What Makes a Good Graph?

Susan P Duke, MS, MS

Manager, Benefit Risk Evaluation

GlaxoSmithKline

&

Micky Salgado-Gomez

Contributed by Susan Duke (https://www.ctspedia.org/StatGraphHome)

Abstract

- From data entry to advanced analytic methods, we've all seen the inroads computerization has made. A strong exception to this is visual perception, one of the few remaining areas where human brains far outperform computers. Consider: a graph is simply another type of statistical method.
- Have you noticed how your eye is drawn to figures more than words and tables? This very fact can be powerfully used to transparently and succinctly communicate the key points of a clinical study or submission.
- The presentation will highlight best practices for effective graph design and creation, including some examples from the FDA/Industry/Academia Safety Graphics Working Group (https://www.ctspedia.org/StatGraphHome).

Outline

- Why graphs? *After all, the data is already in a table*

- Barriers to graph creation and what we can do about it *Best practices and standard graphs that elegantly answer commonly asked questions*

Contributed by Susan Duke
(https://www.ctspedia.org/StatGraphHome)

Why Graphs?

Is Your Brain Frozen?

- An article in Newsweek confirms what we all experience
- From brain scans - parts of the brain best at decision-making can get overloaded
- Too much information results in poorer decisions

The Science of Making Decisions Newsweek 27 Feb 2011

Contributed by Susan Duke
(https://www.ctspedia.org/StatGraphHome)

Why Graphs?

Event	Drug A (%)	Drug B (%)	RR/RR	Low95%	Upp95%
ARTHRALGIA	3.5	0.5	7.0	1.6	31.5
NAUSEA	10.0	4.5	4.1	2.5	6.9
ANOREXIA	3.5	0.9	3.9	1.2	13.1
HEMATURIA	3.2	0.9	3.6	1.0	12.2
INSOMNIA	6.0	1.9	3.2	1.3	7.5
VOMITING	8.6	2.8	3.1	1.5	6.2
DYSPEPSIA	9.7	3.7	2.6	1.4	4.9
WEIGHT DECREASE	2.1	0.9	2.3	0.6	9.0
PAIN	3.9	1.9	2.1	0.6	5.3
DIARRHEA	20.9	10.6	2.0	1.4	2.9
FATIGUE	3.7	1.9	1.9	0.7	5.1
FLATULENCE	4.6	2.8	1.6	0.7	3.7
DIZZINESS	6.7	4.2	1.6	0.8	3.1
ABDOMINAL PAIN	14.2	9.3	1.5	1.0	2.4
RESPIRATORY DISORDER	2.6	1.9	1.4	0.5	4.0
HEADACHE	8.4	6.5	1.3	0.7	2.3
INJURY	7.0	5.6	1.2	0.7	2.3
GASTROESOPHAGEAL REFLUX	2.8	2.3	1.2	0.4	3.3
BACK PAIN	5.3	4.5	1.2	0.6	2.3
HYPERKALEMIA	2.1	1.9	1.1	0.4	3.4
RASH	2.1	1.9	1.1	0.4	3.4
SINUSITIS	6.5	6.0	1.1	0.6	2.0
INFECTION VIRAL	6.0	5.8	1.1	0.6	2.1
UPPER RESP TRACT INFECTION	15.6	15.3	1.0	0.7	1.5
MYALGIA	2.8	2.8	1.0	0.4	2.6
URINARY TRACT INFECTION	2.8	2.8	1.0	0.4	2.6
COUGHING	6.0	6.0	1.0	0.5	1.9
MELENA	2.6	3.2	0.9	0.3	2.2
RHINITIS	3.9	5.1	0.8	0.4	1.7
BRONCHITIS	2.6	3.7	0.7	0.3	1.8
CHEST PAIN	2.8	4.2	0.7	0.3	1.6
CHRONIC OBSTRUCTIVE AIRWAY	22.0	35.2	0.6	0.5	0.8
DYSPNEA	2.1	6.9	0.3	0.1	0.6

From Ken Koury, Merck

Contributed by Susan Duke
(https://www.ctspedia.org/StatGraphHome)

