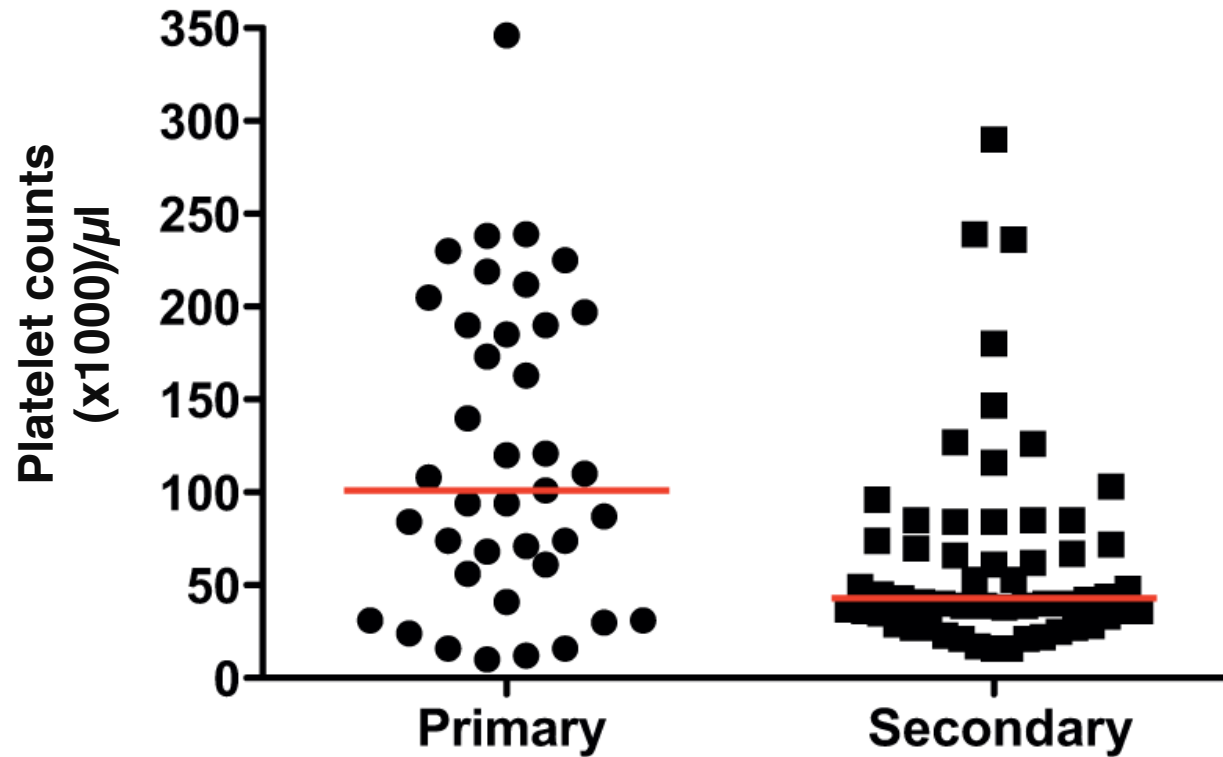


Figure 12. Platelet counts for patients with primary or secondary infections.

