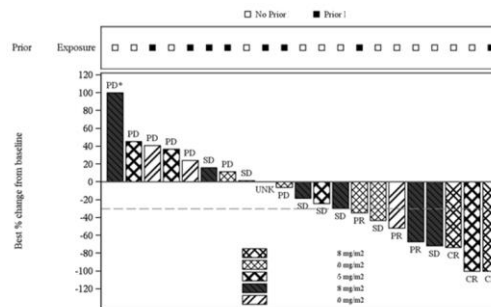




Figure 11-1 (Page 1 of 1)
 Best percentage change from baseline in sum of longest diameters and best overall response as per investigator by prior treatment (Full analysis set)



* Denotes the percentage change from baseline greater than 100.
 Source: Table 11-4, Listing 14.2-1.2 and Listing 16.2.4-1.5

Lessons from a company wide data visualization initiative

Mark Baillie, AMDS Novartis

EFSP/PSDM data visualization workshop, Utrecht 19th April 2023

<https://graphicsprinciples.github.io/>

How can we make better graphs?

During 2014, Novartis formed a cross-functional Graphics Workstream to:

- Enhance understanding and use of **good graphical principles**
- Promote **graphical thinking**
- Facilitate **easy access** to high quality graphics
- Provide **state-of-the-art** ideas on graphical presentation of data for exploratory analyses
- Outline a **view of graphics at Novartis tomorrow**



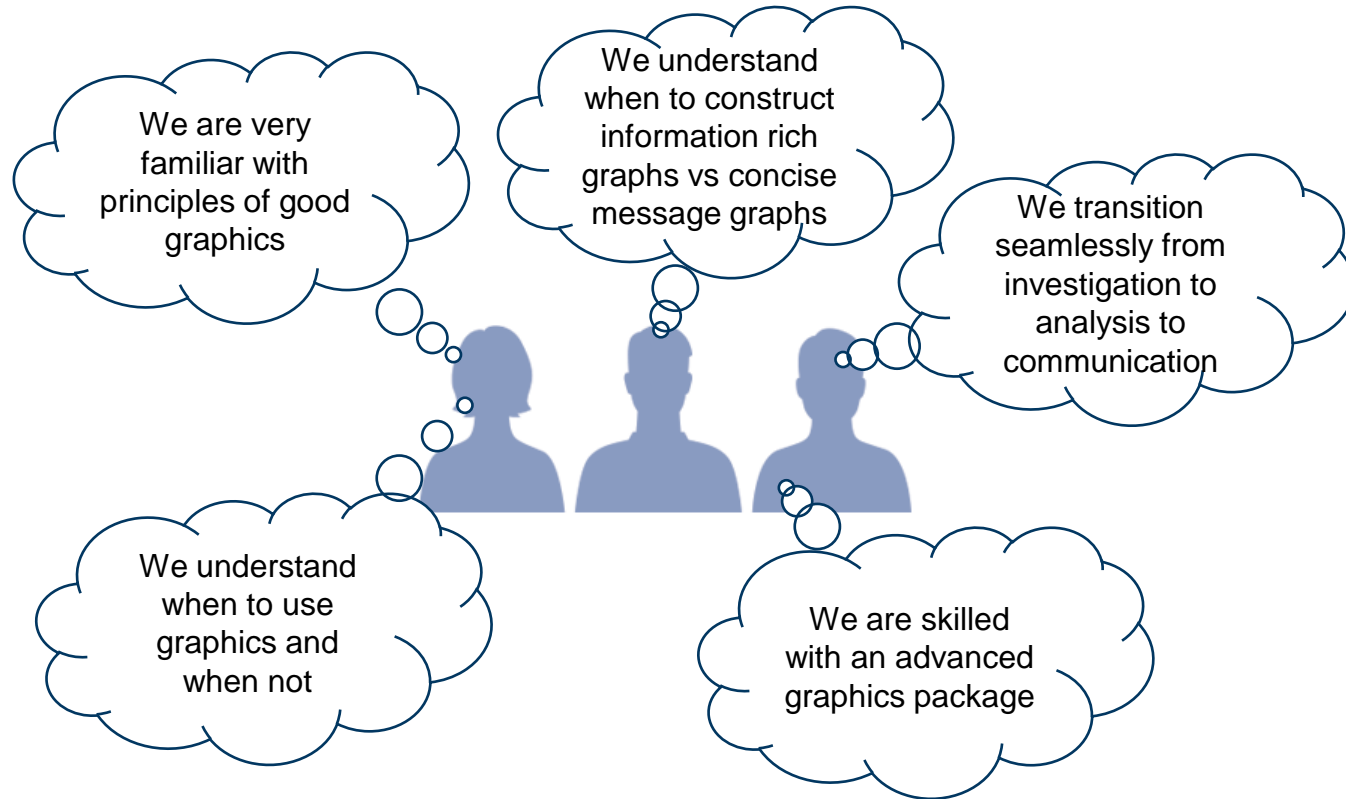
MAIN PAPER | [Full Access](#)

How can we make better graphs? An initiative to increase the graphical expertise and productivity of quantitative scientists

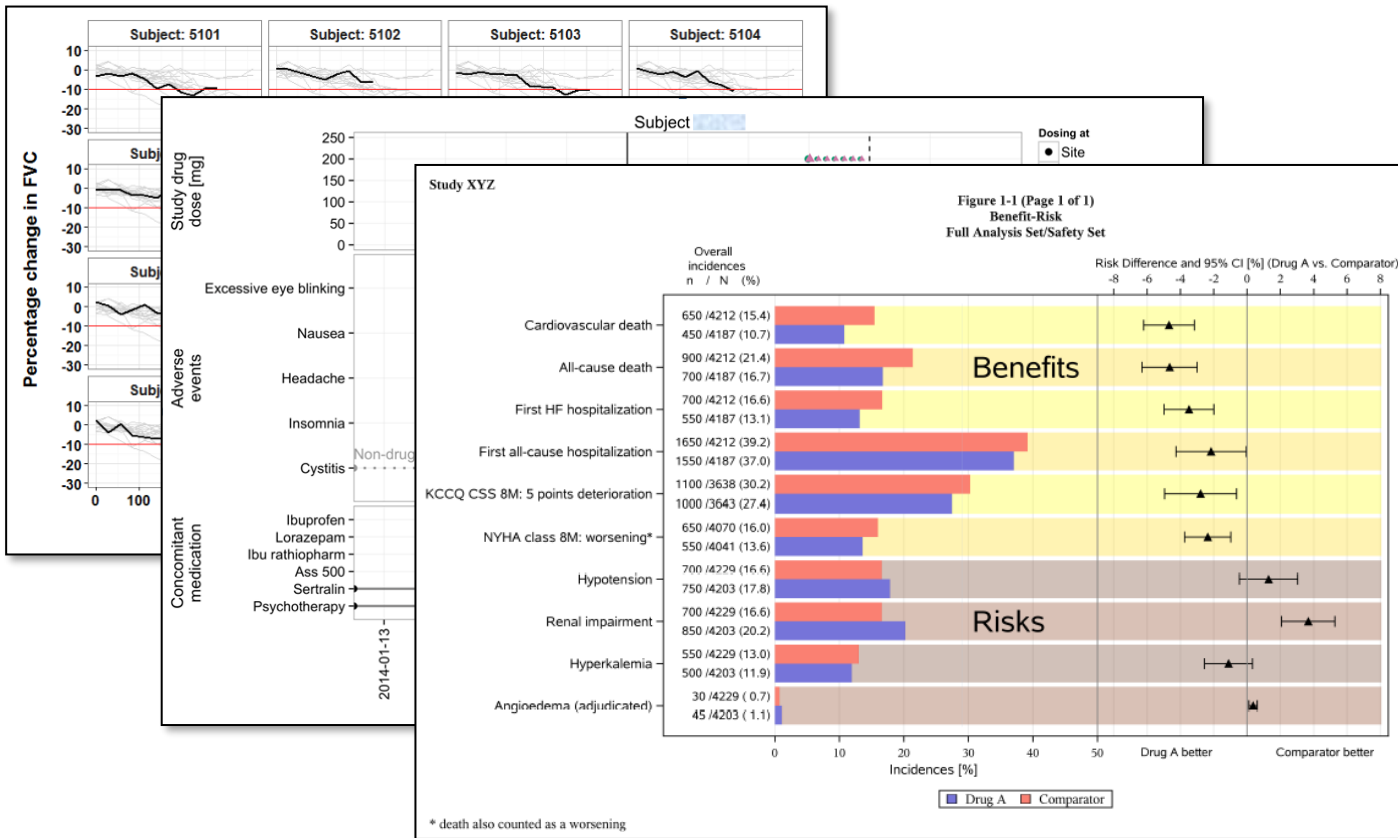
Marc Vandemeulebroecke ✉, Mark Baillie, David Carr, Linda Kanitra, Alison Margolskee, Andrew Wright, Baldur Magnusson

<https://onlinelibrary.wiley.com/doi/full/10.1002/pst.1912>

Vision: Tomorrow's ideal applied statistician



...and more of these



How do we get there?

- How do we tell a good graph from a bad one?
- How do we ensure we design a graph that is fit for purpose?
- Recall workstream goal: enhance understanding and use of **good graphical principles**

Elements of the initiative

Graphical principles and thinking

1. Graphics Principles Cheat Sheet
2. Newsletter

Easing the implementation

3. Graph Gallery
4. Analysis Results Datasets
5. Standardization of most common/important graphs

Graphics tomorrow ... or today?