Motivation

- **Tables:** most presentations of safety data convey information in tables
- often difficult to understand, especially when data are complex
- volume and complexity of presentations complicate regulatory review
- may obscure important relationships
- **Graphical displays:** more effective and more efficient
- communicate critical safety information
- facilitate rapid integration and dissemination of new information

From Ken Koury, Merck

Contributed by Susan Duke
(https://www.ctspedia.org/StatGraphHome)

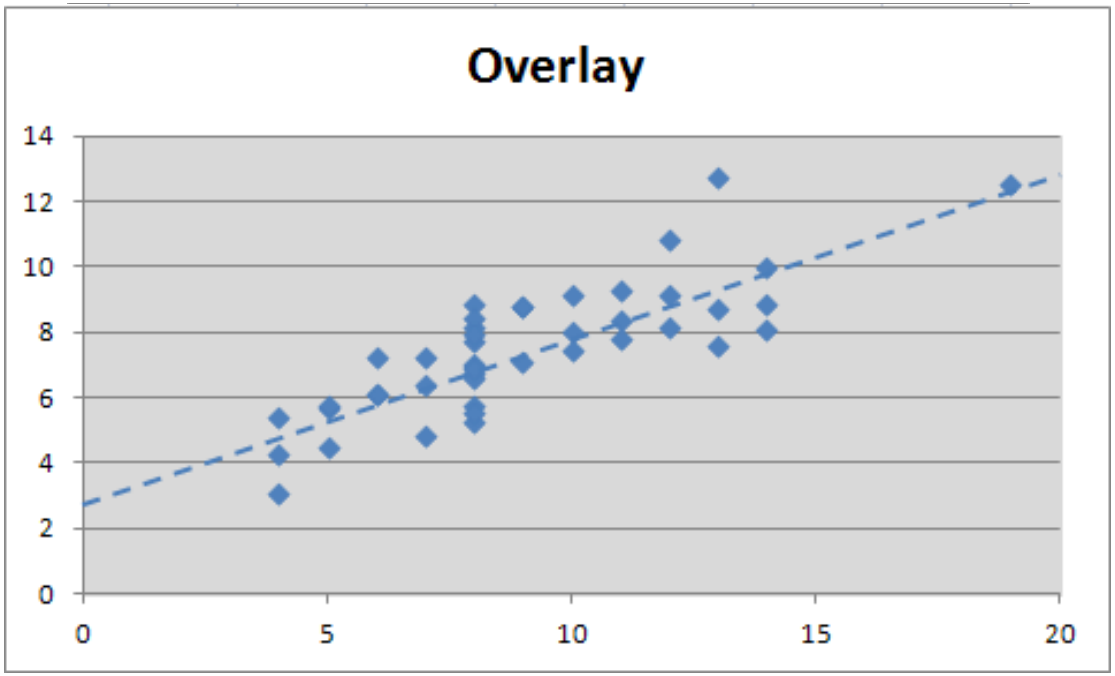
Interpretation – ‘Graphics Reveal Data’ (Tufte)

ANSCOMBE’s QUARTET							
Graph 1		Graph 2		Graph 3		Graph 4	
X	Y	X	Y	X	Y	X	Y
10	8.04	10	9.14	10	7.46	8	6.58
8	6.95	8	8.14	8	6.77	8	5.76
13	7.58	13	8.74	13	12.74	8	7.71
9	8.81	9	8.77	9	7.11	8	8.84
11	8.33	11	9.26	11	7.81	8	8.47
14	9.96	14	8.1	14	8.84	8	7.04
6	7.24	6	6.13	6	6.08	8	5.25
4	4.26	4	3.1	4	5.39	19	12.5
12	10.84	12	9.13	12	8.15	8	5.56
7	4.82	7	7.26	7	6.42	8	7.91
5	5.68	5	4.47	5	5.73	8	6.89

Mean of Y's	7.5
Mean of X's	9
Regression line	Y = 0.5X + 3

From Robert Gordon, J&J

Contributed by Susan Duke
(https://www.ctspedia.org/StatGraphHome)



From Robert Gordon, J&J

Contributed by Susan Duke
(https://www.ctspedia.org/StatGraphHome)

Graphical Perception

"When a graph is constructed, information is *encoded*. The *visual decoding* of this encoded information is *graphical perception*.