Box-and-whiskers plot

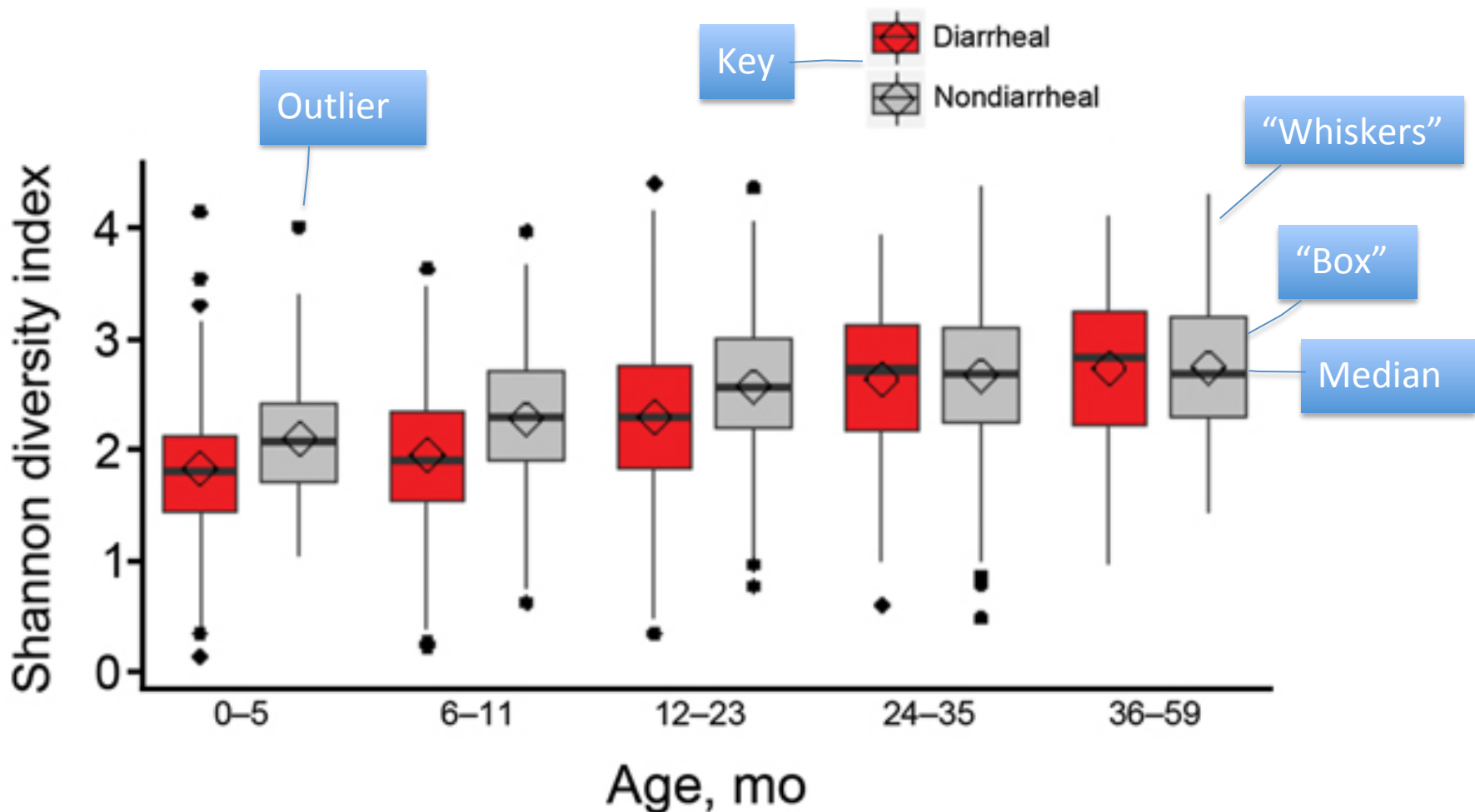


Figure 13. Shannon diversity index for diarrheal and nondiarrheal samples, stratified by age group. ... The upper whisker extends from the 75th percentile to ... The lower whisker extends from ...