6. Question-based visualizations and interactive graphics

...plus overarching stakeholder management and communication

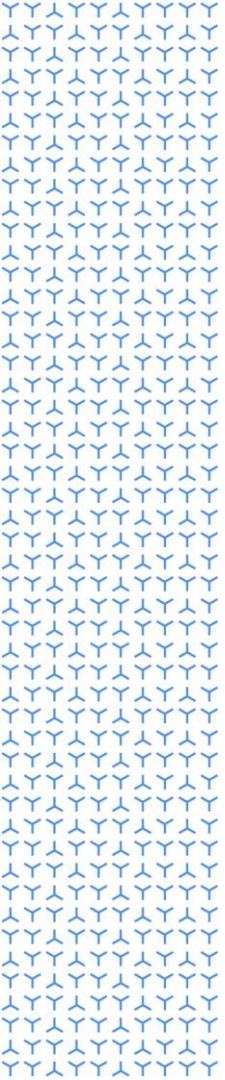


MAIN PAPER |  Full Access

How can we make better graphs? An initiative to increase the graphical expertise and productivity of quantitative scientists

Marc Vandemeulebroecke  Mark Baillie, David Carr, Linda Kanitra, Alison Margolskee, Andrew Wright, Baldur Magnusson

<https://onlinelibrary.wiley.com/doi/full/10.1002/pst.1912>



Graphics Principles Cheat Sheet

Graphics Principles Cheat Sheet

Graphics Principles Cheat Sheet v1.1

Communication

Effective visualizations communicate complex statistical and quantitative information facilitating insight, understanding, and decision making.

But what is an effective graph?

This cheat sheet provides general guidance and points to consider.

Planning

Clearly identify the purpose of the graph, e.g. to deliver a message or for exploration?

Identify the quantitative evidence to support the purpose

Identify the intended audience (specialists, non-specialists, both) and focus the design to support their needs

Identify the intended audience (specialists, non-specialists, both) and focus the design to support their needs

Adapt the design to space or formatting constraints (e.g. critical report, slide deck or publication)

Principles of Effective Graphic Design

Proximity – group related elements together

Alignment – elements on the same vertical or horizontal plane are perceived as having similar properties

Simplicity – cut anything superfluous, only include elements that add value, limit to 2-3 colors or fonts

White space (empty space) – use white space to minimize distraction & provide clarity

Legibility – sans serif fonts are easier to read, use color for emphasis instead of a new typeface

Color – select colors that present enough contrast to make the graph legible. Choose monochromatic color schemes to prevent clashing. Use dark colors and accent colors to emphasize important information

Visual Hierarchy – use color, font, image size, typeface, alignment & placement to create a viewing order

Focal Points – primary area of interest that immediately attracts the eye, emphasize the most important content and make it your focal point. Use contrasting colors to draw attention

Repetition – repeating elements can be visually appealing, repeated shapes, labels, colors

Familiarity – using familiar styles, icons, navigation structure makes viewers feel confident

Consistency – be consistent with heading sizes, font choices, color scheme, and spacing. Use images with similar styles

Effectiveness Ranking

A graph is a representation of data that visually encodes numerical values into attributes such as lines, symbols and colors. The Cleveland-McGill scale can be used to select the most effective attribute(s) for your purpose

Volume Color hue Depth: 3d position Color intensity Area Shape or Angle Length Position on unaligned scale Position on common scale



Least accurate → Most accurate

volume charts, poorly designed heat maps, multivariate density plots, heat maps, bubble charts, mosaic charts, line graphs, pie charts, stacked bar charts, bar charts, waterfall chart, small multiple plots, dot plots, bar charts, parallel coordinate plots

Selecting the right base graph

Consider if a standard graph can be used by identifying suitable designs based on the: (i) purpose (i.e. message to be conveyed or question to answer) and (ii) data (i.e. variables to display).

Deviation	Correlation	Ranking	Distribution	Evolution	Part-to-whole	Magnitude
Chg. from baseline	Scatter plot	Horizontal bar chart	Boxplot	Kaplan Meier	Stacked bar chart	Vertical bar chart
Waterfall	Heat map	Dotplot	Histogram	Line plot	Tree map	Forest plot

Facilitating Comparisons

Proximity improves association

Place labels next to data instead of using legends

Do not use color to differentiate categories of the same variable

Group together elements to be compared directly

Ease visual inspection

Order values to help compare across many categories

Judgments are easier to make on a common vertical scale

Reduce mental arithmetic

Plot the final comparison e.g. mean difference not two means

Use reference lines and other visual anchors

Color for emphasis or distinction

Restrained use of color is highly effective in organizing a narrative and calling attention to certain elements.

Think carefully before introducing additional color. Do you really need it?

Do not use color to differentiate categories of the same variable

Use colors or shades to represent meaningful differences such as positive/negative values, treatments or doses

Be consistent, use the same color to mean the same thing in a series of graphs (e.g. treatment, dose)

Use a bold, saturated or contrasting color to emphasize important details.

Emphasize the data by minimizing unnecessary ink, e.g. soften gridlines with a light color

Utilize existing resources for selection of appropriate palettes such as Color Brewer or Munsell

Implementation Considerations

Plot cause on the x-axis and effect on the y-axis. Use this standard convention in order to avoid misinterpretation.

Aspect ratio can influence interpretation. Aim for a 45 degree angle of change to avoid over-interpretation of slope.

Use position for comparisons rather than length (i.e. dots instead of bars), especially for non-linear scales (e.g. log scale or % change).

Do not plot log-normally distributed variables on a linear scale (e.g. hazard ratio, AUC, CL)

When displaying data measured on the same scale, also plot them on the same scale for easy comparison.

Connected data imply continuity. Do not connect data across a disconnected or uneven time scale.

Visits displayed close together are perceived to be closer in time. Space the visits proportional to the time between such in order to avoid confusion.

Plot data and inferences to support stories about modes.

Display text with enough contrast to be visible. Favor the use of dark on light instead of light on dark whenever possible.

Bold or italics should only be used for layering or emphasis. Emphasizing everything means nothing gets emphasized.

Try not to set text at an angle, as this decreases readability. Think of alternative solutions such as transposing the graph.

Legibility and Clarity

Effective graphs stand alone. They use titles, annotations, labels, shapes, colors, and textures to deliver important information.

Label axes with clear measurement units and provide annotations that support the message.

Use font size to create hierarchy (e.g. set titles 2x larger than all other labels to make them more prominent).

Do not type too small or too condensed. Break long titles into two lines. Shift or adjust size of labels that overlap.

Keep the font style simple – sans serif is easier to read.

Display text with enough contrast to be visible. Favor the use of dark on light instead of light on dark whenever possible.

Bold or italics should only be used for layering or emphasis. Emphasizing everything means nothing gets emphasized.

Try not to set text at an angle, as this decreases readability. Think of alternative solutions such as transposing the graph.

Good graph checklist

Clear Communication

- Is the message of the graph as clear as possible?
- Is it easy for someone unfamiliar with the data to interpret the graph?
- Are the patterns/relationships easily identified?
- Is the graph tailored to its primary purpose and audience?
- Is the correct graph type used?

Facilitating Comparisons

- Are elements to be compared grouped together?
- Are labels placed next to data instead of in legends?
- Have categories been ordered for easy comparison?
- Can the plot be read without doing mental calculations?
- Are the estimates of interest plotted (e.g. mean differences with confidence intervals)?

Color for emphasis or distinction

- Are graphical elements displayed in a dark color on a light background?
- Are grid lines drawn with a thin line and a light color such as grey?
- Are colors used sparingly (e.g. max 3)?
- Do all elements in the graph have a purpose (e.g. colors, textures, grid lines)?
- Are the same colors used to mean the same thing in a series of graphs?

Implementation Considerations

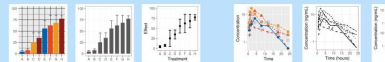
- Are multiple panels plotted on the same scale?
- Are lognormally distributed variables plotted on a log scale?
- Are common baselines used wherever possible?
- Does the orientation of the axes aid interpretation?
- Does the aspect ratio allow the reader to see variations in the data?
- Are data across a disconnected time scale kept disconnected?
- Are data spaced proportionally to the actual time interval (instead of according to visual number)?
- Are data and inferences plotted to support stories about modes?
- Are number of patients by group reported if this adds context?

Legibility and Clarity

- Can all graphical elements be seen?
- Does the graph have a clear title, axis labels, annotations and data units?
- Can the font be read without eye strain or effort?
- Are sans-serif fonts used?
- Do text sizes have correct hierarchy (big to small, main text to subtext)?
- Are the elements of the graph clearly labeled (e.g. points, error bars, lines, shaded regions)?
- Are labels oriented horizontally where possible?

Putting it all together – Remove the clutter & emphasize the message

Creating a graph is an iterative process: produce, review and refine.



Colors, backgrounds, and borders can be removed and gridlines reduced.

It is easier to see differences in position over a difference in length, i.e. a dot over a bar.

Using too many colors can be distracting. Use white background and try using other methods to distinguish different curves.

One solution could be repeating the data in different panels, highlighting individual curves in a darker color.

Resources

Books:
 E. R. Tufte. The visual display of quantitative information. Connecticut, Graphics Press, 2001.
 Cleveland, W.S. and McGill, Robert. Graphical perception: theory, experimentation and application to the development of graphical methods. JASA, Vol. 79, No. 387, pp. 531 – 554, 1984.
 S. Few. Show Me The Numbers: Designing Tables and Graphs to Enlighten (2nd Edition). Burrington, CA, Analytics Press, 2012.
 D. M. Wainwright. The Wall Street Journal Guide to Information Graphics: The Dots and Dots of Presenting Data, Facts, and Figures, December 16, 2013.
 J. Duromont. Trees, Maps, and Theorems: Effective communication for rational minds. PRINCIPAE, N. B. Robbins. Creating More Effective Graphs. Chart House.