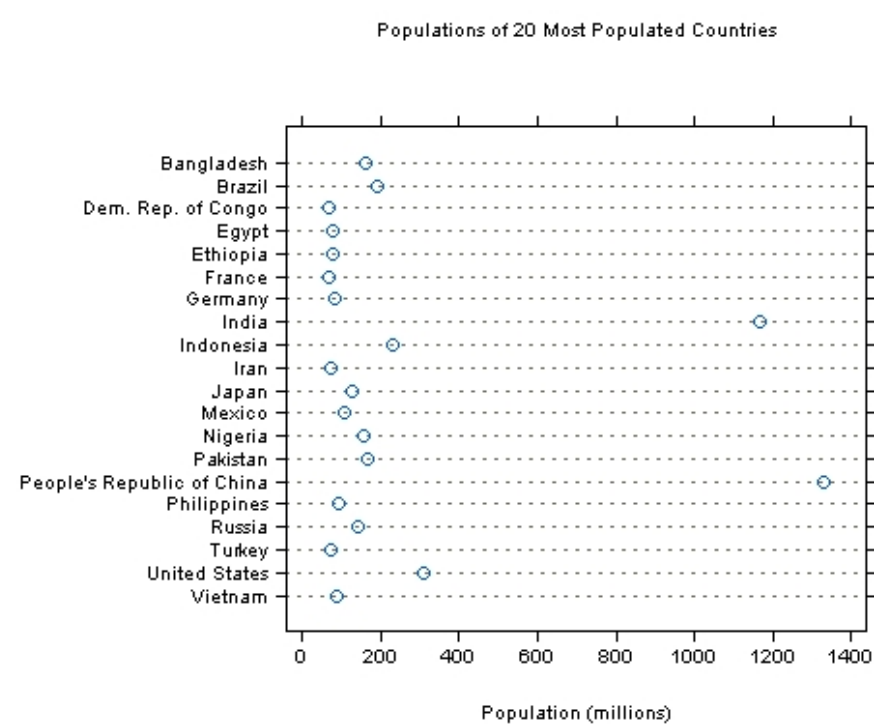
The decoding is the vital link …

No matter how ingenious the encoding … and no matter how technologically impressive the production, a graph is a failure if the visual decoding fails."

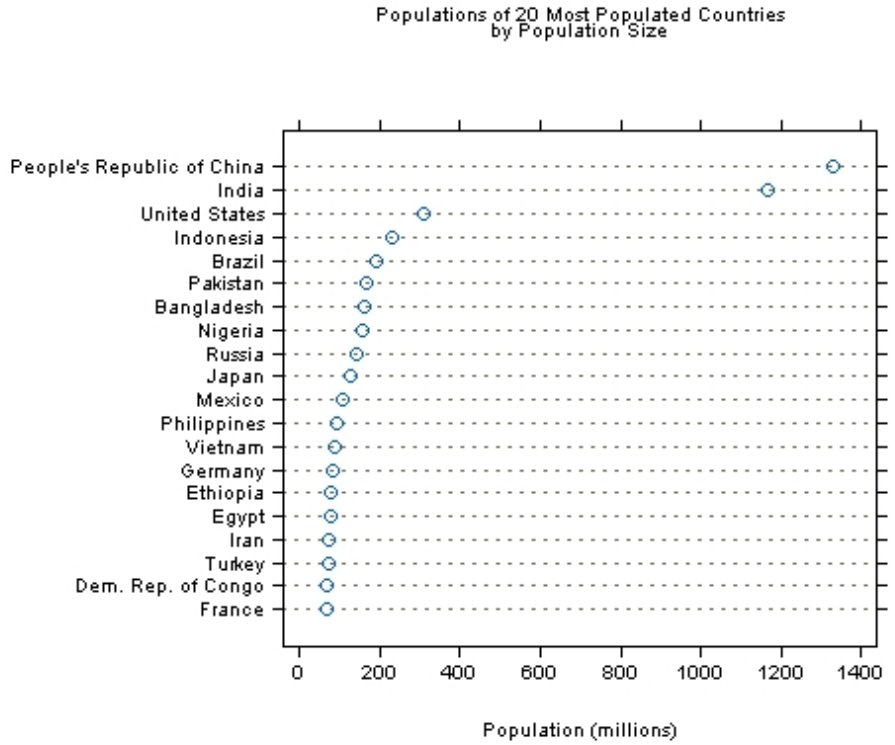
William Cleveland, The Elements of Graphing Data