Frequency histogram

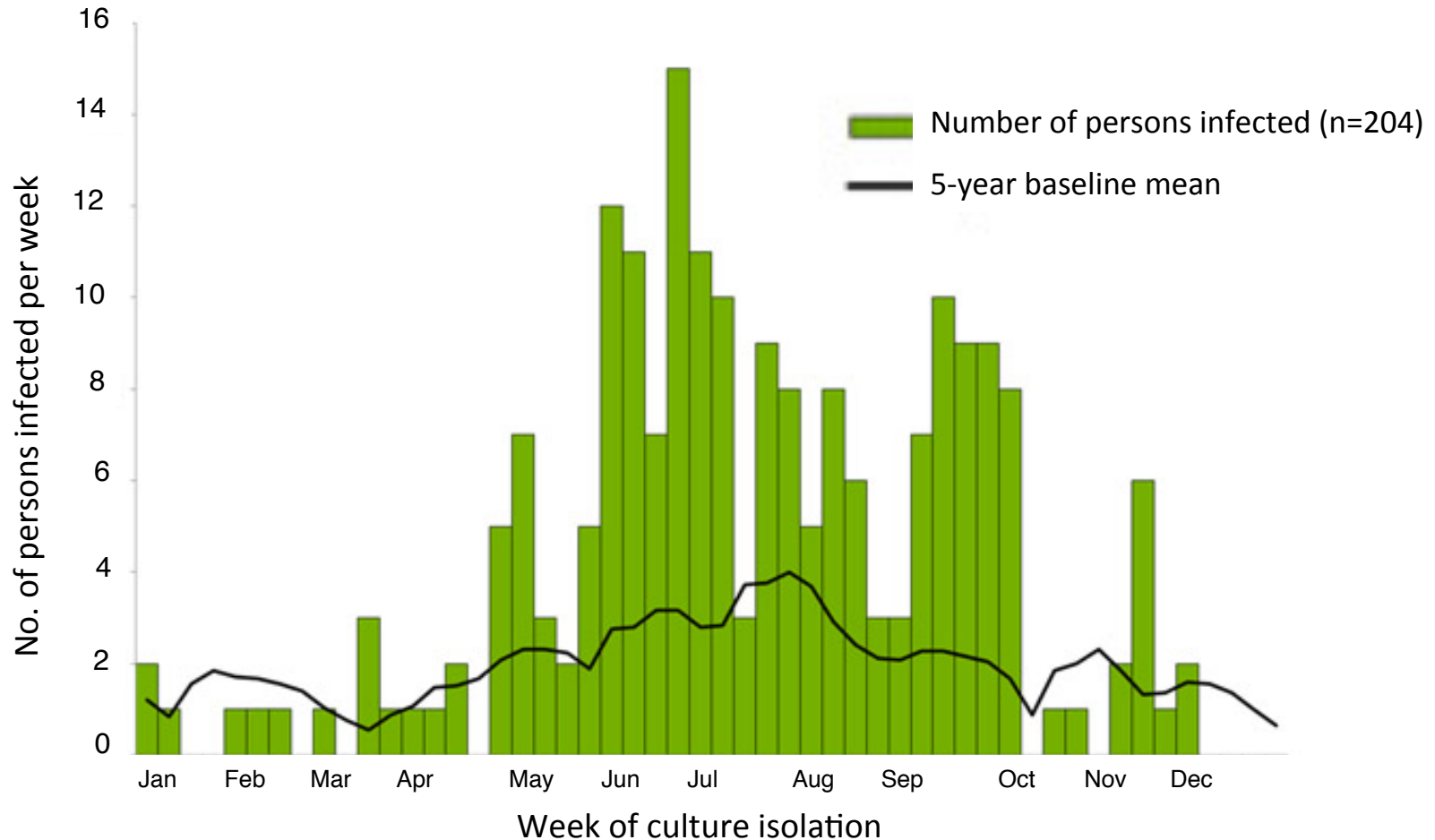


Figure 14. Persons infected with the outbreak strain of Salmonella Heidelberg in six states, 2011.

What we will cover

- When to use a graph
- Types of graphs
- **General principles for effective graphs**

General principles for effective graphs

- Design for clarity
 - Arrange components logically
 - Label clearly
 - Make data easily distinguishable
- Avoid clutter
- Use color wisely
- Include an effective legend
- Be consistent

Label clearly

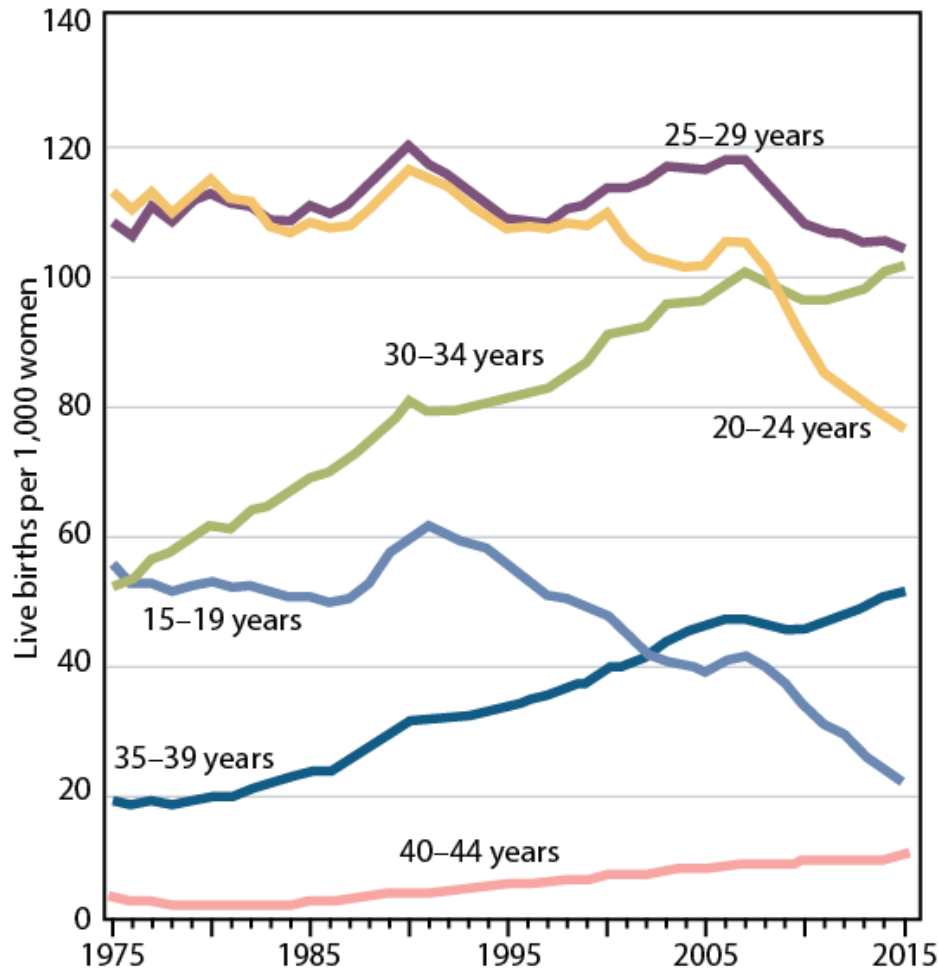


Figure 15. Birth rates, by age of mother: United States, 1975–2015.