Online resources:
<https://www.parsipalindrom.com/> (S. Few)
<https://www.infodivulge.com/> (E. Tufte)
<https://www.strobilisation.com/> (N. Wainwright)
<https://www.funfonctionnel.com/> (A. Duromont)
<https://www.strobilisation.com/> (A. Gemenet)
<https://www.funfonctionnel.com/> (A. Caro)
<https://www.strobilisation.com/> (N. Robbins)

Authors

Alison Margulies, Mark Ball, Bastien Magnussen, Julie Jones, Marc Vandendriessche



© 2017 Novartis



Planning questions



Why

Clearly identify the purpose of the graph, e.g. to deliver a message or for exploration?

What

Identify the quantitative evidence to support the purpose

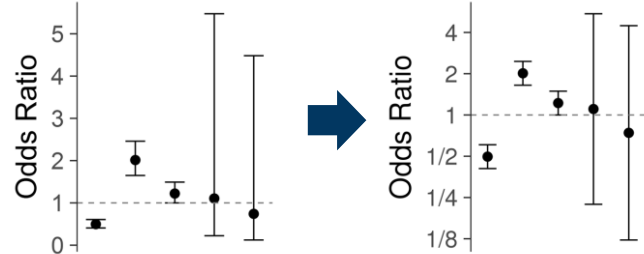
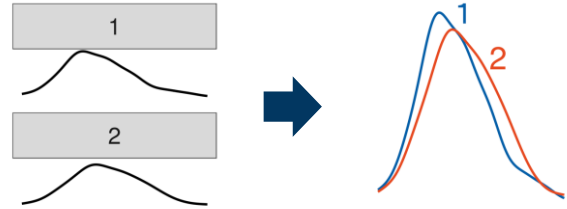
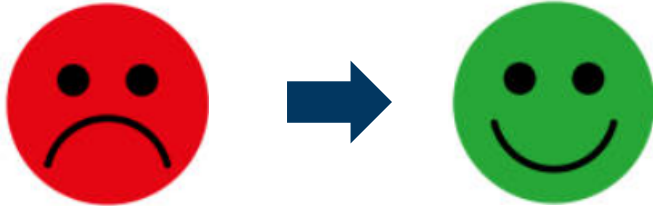
Who

Identify the intended audience (specialists, non-specialists, both) and focus the design to support their needs

Where

Adapt the design to space or formatting constraints (e.g. clinical report, slide deck or publication)

Tips & tricks



Good graph checklist



Good graph checklist	
Clear Communication	Implementation Considerations
<ul style="list-style-type: none"><input type="checkbox"/> Is the message of the graph as clear as possible?<input type="checkbox"/> Is it easy for someone unfamiliar with the data to interpret the graph?<input type="checkbox"/> Are the patterns/relationships easily identified?<input type="checkbox"/> Is the graph tailored to its primary purpose and audience?<input type="checkbox"/> Is the correct graph type used?	<ul style="list-style-type: none"><input type="checkbox"/> Are multiple panels plotted on the same scale?<input type="checkbox"/> Are lognormally distributed variables plotted on a log scale?<input type="checkbox"/> Are common baselines used wherever possible?<input type="checkbox"/> Does the orientation of the axes aid interpretation?<input type="checkbox"/> Does the aspect ratio allow the reader to see variations in the data?<input type="checkbox"/> Are data across a disconnected time scale kept disconnected?<input type="checkbox"/> Are data spaced proportionally to the actual time interval (instead of according to visit number)?<input type="checkbox"/> Are data and inferences plotted to support stories about models?<input type="checkbox"/> Are number of patients by group reported if this adds context?
Facilitating Comparisons	
<ul style="list-style-type: none"><input type="checkbox"/> Are elements to be compared grouped together?<input type="checkbox"/> Are labels placed next to data instead of in legends?<input type="checkbox"/> Have categories been ordered for easy comparison?<input type="checkbox"/> Can the plot be read without doing mental calculations?<input type="checkbox"/> Are the estimates of interest plotted (e.g. mean differences with confidence intervals)?	
Color for emphasis or distinction	Legibility and Clarity
<ul style="list-style-type: none"><input type="checkbox"/> Are graphical elements displayed in a dark color on a light background?<input type="checkbox"/> Are grid lines drawn with a thin line and a light color such as grey?<input type="checkbox"/> Are colors used sparingly (e.g. max 3)?<input type="checkbox"/> Do all elements in the graph have a purpose (e.g. colors, textures, grid lines)?<input type="checkbox"/> Are the same colors used to mean the same thing in a series of graphs?	<ul style="list-style-type: none"><input type="checkbox"/> Can all graphical elements be seen?<input type="checkbox"/> Does the graph have a clear title, axis labels, annotations and data units?<input type="checkbox"/> Can the font be read without eye strain or effort?<input type="checkbox"/> Are sans-serif fonts used?<input type="checkbox"/> Do text sizes have correct hierarchy (big to small, main text to subtext)?<input type="checkbox"/> Are the elements of the graph clearly labeled (e.g. points, error bars, lines, shaded regions)?<input type="checkbox"/> Are labels oriented horizontally where possible?

Graphics Principles Cheat Sheet

Graphics Principles Cheat Sheet v1.1

Communication

Effective visualizations communicate complex statistical and quantitative information facilitating insight, understanding, and decision making.

But what is an effective graph?

This cheat sheet provides general guidance and points to consider.

Planning

Clearly identify the purpose of the graph, e.g. to deliver a message or for exploration?

Identify the quantitative evidence to support the purpose

Identify the intended audience (specialists, non-specialists, both) and focus the design to support their needs

Adapt the design to space or formatting constraints (e.g. critical report, slide deck or publication)

Principles of Effective Graphic Design

Proximity – group related elements together
Alignment – elements on the same vertical or horizontal plane are perceived as having similar properties

Simplicity – cut anything superfluous, only include elements that add value, limit to 2-3 colors or fonts

White space (empty space) – use white space to minimize distraction & provide clarity

Legibility – sans serif fonts are easier to read, use color for emphasis instead of a new typeface
Color – select colors that present enough contrast to make the graph legible. Choose monochromatic color schemes to prevent clashing. Use dark colors and accent colors to emphasize important information

Visual Hierarchy – use color, font, image size, typeface, alignment & placement to create a viewing order

Focal Points – primary area of interest that immediately attracts the eye; emphasize the most important content and make it your focal point. Use contrasting colors to draw attention

Repetition – repeating elements can be visually appealing, repeated shapes, labels, colors

Familiarity – using familiar styles, icons, navigation structure makes viewers feel confident

Consistency – be consistent with heading sizes, font choices, color scheme, and spacing. Use images with similar styles

Effectiveness Ranking

A graph is a representation of data that visually encodes numerical values into attributes such as lines, symbols and colors. The Cleveland-McGill scale can be used to select the most effective attribute(s) for your purpose

Volume	Color hue	Depth: 3d position	Color intensity	Area	Slope or Angle	Length	Position on unaligned scale	Position on common scale

Least accurate

poorly designed heat maps
 multivariate density plots
 heat maps
 bubble charts, mosaic charts
 line graphs, pie charts, waterfall chart
 stacked bar charts, pie charts, small multiple plots
 dot plots, bar charts, parallel coordinate plots

Most accurate

Selecting the right base graph

Consider if a standard graph can be used by identifying suitable designs based on the: (i) purpose (i.e. message to be conveyed or question to answer) and (ii) data (i.e. variables to display).

Example plots categorized by purpose

Deviation	Correlation	Ranking	Distribution	Evolution	Part-to-whole	Magnitude
Chg. from baseline	Scatter plot	Horizontal bar chart	Boxplot	Kaplan Meier	Stacked bar chart	Vertical bar chart

Facilitating Comparisons

Proximity improves association

Place labels next to data instead of using legends

Do not use color to differentiate categories of the same variable

Ease visual inspection

Order values to help compare across many categories

Judgments are easier to make on a common vertical scale

Reduce mental arithmetic

Plot the final comparison e.g. mean difference not two means

Use reference lines and other visual anchors

Color for emphasis or distinction

Restrained use of color is highly effective in organizing a narrative and calling attention to certain elements. Think carefully before introducing additional color. Do you really need it?

Do not use color to differentiate categories of the same variable

Use colors or shades to represent meaningful differences such as positive/negative values, treatments or doses

Be consistent, use the same color to mean the same thing in a series of graphs (e.g. treatment, dose)

Use a bold, saturated or contrasting color to emphasize important details

Emphasize the data by minimizing unnecessary ink, e.g. soften gridlines with a light color

Utilize existing resources for selection of appropriate palettes such as Color Brewer or Munsell

Implementation Considerations

Plot cause on the x-axis and effect on the y-axis. Use this standard convention in order to avoid misinterpretation.

Aspect ratio can influence interpretation. Aim for a 45 degree angle of change to avoid over-interpretation of slope.

Use position for comparisons rather than length (i.e. dots instead of bars), especially for non-linear scales (e.g. log scale or % change).

Do not plot log-normally distributed variables on a linear scale (e.g. hazard ratio, AUC, CL)

When displaying data measured on the same scale, also plot them on the same scale for easy comparison.

Connected data imply continuity. Do not connect data across a disconnected or uneven time scale.

Visits displayed close together are perceived to be closer in time. Space the visits proportional to the time between such in order to avoid confusion.

Plot data and inferences to support stories about modes.

Exception: baseline or pre-dose

Use a bold, saturated or contrasting color to emphasize important details

Emphasize the data by minimizing unnecessary ink, e.g. soften gridlines with a light color

Utilize existing resources for selection of appropriate palettes such as Color Brewer or Munsell

Legibility and Clarity

Effective graphs stand alone. They use titles, annotations, labels, shapes, colors, and textures to deliver important information.