Contributed by Susan Duke
(https://www.ctspedia.org/StatGraphHome)

Table Look-Up and Pattern Perception



Source: Wikipedia



Concept from William Cleveland, The Elements of Graphing Data
Graphs by Susan Duke, GSK

Contributed by Susan Duke
(https://www.ctspedia.org/StatGraphHome)

Graphical Excellence

- Graphical excellence is the well-designed presentation of interesting data – a matter of substance, of statistics, and of design.
- Graphical excellence consists of complex ideas communicated with clarity, precision, and efficiency.
- Graphical excellence is that which gives to the viewer the greatest number of ideas in the shortest time with the least ink in the smallest space.
- Graphical excellence is nearly always multivariate.
- And graphical excellence requires telling the truth about the data.

From Edward Tufte, 2001. The visual display of quantitative information

Contributed by Susan Duke

Barriers to Graph Creation

- What's in it for me/my group?
  - The value of clear communications inherent in graphs is typically downstream from the group who creates them

- Graphs take time!
  - **Thought time**: how best to design the graph for its intended purpose?
    - What is the purpose of this graph? Who is the audience? What type of data?
  - **Programming time**
    - How to reduce the time needed for programming?

Contributed by Susan Duke
(https://www.ctspedia.org/StatGraphHome)

9 Best Practices for Making Graphs

Best Practices Recommendations

Contributed by: SusanDuke

1. **Content** Every graph should stand on its own
    1. It should tell its story without a need for detailed explanatory text or supporting documents.
    2. It should be clear, effective and informative for the intended audience.
2. **Communication** Tailor each graph to its primary communication purpose
    1. What is insight the graph is intended to convey? Is it intuitive?
    2. Avoid packing too much information into a single display and distracting from the main message.
3. **Information** Maximize the data-to-ink ratio
    1. Each spot of ink should be necessary for imparting the main message
    2. Do not clutter a graph with what you don't need. Less is more.
4. **Annotation** Provide legible text and information
    1. Position annotation (including legends) so that it aids interpretation and does not distract from the message.
    2. Use legible font that can be read without eye strain or a great deal of effort. Consider the format (presentation or document)

Contributed by Safety Graphics Working Group General Principles subteam, lead by Susan Duke, GSK

9 Best Practices for Making Graphs

5. **Axes** Design axes to aid interpretation of a graph
    1. Scale axes to show the interesting features of the data; for example, for longitudinal data, use time (on a continuous scale) instead of visit number (on an ordinal scale).
    2. Give careful consideration to inclusion of the zero of each axis; if excluded, ensure its absence is clearly sign-posted.
    3. Avoid crowded axes.
    4. Use the same axis scales on graphs that need to be compared.
    5. Choose the appropriate style of axes. For example, select between a box, X and Y axes, X only, Y only; consider grid lines; ensure intelligent placing of tick marks.
    6. If the nature of the data suggests the shape of the graphics, follow that suggestion; otherwise, use horizontal graphics about 50% wider than tall.
6. **Styles** Make symbols and plot lines distinct and readable
    1. Choose plot symbols with simple, familiar shapes and intuitive interpretation (eg 'A' for active and 'P' for placebo)
    2. If a graph is to be displayed by projection onto a screen, or in a poster, use thick lines, large symbols and large fonts to achieve legible display.
    3. Where possible and appropriate, data representations (such as styles of symbols, lines and bars) should have the same meaning across all similar graphs within a package; for example, if one line graph uses a solid blue line to represent Placebo, all graphs in the package should use a solid blue line for Placebo.