Label clearly

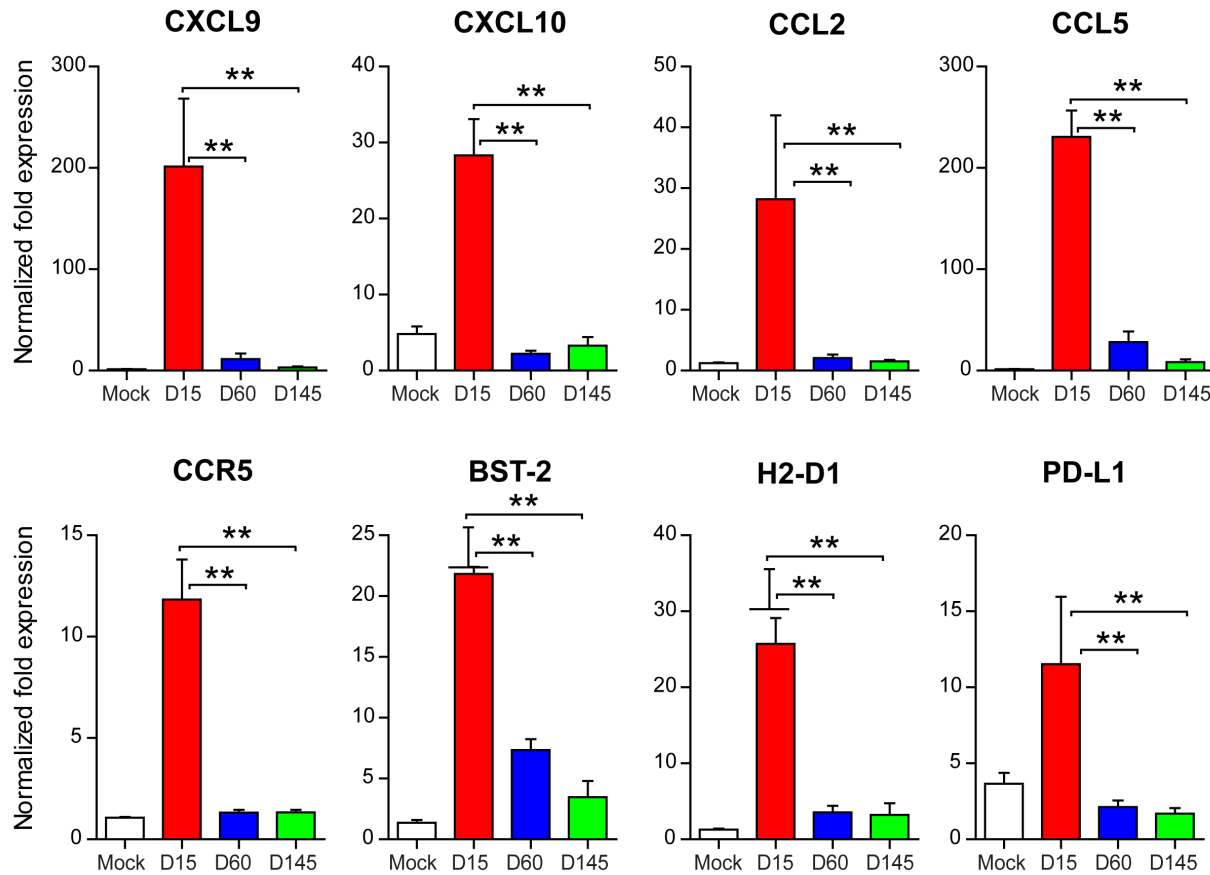


Figure 16. Validation of expression of innate immune genes during viral persistence.