Label axes with clear measurement units and provide annotations that support the message.

Use font size to create hierarchy (e.g. set titles 2x larger than all other labels to make them more prominent).

Do not type too small or too condensed. Break long titles into two lines. Shift or adjust size of labels that overlap.

Keep the font style simple – sans serif is easier to read.

Display text with enough contrast to be visible. Favor the use of dark on light instead of light on dark whenever possible.

Bold or italics should only be used for layering or emphasis. Emphasizing everything means nothing gets emphasized.

Try not to set text at an angle, as this decreases readability. Think of alternative solutions such as transposing the graph.

Putting it all together – Remove the clutter & emphasize the message

Creating a graph is an iterative process: produce, review and refine.

Colors, backgrounds, and borders can be removed and gridlines reduced.

It is easier to see differences in position over a difference in length, i.e. a dot over a bar.

Using too many colors can be distracting. Use white background and try using other methods to distinguish different curves.

One solution could be repeating the data in different panels, highlighting individual curves in a darker color.

Good graph checklist

- | Clear Communication | Implementation Considerations |
|---|--|
| <input type="checkbox"/> Is the message of the graph as clear as possible? | <input type="checkbox"/> Are multiple panels plotted on the same scale? |
| <input type="checkbox"/> Is it easy for someone unfamiliar with the data to interpret the graph? | <input type="checkbox"/> Are lognormally distributed variables plotted on a log scale? |
| <input type="checkbox"/> Are the patterns/relationships easily identified? | <input type="checkbox"/> Are common baselines used wherever possible? |
| <input type="checkbox"/> Is the graph tailored to its primary purpose and audience? | <input type="checkbox"/> Does the orientation of the axes aid interpretation? |
| <input type="checkbox"/> Is the correct graph type used? | <input type="checkbox"/> Does the aspect ratio allow the reader to see variations in the data? |
| Facilitating Comparisons | <input type="checkbox"/> Are data across a disconnected time scale kept disconnected? |
| <input type="checkbox"/> Are elements to be compared grouped together? | <input type="checkbox"/> Are data spaced proportionally to the actual time interval (instead of according to visual number)? |
| <input type="checkbox"/> Are labels placed next to data instead of in legends? | <input type="checkbox"/> Are data and inferences plotted to support stories about modes? |
| <input type="checkbox"/> Have categories been ordered for easy comparison? | <input type="checkbox"/> Are number of patients by group reported if this adds context? |
| <input type="checkbox"/> Can the plot be read without doing mental calculations? | Legibility and Clarity |
| <input type="checkbox"/> Are the estimates of interest plotted (e.g. mean differences with confidence intervals)? | <input type="checkbox"/> Can all graphical elements be seen? |
| Color for emphasis or distinction | <input type="checkbox"/> Does the graph have a clear title, axis labels, annotations and data units? |
| <input type="checkbox"/> Are graphical elements displayed in a dark color on a light background? | <input type="checkbox"/> Can the font be read without eye strain or effort? |
| <input type="checkbox"/> Are grid lines drawn with a thin line and a light color such as grey? | <input type="checkbox"/> Are sans-serif fonts used? |
| <input type="checkbox"/> Are colors used sparingly (e.g. max 3)? | <input type="checkbox"/> Do text sizes have correct hierarchy (big to small, main text to subtext)? |
| <input type="checkbox"/> Do all elements in the graph have a purpose (e.g. colors, textures, grid lines)? | <input type="checkbox"/> Are the elements of the graph clearly labeled (e.g. points, error bars, lines, shaded regions)? |
| <input type="checkbox"/> Are the same colors used to mean the same thing in a series of graphs? | <input type="checkbox"/> Are labels oriented horizontally where possible? |

Resources

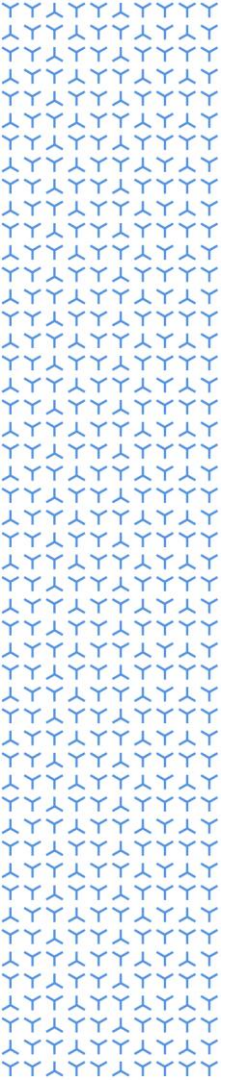
Books:
 E. R. Tufte. The visual display of quantitative information. Connecticut: Graphics Press, 2001.
 Cleveland, W.S. and McGill, Robert. Graphical perception: theory, experimentation and application to the development of graphical methods. JASA, Vol. 79, No. 387, pp. 531 – 554, 1984.
 S. Few. Show Me The Numbers: Designing Tables and Graphs to Enlighten (2nd Edition). Burlingame, CA: Analysis Press, 2012.
 D. M. Wiegman. The Wall Street Journal Guide to Information Graphics: The Dos and Don'ts of Presenting Data, Facts, and Figures. December 16, 2013.
 J. Duromout. Trees, Maps, and Theorems: Effective communication for rational minds. PRINCIPAE. N. B. Robbins. Creating More Effective Graphs. Chart House.

Online resources:
<https://www.parsipalindrom.com/> (S. Few)
<https://www.visualization.tufts.edu/> (E. Tufte)
<https://www.d3.js.org/> (M. Van Wazer)
<https://www.strobilium.be/> (J. Duromout)
<https://www.chartio.com/> (A. Gemeny)
<https://www.thefunctionalcolor.com/> (A. Caro)
<https://www.color-matters.com/> (N. Robbins)

Authors

Allison Margulies, Mark Sallis, Bastien Magnussen, Julie Jones, Marc Vandendriessche



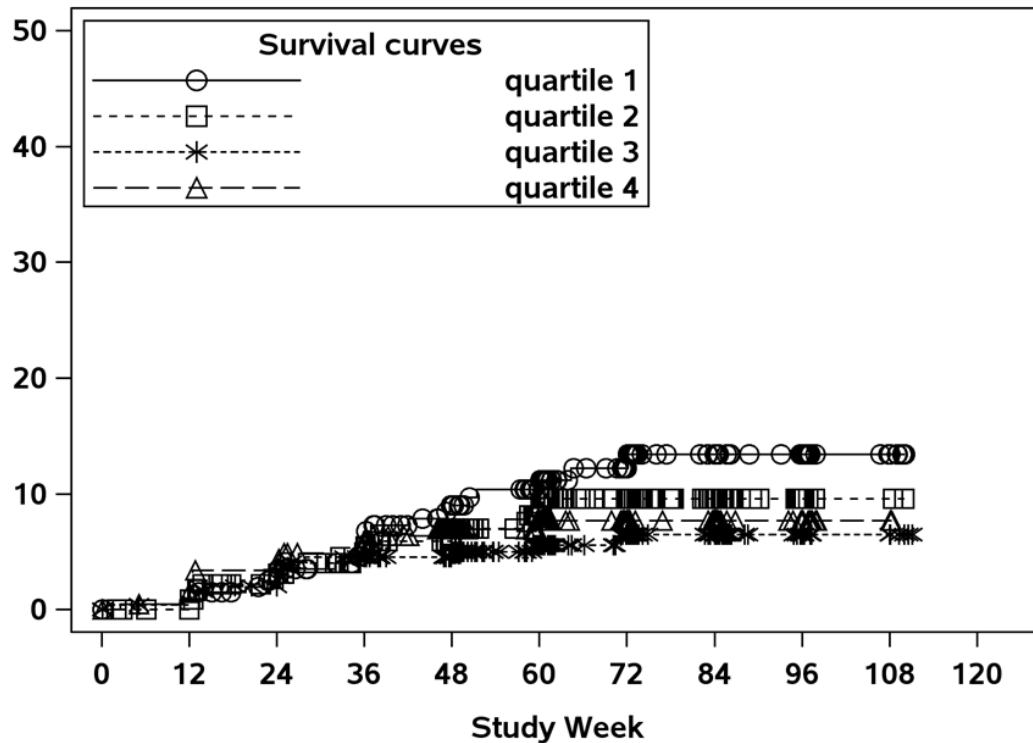


Principles of effective visual communication

This is a continual process

Planned Treatment:

mg



How do we get there?

- How do we tell a good graph from a bad one?
- How do we ensure we design a graph that is fit for purpose?
- Recall workstream goal: enhance understanding and use of **good graphical principles**

Effective Visual Communication (EVC)!

CPT: Pharmacometrics & Systems Pharmacology

Tutorial | [Open Access](#) | [CC](#) | [i](#) | [S](#)

Effective Visual Communication for the Quantitative Scientist

Marc Vandemeulebroecke [✉](#), Mark Baillie, Alison Margolskee, Baldur Magnusson

First published: 22 July 2019 | <https://doi.org/10.1002/psp4.12455> | Citations: 7



3 principles for better visual communication

1. Have a clear **purpose**

- Understand the question you are trying to answer
- Identify the quantitative evidence to answer that question
- Know your audience and focus the design to support their needs

2. Show the data **clearly**

- Choose the appropriate graph type to display your data
- Avoid misrepresentation (use appropriate scales)
- Maximize data to ink ratio (reduce distraction, less is more)

3. Make the **message** obvious

- Minimize mental arithmetic (e.g. plot the difference)
- Use proximity and alignment to aid in comparisons
- Use colors and annotations to highlight important details



Learning outcomes for today's workshop

- Appreciate why Effective Visual Communication (EVC) is a key competency for the quantitative scientist.
- Explain the three principles of EVC (purpose, clarity and message).
- Design a visualization based on a specific purpose.
- Redesign a visualization to show data clearly.
- Enhance the message of a visualization.
- Recognize where to apply the three EVC principles in your daily work.