Contributed by Safety Graphics Working Group General Principles subteam, lead by Susan Duke, GSK

9 Best Practices for Making Graphs

7. **Colors** Make use of color if appropriate for the medium of communication
    1. Use color only when it decodes information. When color is used, choose contrasting and clearly visible colors; avoid yellow, and contrasts with red, green or brown which are difficult for people with color-deficient vision.
    2. If a graph may be viewed in black and white, ensure that all distinctions made by color are also made by other features such as symbols and line-styles.
    3. For black-and-white media, make use of line-styles (dashing and gray levels) that are easy to distinguish.
    4. Design backgrounds to set off the graph, not compete with it.
    5. Choose area fills that are distinct but compatible.
    6. Make secondary plot lines lighter in weight, color or style.
    7. Keep reference lines and grids distinct from other data lines.
    8. Color Brewer is an excellent reference for choice of colors.
8. **Techniques** Use established techniques to clarify the message
    1. Show causality: when a causal relationship exists between variables make sure it is easily discernable from the graph.
    2. Make comparisons from a common baseline.
    3. Sort categories according to relevant features of the data.
    4. Do not introduce spurious dimensions to a graph, as they reduce clarity.
    5. Combine multiple images into a single display when information needs to be presented together.
    6. When a graph summarizes data at an aggregate level, always plot estimates of variability in the data.

Contributed by Safety Graphics Working Group General Principles subteam, lead by Susan Duke, GSK

9 Best Practices for Making Graphs

9. **Types of plots** Use the simplest plot that is appropriate for the information to be displayed (see Select the Right Graph for My Question)
    1. To show a distribution of values, use whichever form is most appropriate: rugplot, strip plot, dotplot, boxplot, histogram, CDF plot, or more specialized display.
    2. Use scatter and line plots to show association between a pair of variables, thinking carefully about the representation of variability of actual data.
    3. Use trellis displays to show changes in association between a pair of variables with respect to a third variable.

Adapted from: GlaxoSmithKline Graphics Principles (used with permission)

Revised by: General Principles subteam, FDA/Industry/Academia Safety Graphics team (24Mar2011)

Contributed by Safety Graphics Working Group General Principles subteam, lead by Susan Duke, GSK

How to reduce programming time?

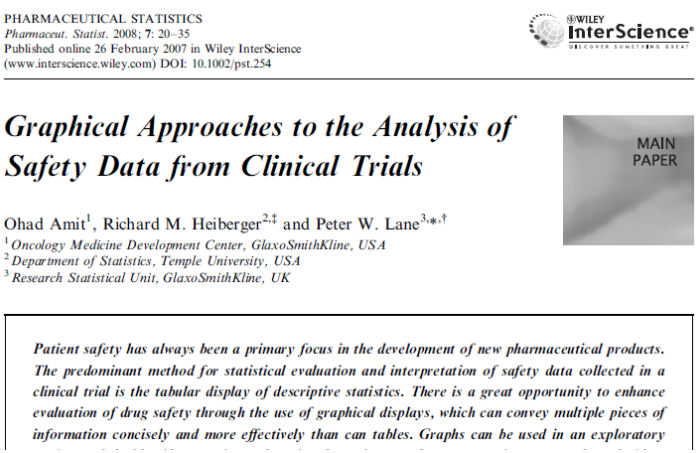
- Let's face it! Time needed to make graphs is typically longer than time needed for tables and listings

- Learnings from GSK Graphics Initiative Experience:
  - Stats & programming teams with a reputation for sustainably creating impactful graphs have a "graphics guru" (or two)
  - Anything to make graph creation easier
    - We focused on safety, specifically standard graphs for common safety questions
    - We use a software tool (TSCG) that has a template for each of the standard graphs. Graphs are created with a GUI interface and easy to put in our reporting system

- Industry-wide, FDA has encouraged a similar approach with the FDA/Industry/Academia Safety Graphics Working Group

How to Make Quality Graphs More Quickly?
Use Standard Graphs for Common Safety Questions