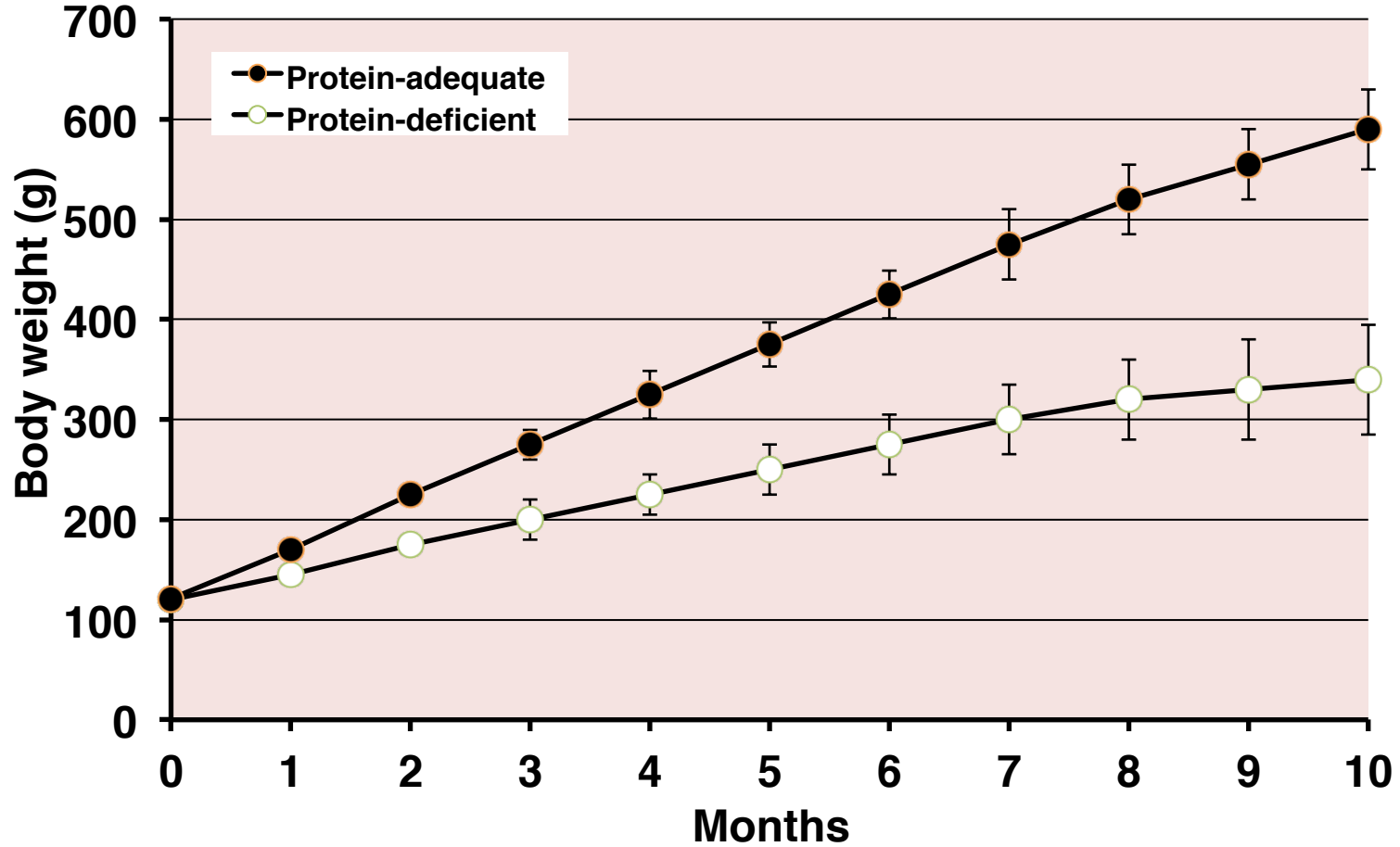
Label clearly

- Make sure labels are large enough to read (no smaller than 8 pt at printed size)
- Use a consistent capitalization style
 - Body weight, g
 - Body Weight (g)
- Indicate units

Make data easily distinguishable

- Data points: open circles, triangles, squares; closed circles, triangles, squares
- Lines
 - A few lines: black solid, dashed, dotted
 - A lot of lines: may need color
- Bars
 - A few bars: black, white, gray (avoid patterned fills)
 - A lot of bars: may need color