Purpose



What is the purpose of the visualization?

What is the main objective of the visualization?

The visualization is to display supporting evidence that LNPO23 has demonstrated proof of concept and is a good candidate to take into phase 3 development.

List the (scientific) question(s) the visualization is trying to answer. Try to be specific.

- *Is there a decrease in LDH to "normal levels" post LNPO23 dose as a mono and combo therapy?*
- *Does LNPO23 increase hemoglobin levels?*

What is the key evidence that is available to answer the question?

Two studies.

Two different dose cohorts in one study. Mono and combo.

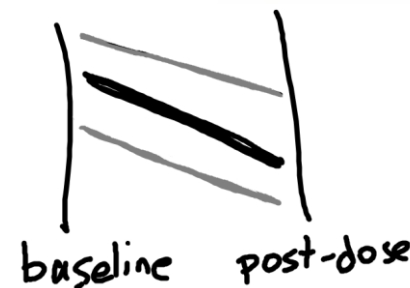
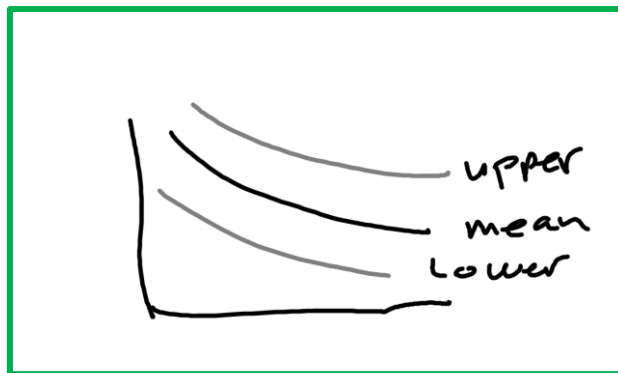
LDH is a surrogate measure of efficacy for PNH.

Consistency across gender for Hemoglobin improvement.

2nd principle – select the appropriate graph

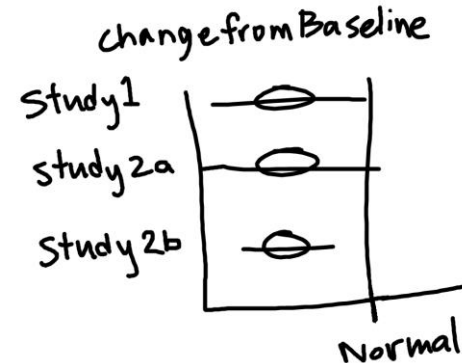
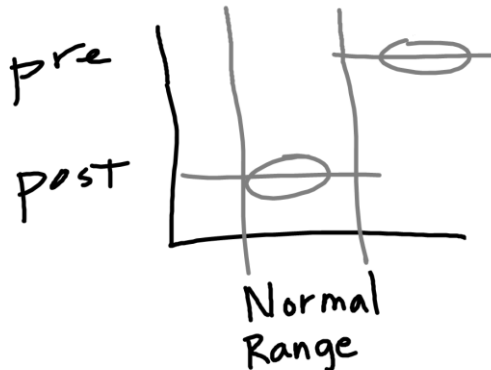


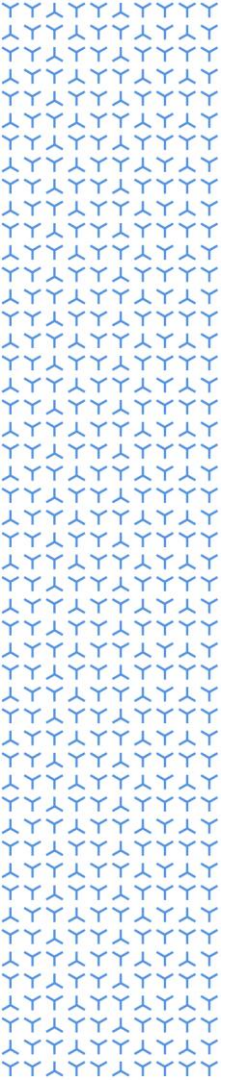
- Brainstorm different ways to display the data
- The more ideas the better!
- Select the display that supports the key message



Continuing with LNP example


- What is the key message:
 - LNP023 reduces LDH levels to normal
- The key evidence to support this:
 - Two studies with different dose cohorts
 - LDH as a surrogate for efficacy





3. Graph Gallery

3. Graph Gallery: Providing access to example code and data



GRAPH GALLERY

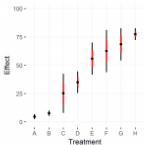
Email Password

[Register](#), if you have not already; [Reset password](#), if you forgot it

[Home](#) [Samples](#) [Gallery](#) [Discussion Forum](#) [Upload Request](#) [Request for Help](#) [Links](#) [FAQ](#)

127 Results

Order by: [Views](#) | [Downloads](#) - [Rating](#)

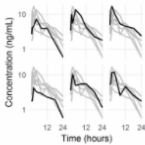


1593: Revising a dynamite plot

Creating a graph is an iterative process: produce, review and refine.

[R Version 3.2.3](#) • [Barchart](#) • [Cheatsheet](#) • [Continuous](#) • [Dotplot](#) • [Efficacy](#) • [Forestplot](#) • [GGPLOT2](#) • [Model-based](#) • [R](#) • [Safety](#)

Views: 108



1586: Demonstration of a multi panel (small multiples) line plot

Creating a graph is an iterative process: produce, review and refine.

[R Version 3.2.3](#) • [Cheatsheet](#) • [Continuous](#) • [GGPLOT2](#) • [Lineplot](#) • [Model-based](#) • [Panel](#) • [R](#) • [Time Line](#)

Views: 100

Search

Filter

Logic

AND OR

Language

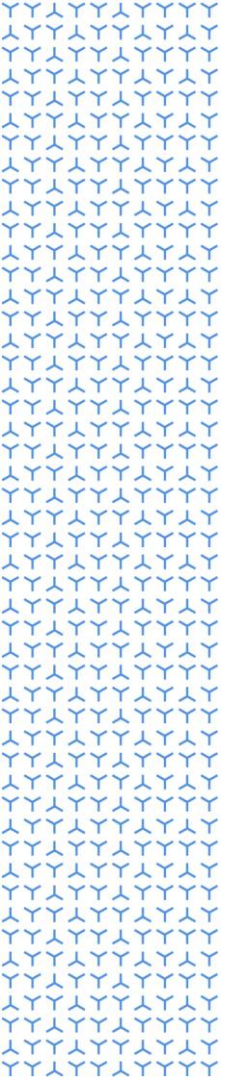
R SAS
 MATLAB all

Data Types

Binary Continuous
 Mixed Model-based
 Ordinal Time-To-Event

Sample Type

Barchart Boxplot
 Distribution Dotplot
 Forestplot Lineplot
 Other Panel
 Patient profile/listing Scatterplot
 Survival Spaghetti
 Time Line



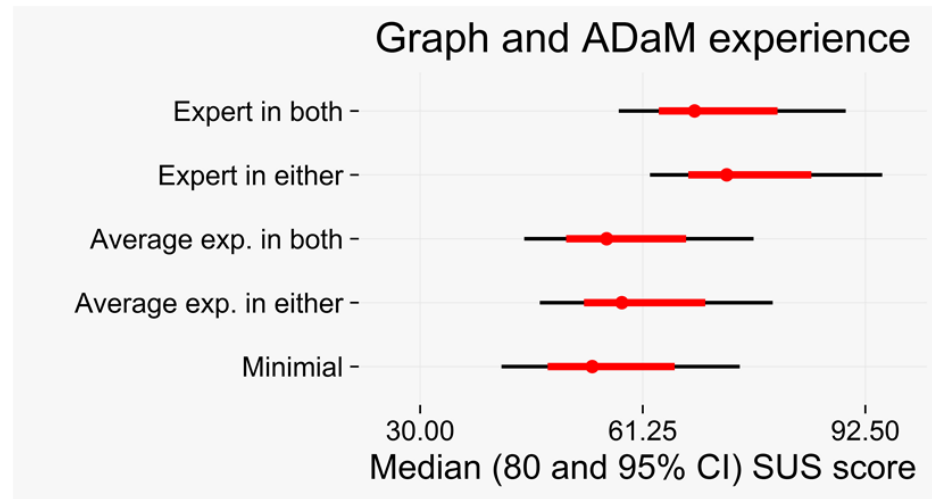
4. Analysis results data sets

How can we help with implementation across varying skill levels?