- Two references:
  - Graphical approaches to the analysis of safety data from clinical trials (Amit, Heiberger & Lane, 2008)



PHARMACEUTICAL STATISTICS
Pharmaceut. Statist. 2008; **7**: 20–35
Published online 26 February 2007 in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/pst.254

Graphical Approaches to the Analysis of Safety Data from Clinical Trials

MAIN PAPER

Ohad Amit[1], Richard M. Heiberger[2,‡] and Peter W. Lane[3,*,†]
[1] *Oncology Medicine Development Center, GlaxoSmithKline, USA*
[2] *Department of Statistics, Temple University, USA*
[3] *Research Statistical Unit, GlaxoSmithKline, UK*

Patient safety has always been a primary focus in the development of new pharmaceutical products. The predominant method for statistical evaluation and interpretation of safety data collected in a clinical trial is the tabular display of descriptive statistics. There is a great opportunity to enhance evaluation of drug safety through the use of graphical displays, which can convey multiple pieces of information concisely and more effectively than can tables. Graphs can be used in an exploratory

  - FDA/Industry/Academia Safety Graphics Working Group
    - Each graph entry in the wiki has a description of use, sample program code & data
    - The wiki is searchable, has a glossary and many other features
    - ctspedia.org/StatGraphHomeLink to Safety Graphics

Graphics Glossary

Glossary of Graphics Information

The following is a list of terms used in the graphing of data.

A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U, V, W, X, Y, Z

A

- **Aspect ratio**: The aspect ratio of a graph is the height of the data rectangle divided by the width. An aspect ratio of 1 means the data rectangle is square with the height and width being the same length.

B

- **Bar chart**: A bar chart is a chart with rectandular bars with lengths proportional to the values that they represent. The bars can be plotted vertically or horizontally. Bar charts are used for plotting *discrete* data.
- **Boxplot**: A boxplot is a graph of the statistical five number summary. A central box spans the the first (25th percentile) and third (75th percentile) **quantiles**. A line bisects the box at the median (50th percentile). Lines extend from the box to the smallest and largest observations. An alternate version of the box plot, the **modified boxplot**, has lines that extend from the box to 1.5 **IQR** (see below).

C

Contributed by Peter Lane, GSK (retired).

How to Make Quality Graphs More Quickly?

Use Standard Graphs for Common Safety Questions

You are here: CTSPedia > CTSpedia Web > StatGraphHome (23 Aug 2012, MaryBanach) Edit Attach

Tags: create new tag , view all tags, tagging instructions

Welcome to the Clinical Trials Safety Graphics Home Page

Graphs that answer common clinical trial safety questions

Recommendations from the FDA/Industry/Academia Safety Graphics Working Group

- Labs / Liver Toxicity
- General Adverse Events
- ECG/Vital Signs

for general information about graph types and where to use them

Select the Right Graph for Your Data

See all graphical entries in the library

Search for a graph entry

Resources:

- 9 Best Practices for Making Graphs
- Graphics Glossary
- FDA/Industry/Academia Safety Graphics Presentation Archive
- Graphics References

CTSpedia

Log In or Register

Webs on CTSpedia

- CTSpedia
- Main
- ResearchEthics
- Sandbox
- System

Recent Changes
Help

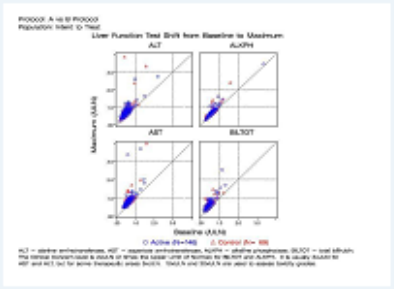
Thanks to Mary Banach, UC Davis, for her webmaster role on the FDA/Industry/Academia Safety Graphics Working Group

Graphs that answer common lab questions

Baseline and Trending over Time

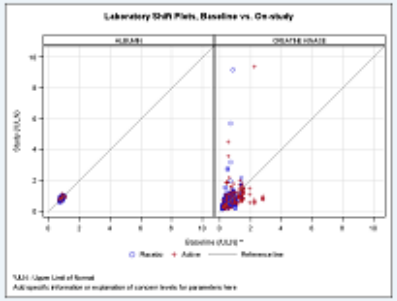
1. What are the changed and percent changes from baseline over time? ie, are abnormal lab values a result of an abnormal baseline or have values changed on study?