Avoid clutter



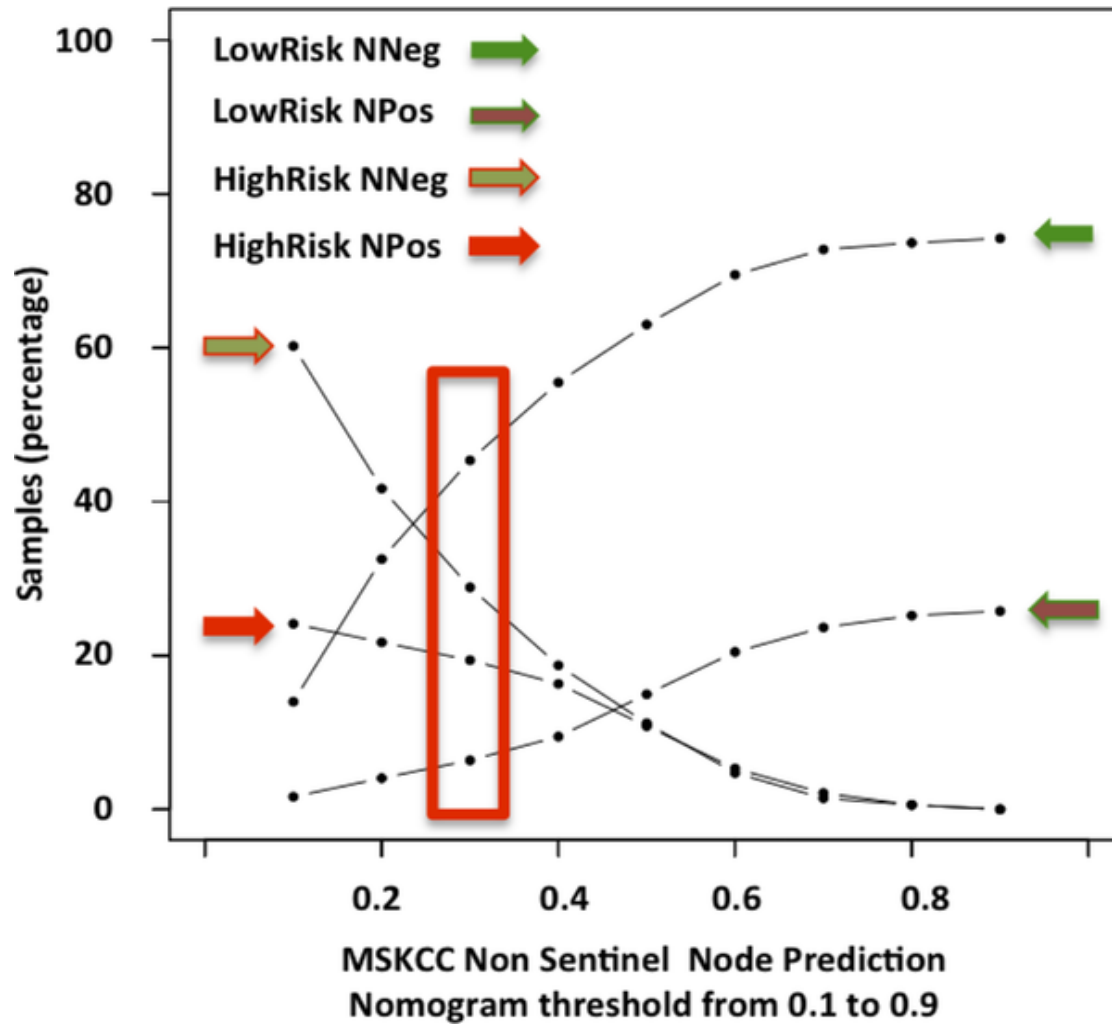
Use color wisely

- *Ann Oncol*: No charge for color figures in online article; **\$600**/figure for color figures in print article
- *Cancer*: No charge for color
- *Cancer Cell*: **\$1,000** for first color figure; **\$275**/figure for others
- *Cancer Res, Clin Cancer Res*: **\$675**/figure for color figures
- *J Clin Oncol*: No charge for color