A survey was sent to associates working with clinical data

The purpose to:

- evaluate ADaM as a data standard for graph production
- Identify key issues associates currently experience
- Explore issues related to role and experience level
- 85 respondents

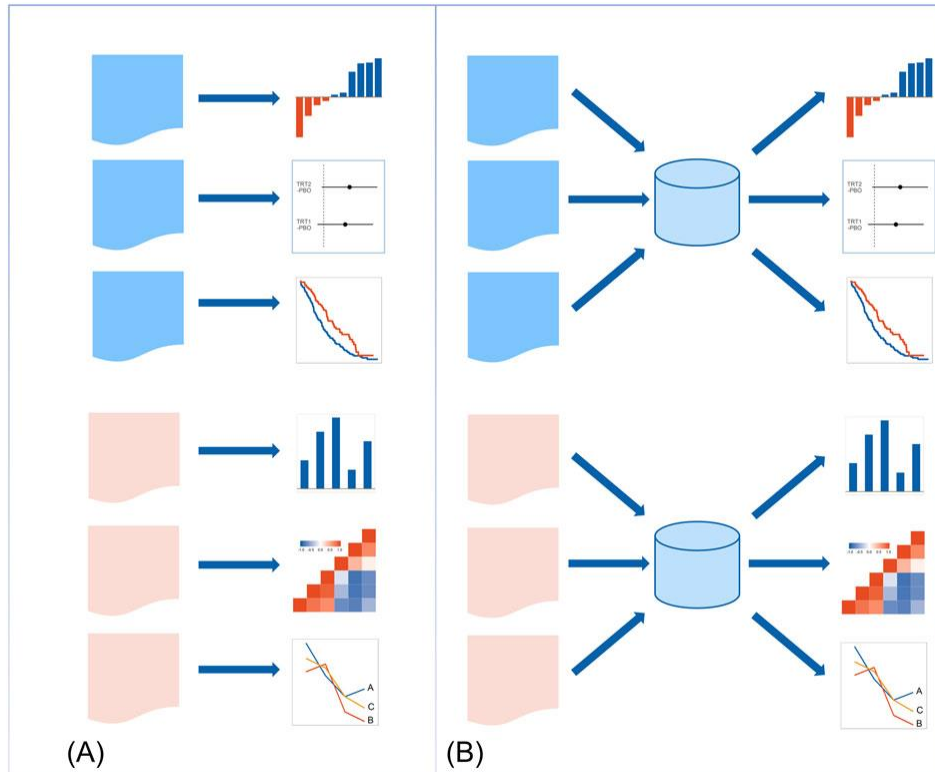


How can we help with implementation across varying skill levels?

(Negative) qualitative comments

- ***Complex graphs such as Forest Plots will need a fair amount of data manipulation to get all needed for input to the graph***
- ***some variables needed for graphs are not in ADaM datasets***
- ***Trying to figure out what the different parameters mean and **extracting the information relevant to my task.*****

Analysis results data sets



Aliskiren, Enalapril, or Aliskiren and Enalapril in Heart Failure

John J.V. McMurray, M.D., Henry Krum, M.B., B.S., Ph.D., William T. Abraham, M.D., Kenneth Dickstein, M.D., Ph.D., Lars V. Køber, M.D., D.M.Sc., Akshay S. Desai, M.D., M.P.H., Scott D. Solomon, M.D., Nicola Greenlaw, M.Sc., M. Atif Ali, B.A., Yanntong Chiang, Ph.D., Qing Shao, Ph.D., Georgia Tarnesby, M.B., B.Chir., *et al.*, for the ATMOSPHERE Committees Investigators[†]

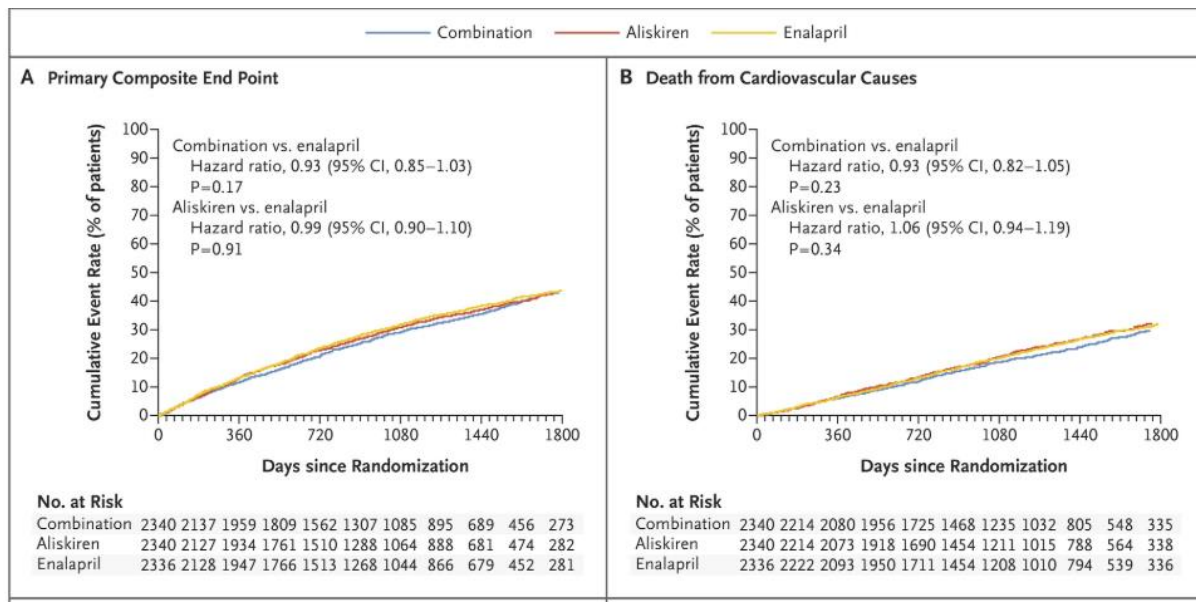


Table 2. Protocol-Specified Primary and Secondary Outcomes.*

Outcome	Combination Therapy (N = 2340)	Aliskiren (N = 2340)	Enalapril (N = 2336)	Combination Therapy vs. Enalapril		Aliskiren vs. Enalapril	
				Hazard Ratio or Difference (95% CI)	P Value	Hazard Ratio or Difference (95% CI)	P Value
Primary composite outcome: death from cardiovascular causes or first hospitalization for worsening heart failure — no. (%)	770 (32.9)	791 (33.8)	808 (34.6)	0.93 (0.85 to 1.03)	0.17	0.99 (0.90 to 1.10)	0.91
Death from cardiovascular causes	512 (21.9)	562 (24.0)	547 (23.4)	0.93 (0.82 to 1.05)	0.23	1.06 (0.94 to 1.19)	0.34
First hospitalization for worsening heart failure	430 (18.4)	442 (18.9)	452 (19.3)	0.93 (0.82 to 1.06)	0.29	0.99 (0.87 to 1.13)	0.91
Secondary outcome: change in KCCQ clinical summary score at 12 mo [†]	-5.04±0.56	-6.03±0.57	-5.01±0.55	-0.03 (-1.56 to 1.50)	0.97	-1.02 (-2.56 to 0.52)	0.20
Other prespecified exploratory outcomes — no. (%) [‡]							
Death from cardiovascular causes, hospitalization for heart failure, nonfatal myocardial infarction, nonfatal stroke, or resuscitated cardiac arrest	841 (35.9)	874 (37.4)	877 (37.5)	0.94 (0.86 to 1.04)	0.23	1.01 (0.92 to 1.11)	0.80
Fatal or nonfatal myocardial infarction	88 (3.8)	84 (3.6)	100 (4.3)	0.87 (0.66 to 1.16)	0.36	0.85 (0.64 to 1.14)	0.28
Fatal or nonfatal stroke	87 (3.7)	103 (4.4)	93 (4.0)	0.93 (0.70 to 1.25)	0.65	1.12 (0.85 to 1.49)	0.42
First resuscitated cardiac arrest	31 (1.3)	35 (1.5)	32 (1.4)	0.96 (0.58 to 1.57)	0.86	1.10 (0.68 to 1.78)	0.69
Death from any cause	595 (25.4)	654 (27.9)	646 (27.7)	0.91 (0.82 to 1.02)	0.12	1.04 (0.93 to 1.16)	0.46
Composite renal outcome — no. (%) [§]	39 (1.7)	26 (1.1)	18 (0.8)	2.17 (1.24 to 3.79)	0.007	1.50 (0.82 to 2.74)	0.18