Example 1



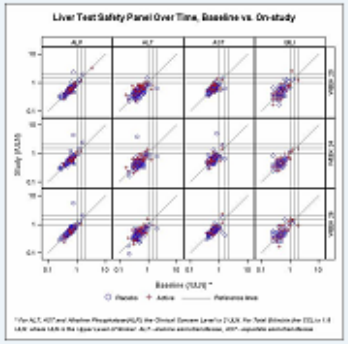
Click here for Example 1 Data

Example 2



Click here for Example 2 Data

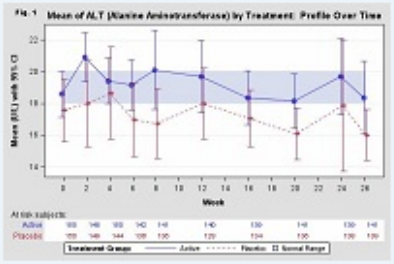
Example 3



Click here for Example 3 Data

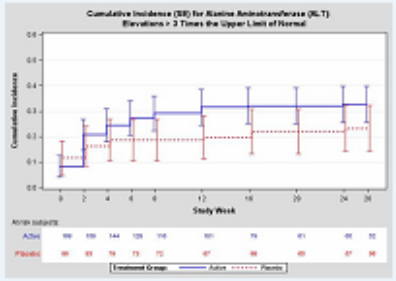
2. Is there a temporal relationship between treatment and lab values?

Example 1



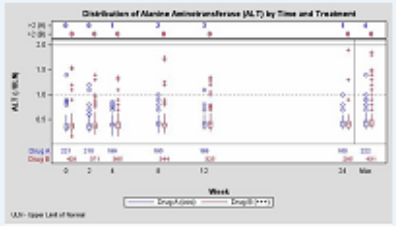
Click here for Example 1 Data

Example 2



Click here for Example 2 Data

Example 3



Click here for Example 3 Data

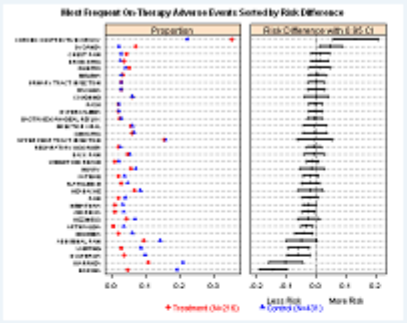
See more questions and graphs at ctspedia.org/StatGraphHome

Contributed by Safety Graphics Working Group Lab/Liver subteam, lead by Robert Gordon, J&J

Graphs that answer common adverse event questions

1. Which AEs are elevated in treatment vs. control? Which AE could be a safety signal?

Example 1

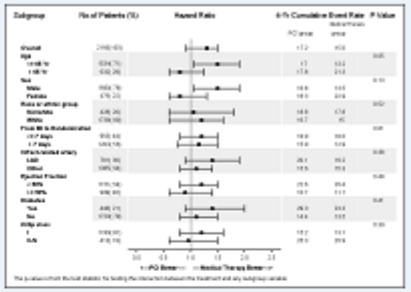


Click here for Example 1 Data

Example 2

Click here for Example 2 Data

Example 3

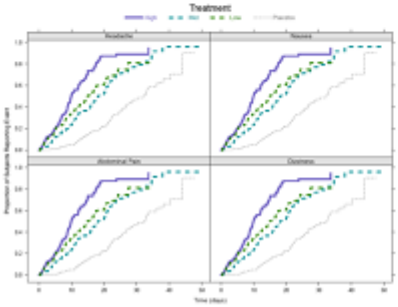


Click here for Example 3 Data

2. What is the risk trend of an Adverse Event of Special Interest?

3. Is there a difference in the time to the first event across treatment groups?

Example 1



Click here for Example 1 Data

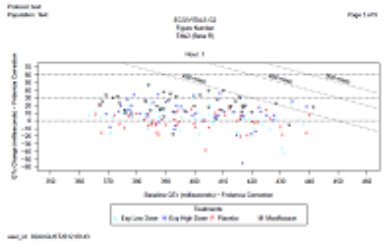
See more questions and graphs at ctspedia.org/StatGraphHome