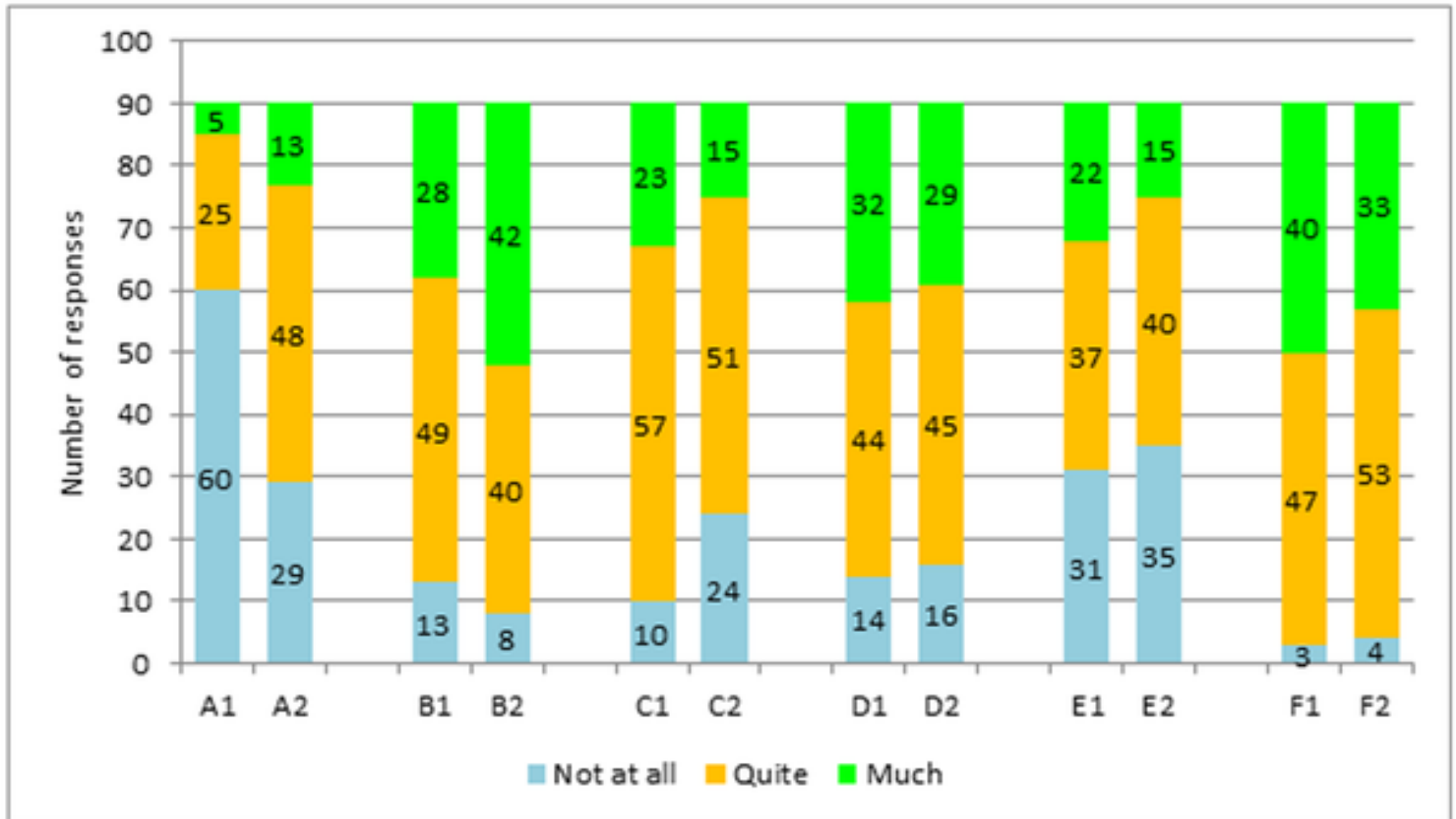
Use color wisely

- Black, red, and blue reproduce well
- Avoid yellow
- Avoid red and green together
- Avoid “loud” colors
- Keep it simple