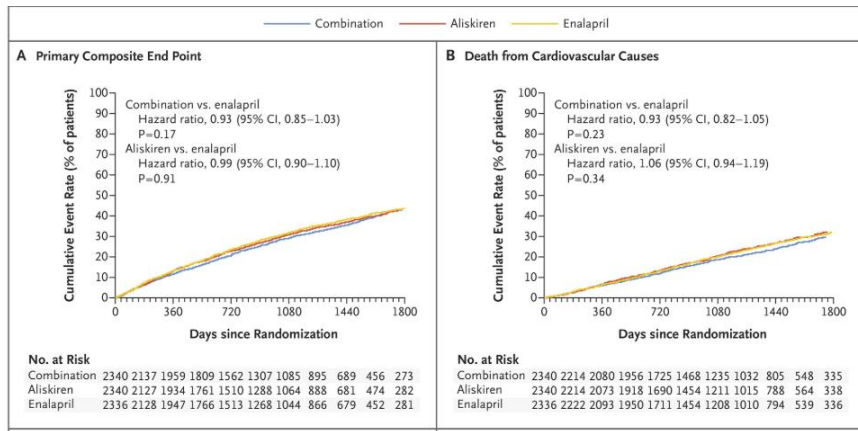


Table 2. Protocol-Specified Primary and Secondary Outcomes.*

Outcome	Combination Therapy (N=2340)	Aliskiren (N=2340)	Enalapril (N=2336)	Combination Therapy vs. Enalapril		Aliskiren vs. Enalapril	
				Hazard Ratio or Difference (95% CI)	P Value	Hazard Ratio or Difference (95% CI)	P Value
Primary composite outcome: death from cardiovascular causes or first hospitalization for worsening heart failure — no. (%)	770 (32.9)	791 (33.8)	808 (34.6)	0.93 (0.85 to 1.03)	0.17	0.99 (0.90 to 1.10)	0.91
Death from cardiovascular causes	512 (21.9)	562 (24.0)	547 (23.4)	0.93 (0.82 to 1.05)	0.23	1.06 (0.94 to 1.19)	0.34
First hospitalization for worsening heart failure	430 (18.4)	442 (18.9)	452 (19.3)	0.93 (0.82 to 1.06)	0.29	0.99 (0.87 to 1.13)	0.91
Secondary outcome: change in KCCQ clinical summary score at 12 mo†	-5.04±0.56	-6.03±0.57	-5.01±0.55	-0.03 (-1.56 to 1.50)	0.97	-1.02 (-2.56 to 0.52)	0.20
Other prespecified exploratory outcomes — no. (%)‡							
Death from cardiovascular causes, hospitalization for heart failure, nonfatal myocardial infarction, nonfatal stroke, or resuscitated cardiac arrest	841 (35.9)	874 (37.4)	877 (37.5)	0.94 (0.86 to 1.04)	0.23	1.01 (0.92 to 1.11)	0.80
Fatal or nonfatal myocardial infarction	88 (3.8)	84 (3.6)	100 (4.3)	0.87 (0.66 to 1.16)	0.36	0.85 (0.64 to 1.14)	0.28
Fatal or nonfatal stroke	87 (3.7)	103 (4.4)	93 (4.0)	0.93 (0.70 to 1.25)	0.65	1.12 (0.85 to 1.49)	0.42
First resuscitated cardiac arrest	31 (1.3)	35 (1.5)	32 (1.4)	0.96 (0.58 to 1.57)	0.86	1.10 (0.68 to 1.78)	0.69
Death from any cause	595 (25.4)	654 (27.9)	646 (27.7)	0.91 (0.82 to 1.02)	0.12	1.04 (0.93 to 1.16)	0.46
Composite renal outcome — no. (%)§	39 (1.7)	26 (1.1)	18 (0.8)	2.17 (1.24 to 3.79)	0.007	1.50 (0.82 to 2.74)	0.18

UANALID	TRTVAR	TRTVAl	AVISITN	PARAMCD	ROWCAT1	ANLTP1	ANLTP2	STAT	STATVAL	ANLMETH
<STUDYID>_<RA>_XXX1	TRT01P	Combination		PCE	Combination vs. Enalapril	RESPONSE	Experimental	SMALLN	770	Lifetest KM
<STUDYID>_<RA>_XXX1	TRT01P	Combination		PCE	Combination vs. Enalapril	RESPONSE	Experimental	BIGN	2340	Lifetest KM
<STUDYID>_<RA>_XXX1	TRT01P	Combination		PCE	Combination vs. Enalapril	RESPONSE	Experimental	PERCENT	32.9	Lifetest KM
<STUDYID>_<RA>_XXX1	TRT01P	Enalapril		PCE	Combination vs. Enalapril	RESPONSE	Enalapril	SMALLN	808	Lifetest KM
<STUDYID>_<RA>_XXX1	TRT01P	Enalapril		PCE	Combination vs. Enalapril	RESPONSE	Enalapril	BIGN	2336	Lifetest KM
<STUDYID>_<RA>_XXX1	TRT01P	Enalapril		PCE	Combination vs. Enalapril	RESPONSE	Enalapril	PERCENT	34.6	Lifetest KM
<STUDYID>_<RA>_XXX1	TRT01P			PCE	Combination vs. Enalapril	COMPARISON		Hazard	0.93	Lifetest KM
<STUDYID>_<RA>_XXX1	TRT01P			PCE	Combination vs. Enalapril	COMPARISON		95CILOW	0.846	Lifetest KM
<STUDYID>_<RA>_XXX1	TRT01P			PCE	Combination vs. Enalapril	COMPARISON		95CIHIGH	1.03	Lifetest KM
<STUDYID>_<RA>_XXX1	TRT01P			PCE	Combination vs. Enalapril	COMPARISON		1sidedp	0.0862	Lifetest KM
<STUDYID>_<RA>_XXX1	TRT01P			PCE	Combination vs. Enalapril	COMPARISON		2sidedp	0.1724	Lifetest KM
<STUDYID>_<RA>_XXX1	TRT01P			PCE	Combination vs. Enalapril	COMPARISON		adj1sidedp	0.3448	Lifetest KM
<STUDYID>_<RA>_XXX1	TRT01P	Combination		PCE	Non Diabetes: Combination vs Enalapril(1)	RESPONSE	Experimental	SMALLN		Lifetest KM
<STUDYID>_<RA>_XXX1	TRT01P	Combination		PCE	Non Diabetes: Combination vs Enalapril(1)	RESPONSE	Experimental	BIGN		Lifetest KM
<STUDYID>_<RA>_XXX1	TRT01P	Combination		PCE	Non Diabetes: Combination vs Enalapril(1)	RESPONSE	Experimental	PERCENT		Lifetest KM

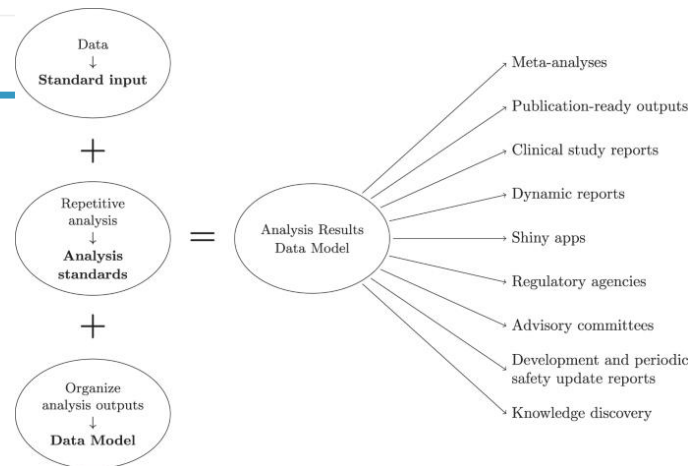
Rethinking clinical study data: why we should respect analysis results as data

[Joana M. Barros](#) , [Lukas A. Widmer](#), [Mark Baillie](#)  & [Simon Wandel](#)

[Scientific Data](#) **9**, Article number: 686 (2022) | [Cite this article](#)

2762 Accesses | **7** Altmetric | [Metrics](#)

<https://www.nature.com/articles/s41597-022-01789-2>



Pre-clinical

Organize

SEND

Tabulation
for animal
studies

Clinical

Plan

PRM

Model for
planning

Collect

CDASH

Model
for data
collection

Organize

SDTM

Model for
tabulation of
study data

Analyze

ADaM

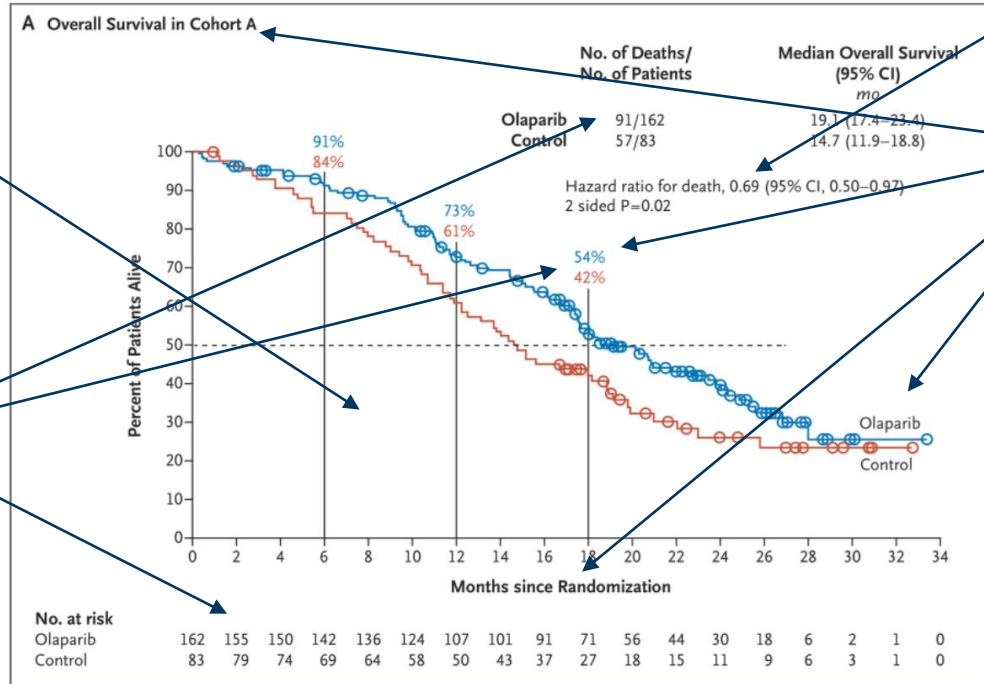
Analysis
data model

The anatomy of a KM plot

Additional inferential statistics

Data area

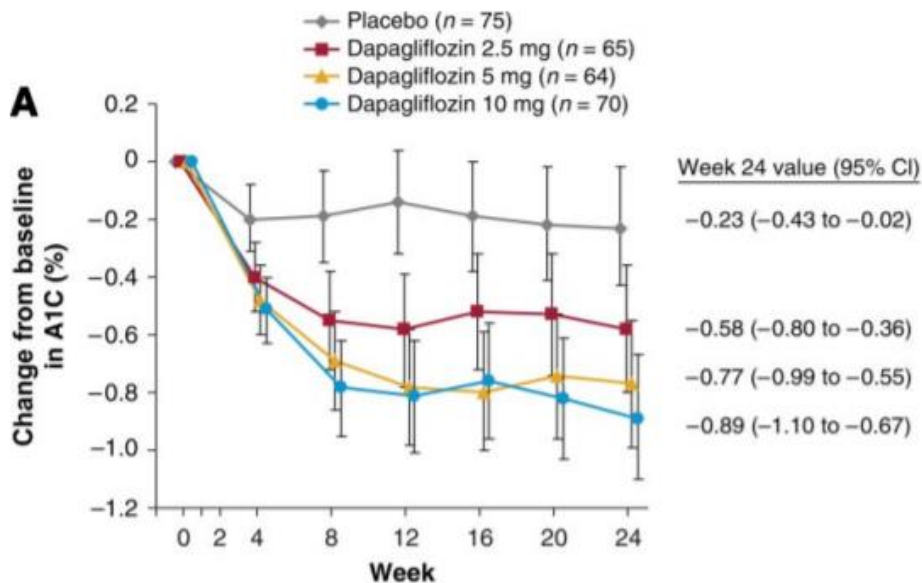
Additional summary information for context