Contributed by Safety Graphics Working Group AE subteam, lead by Qi Jiang, Amgen & Liping Huang, CSL Behring

Graphs that answer common ECG/Vital Signs questions

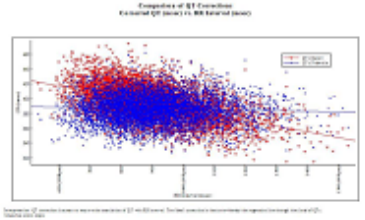
1. What are the longitudinal trends in the data?

2. Are there outlier individuals that have large changes or raw values?



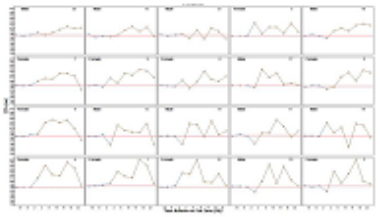
Click here for more information

3. How do the different QT correction compare w.r.t relationship with RR?



Click here for more information

4. How do individuals' values track over time?



Click here for more information

See more questions and graphs at ctspedia.org/StatGraphHome

Contributed by Safety Graphics Working Group ECG/Vitals subteam, lead by Rich Anziano, Pfizer

Solutions to Graph Creation

- What's in it for me/my group?
  - See it as an opportunity for empowerment!
- Graphs take time!
  - **Thought time**: how best to design the graph for its intended purpose?
    - What is the purpose of this graph? Who is the audience? What type of data?
    - **Consider the 9 Best Practices for Making Graphs…**
  - **Programming time**
    - How to reduce the time needed for programming?
    - **Use the Safety Graphics Wiki for ideas and code**
    - **Use software that's designed for graphing (eg, R)**
    - **Work together (many teams have a 'graphics guru')**

Contributed by Susan Duke (https://www.ctspedia.org/StatGraphHome)

Conclusions

- Clear and informative graphs enhance the ability to understand the data
- Suitable graphical presentation could increase the likelihood of detecting safety signals
- Graphs convey information more efficiently and better meet regulatory requirements for ongoing safety evaluation

From Ken Koury, Merck

Contributed by Susan Duke
(https://www.ctspedia.org/StatGraphHome)

Thanks to

- The FDA/Industry/Academia Safety Graphics Team
  - ***Regulatory*: Mat Soukup,** George Rochester, Antonio Paredes, Chuck Cooper, Eric Frimpong, Hao Zhu, Janelle Charles, Jeff Summers, Joyce Korvick, Leslie Kenna, Mark Walderhaug, Pravin Jadjav, Richard Forshee, Robert Fiorentino, Suzanne Demko, Ted Guo, Yaning Wang, Robert Makowsky
  - ***Industry*: Brenda Crowe, Ken Koury,** Andreas Brueckner, Andreas Krause, Fabrice Bancken, Larry Gould, Liping Huang, Mac Gordon, Matthew Gribbin, Navdeep Boparai, Qi Jiang, Rich Anziano, Susan Duke, Sylvia Engelen,
  - ***Academia*:** Frank Harrell, Mary Banach
  - ***Programmers:*** Max Cherny, Sanjay Mantage, Sally Rodriquez
- GSK Graphics Team (variously lead by Ohad Amit, Peter Lane, Susan Duke, Mark Jones, Jay Hilliard and Michael Durante)
- Special thanks to Peter Lane (formerly GSK, now retired)
  - Peter had the original idea to create a Graphics Catalogue at GSK. He drafted the initial best practices, created the glossary and wrote or reviewed many of the GSK graphics entries.
  - Much of Peter's work has found it's way into the Safety Graphics wiki.

Contributed by Susan Duke (https://www.ctspedia.org/StatGraphHome)