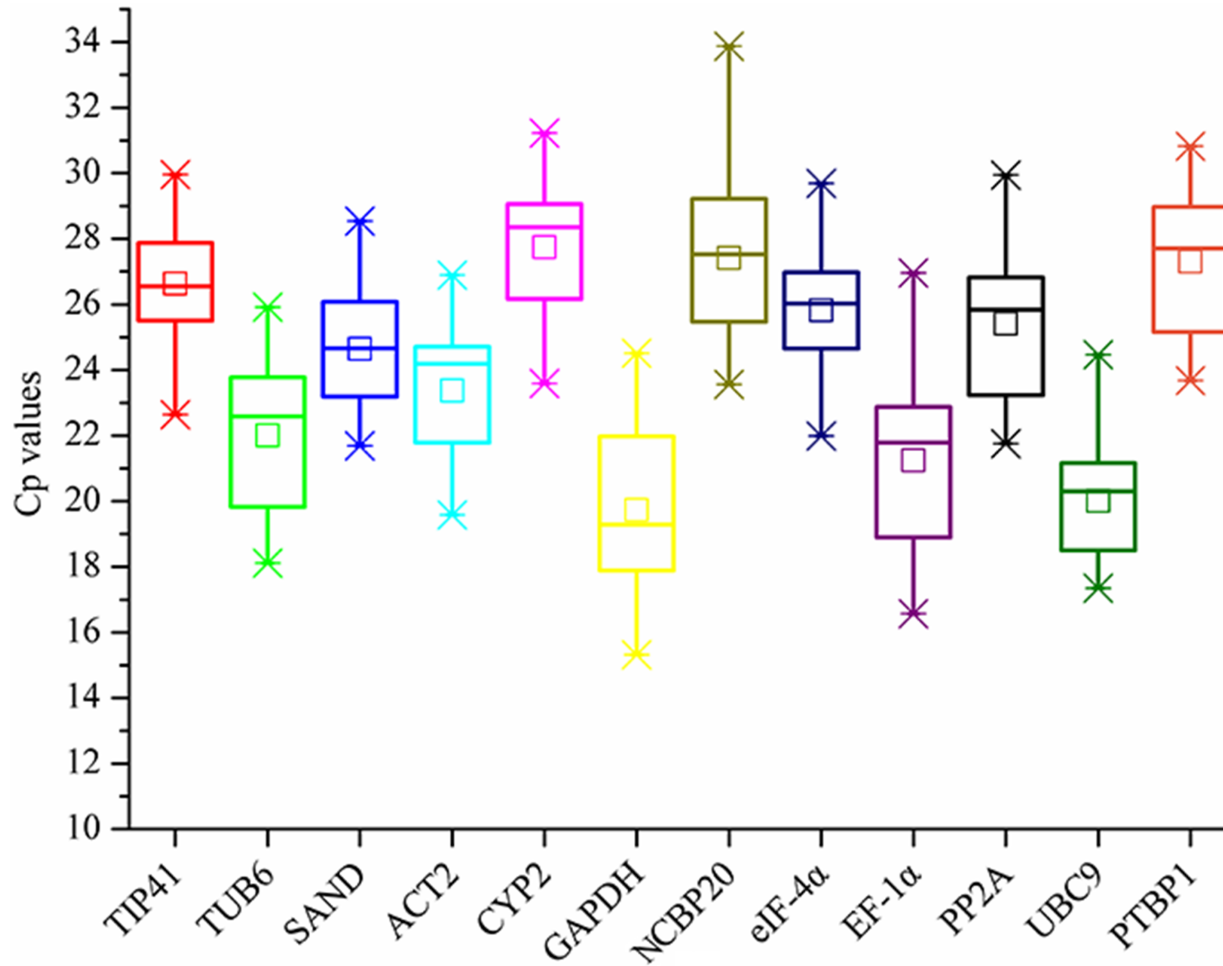
Use color wisely



Use color wisely



Use color wisely



Include an effective legend

- Overall subject or main message of graph
- Number of patients or samples
- Definitions of abbreviations and symbols
- Method
- Statistical test

Include an effective legend

Figure 12. Platelet counts for patients with primary infections (n=39) or secondary infections (n=57). Counts were measured at the time of admission. The median values are indicated by red lines. ** $P < 0.005$, Mann-Whitney test.

Be consistent

- Use the same software
- Use a consistent order of categories
- Use the same font and symbols
- Use consistent nomenclature, spelling, and capitalization

For more information

- American Medical Association. *AMA Manual of Style: A Guide for Authors and Editors*. 10th ed. New York, NY: Oxford University Press; 2007:98-106.
- Lang TA. *How to Write, Publish, & Present in the Health Sciences: A Guide for Clinicians & Laboratory Researchers*. Philadelphia, PA: American College of Physicians Press; 2010:67-86.
- Day RA, Gastel B. *How to Write and Publish a Scientific Paper*. 8th ed. Santa Barbara, CA: Greenwood; 2016:.
- Tufte ER. *The Visual Display of Quantitative Information*. 2nd ed. Cheshire, CN: Graphics Press, 2001.

Graph sources

Line graph (Fig. 3): CDC, <https://www.cdc.gov/mmwr/preview/mmwrhtml/mm6053a1.htm> [public domain]

Pie chart (Fig. 4): Minino AM, <https://www.cdc.gov/nchs/products/databriefs/db115.htm> [public domain]

Survival curve (Fig. 6): Gross et al, <https://doi.org/10.1371/journal.pone.0038960> [CC BY 4.0]

Scatterplot (Fig. 7): Chen et al, <https://doi.org/10.1371/journal.pone.0144927> [CC BY 4.0]

Stacked bar graph (Fig. 11): CDC, <https://www.cdc.gov/nchs/products/databriefs/db182.htm> [public domain]

Dot plot (Fig. 12): Singla et al, <https://doi.org/10.1371/journal.pntd.0004497> [CC BY 4.0]

Graph sources

Box-and-whiskers plot (Fig. 13): Lindsay et al, https://wwwnc.cdc.gov/eid/article/21/2/14-0795_article [public domain]

Histogram (Fig. 14): CDC, <https://www.cdc.gov/salmonella/heidelberg-chickenlivers/011112/epi.html> [public domain]

Line graph (Fig. 15): CDC, <https://www.cdc.gov/nchs/data/hus/hus16.pdf> [public domain]

Bar graph (Fig. 16): Nayak et al, <https://doi.org/10.1371/journal.ppat.1003395> [CC0]

Line graph (Fig. 18): Zhao et al, <https://doi.org/10.1371/journal.pone.0169962> [CC BY 4.0]

Stacked bar chart (Fig. 19): Numico et al, <https://doi.org/10.1371/journal.pone.0101170> [CC BY 4.0]

Box-and-whiskers plot (Fig. 20): Zhao et al, <https://doi.org/10.1371/journal.pone.0152356> [CC BY 4.0]

Scientific Publications

- <http://inside.mdanderson.org/departments/scipub/index.html>
- Tel: 713-792-3305
- Email:
scientificpublications@mdanderson.org
- Location: Mid-Campus Building, 16th floor