Titles, labels and annotations

Activity 1 | Results from Analysis of Primary Endpoint

Study 2013



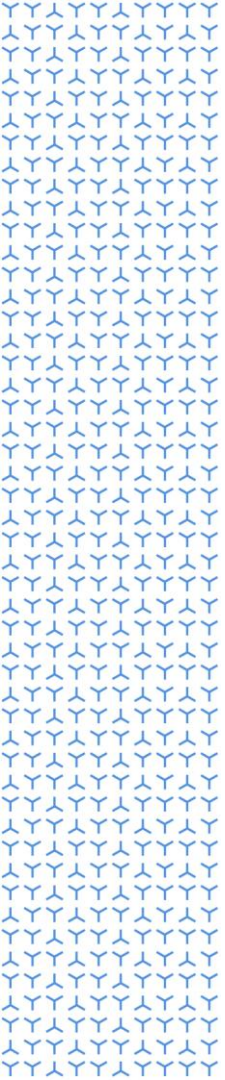
Diabetes Care 33:2217-2224, 2010

Treatment	Treatment difference versus placebo*	P-value	% Patients with rescue medication
placebo			12% (9/75)
2.5 mg	-0.35	p < 0.05	11% (7/65)
5 mg	-0.54	p < 0.001	2% (1/64)
10 mg	-0.66	p < 0.001	0% (0/70)

* Difference between means in change from baseline HbA1c at Week 24

Remember: HbA1c values collected after rescue were **excluded** from the analysis.

Interesting point: % of patients with rescue were not presented in the primary manuscript!



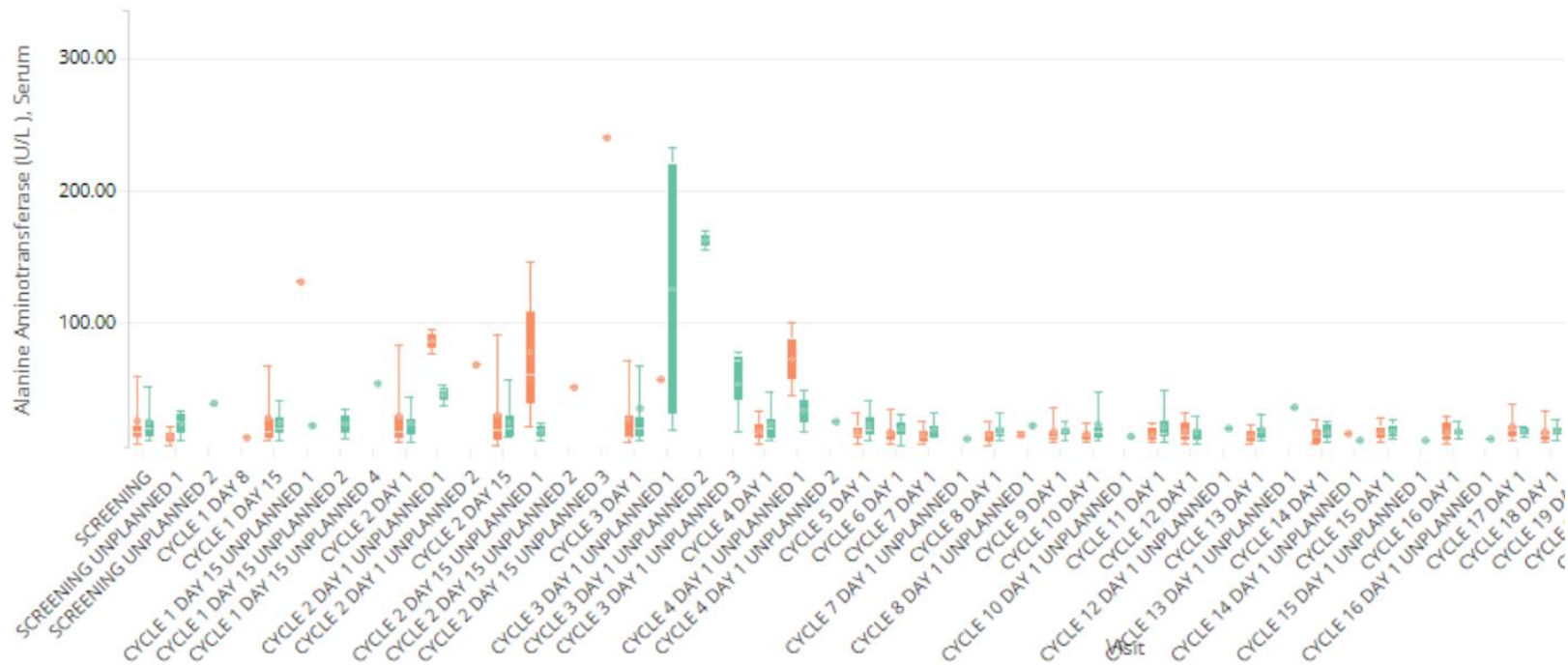
Lessons from a company wide data visualization initiative

Why does it matter?

- Data visualization and more importantly Effective Visual Communication (EVC) is essential in every step of a quantitative workflow, from scoping to execution through to communicating results and conclusions
- Within the quantitative sciences, traditional university and professional training curricula have not focused on EVC
- One hears regularly “I don’t understand this graph”
- A visual representation can be very powerful if it is clear and aligned with the intended purpose

This is a continual process

50 of 50 participant(s) shown (100.0%)



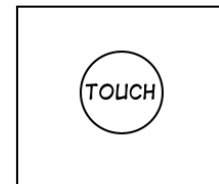
R for Pharma



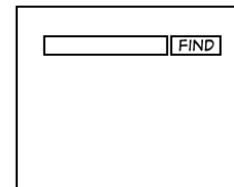
Pilot 2 Shiny Application



TYPICAL APPLE PRODUCT...



A GOOGLE PRODUCT...



YOUR COMPANY'S APP...

FIRST NAME: TYPE CD:

LAST NAME: TQP STAT:

SSN: FT/PT: VER:

ID: CAT CD:

PHONE 1: CITY:

PHONE 2: STATE:

ADDR 1: ZIP:

ACCT #: ORD #:

4 - K
AA2-
DK9B
KKA?
CN3
AA-9

NEW
DEL

OKAY APPLY SAVE UNDO HELP DELETE EDIT

SELECT BROWSE ERRORS

Effective visualization is a practice!

- Tufte: “Graphical competence demands three quite different skills: the substantive, statistical, and artistic.”
- These skills cannot be learned by reading an article
- Adopt visualization in every part of your workflow; make it a habit
- Think graphically
- Use pencil and paper before coding in software
- Calculate and communicate
- Test and repeat

Acknowledgements

- Alison Margolskee
 - Marc Vandemeulebroecke
 - Baldur Magnusson
 - David Carr
 - Andrew Wright
 - Jürgen Löffler
 - Linda Kanitra
 - Baldur Magnusson
 - Matthew Brierley
 - David Ohlssen
 - Ivo Masseling
 - Armin Griesel
 - Janice Branson
 - Lorenz Uhlmann
 - Shaun Butcher
 - Julie Jones
 - Walter Hufford
 - Ruquan You
 - Ivan-Toma Vranesic
 - Ian Rees
 - Nicolas Guerro
 - Keo Chanthavinout
 - Frank Bretz
 - Bonnie Petrucka
 - Malika Cremer
- And many more...

Resources, where to find out more?

Graphics principles

Tutorial

Effective Visual Communication is essential in every step of a quantitative workflow, from scoping to execution and communicating results and conclusions. The goal of the [tutorial](#) is to convey this competency.

In the [tutorial](#) we posit three laws of Effective Visual Communication for the quantitative scientist: have a clear purpose, show the data clearly, and make the message obvious. The [tutorial](#) provides more granularity on these laws; it covers the [Graphics Principles Cheat Sheet](#); and it discusses the application of both in four [Case Studies](#).

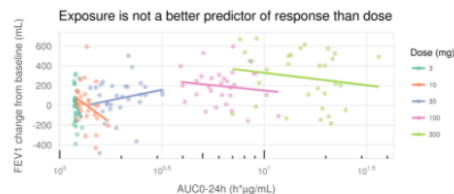


Figure 1: Example visualization from the tutorial

Additional resources:

- [Effective Visual Communication Website](https://graphicsprinciples.github.io/)
- [PKPD Exploratory graphics \(xGx\)](https://opensource.nibr.com/xgx)
- [Tutorial on effective visual communication](https://ascpt.onlinelibrary.wiley.com/doi/full/10.1002/psp4.12455)
- [Video on the three principles](https://youtu.be/pfxulpF9XOw)
- [Presentation checklist](https://opensource.nibr.com/xgx/Resources/Presentation_Checklist_v2_03.pdf)
- [Wonderful Wednesdays](https://www.psiweb.org/sigs-special-interest-groups/visualisation/welcome-to-wonderful-wednesdays)



Thank you

Contact:

mark.baillie@novartis.com