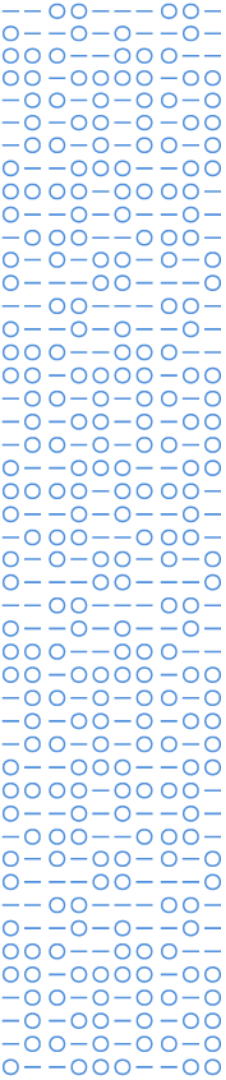


# How do we make better graphs? Effective visual communication for the quantitative scientist

**Mark Baillie**

**July 18<sup>th</sup>, 2019**

**Leuven, STRATOS symposia, ISCB 2019**



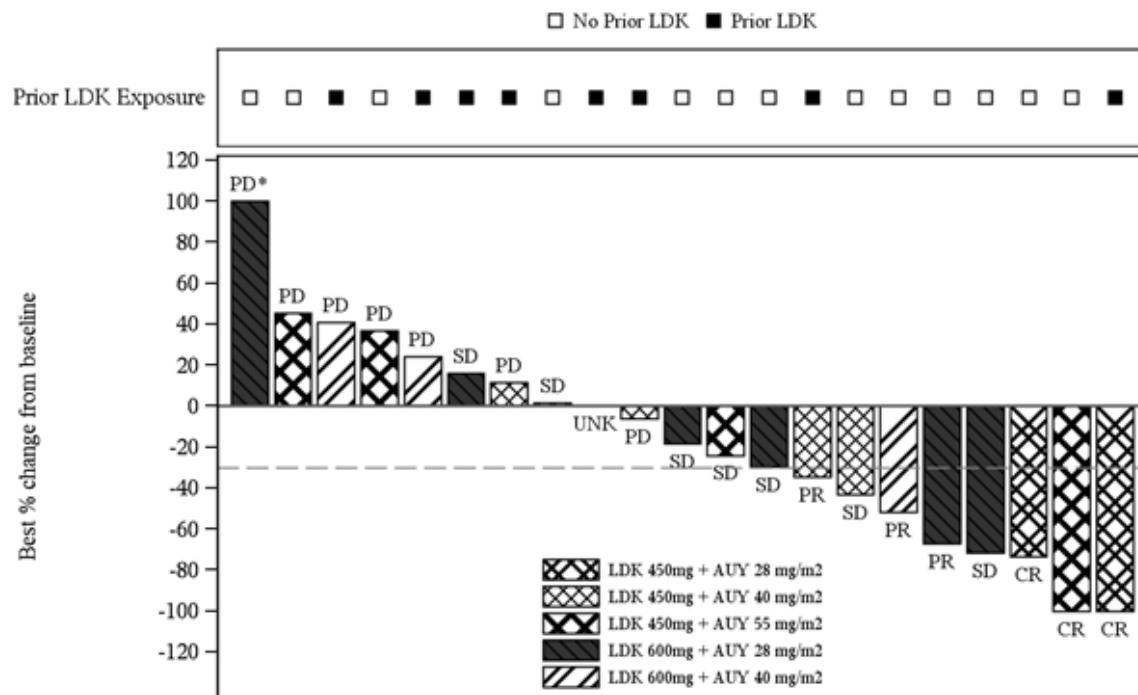
**Table 11-4**  
**Summary of best overall response as per investigator by treatment group**  
**(Full analysis set)**

	LDK 450mg + AUY 28 mg/m <sup>2</sup> N=3 n (%)	LDK 450mg + AUY 40 mg/m <sup>2</sup> N=5 n (%)	LDK 450mg + AUY 55 mg/m <sup>2</sup> N=4 n (%)	LDK 600mg + AUY 28 mg/m <sup>2</sup> N=6 n (%)	LDK 600mg + AUY 40 mg/m <sup>2</sup> N=4 n (%)	All patients N=22 n (%)
Best overall response						
Complete response (CR)	2 ( 66.7)	0	1 ( 25.0)	0	0	3 ( 13.6)
Partial response (PR)	0	1 ( 20.0)	0	1 ( 16.7)	1 ( 25.0)	3 ( 13.6)
Stable disease (SD)	0	2 ( 40.0)	1 ( 25.0)	4 ( 66.7)	0	7 ( 31.8)
Progressive disease (PD)	0	2 ( 40.0)	2 ( 50.0)	1 ( 16.7)	2 ( 50.0)	7 ( 31.8)
Unknown	1 ( 33.3)	0	0	0	1 ( 25.0)	2 ( 9.1)
Overall response rate (ORR)	2 ( 66.7)	1 ( 20.0)	1 ( 25.0)	1 ( 16.7)	1 ( 25.0)	6 ( 27.3)
(CR or PR)						
95% CI	(9.4-99.2)	(0.5-71.6)	(0.6-80.6)	(0.4-64.1)	(0.6-80.6)	(10.7-50.2)
Disease control rate (DCR)	2 ( 66.7)	3 ( 60.0)	2 ( 50.0)	5 ( 83.3)	1 ( 25.0)	13 ( 59.1)
(CR or PR or SD)						
95% CI	(9.4-99.2)	(14.7-94.7)	(6.8-93.2)	(35.9-99.6)	(0.6-80.6)	(36.4-79.3)

**Best overall response is based on investigator's assessment of disease status using RECIST 1.1.**  
**Estimate (95%CI) for ORR and DCR were obtained using exact binomial 95% confidence interval test.**

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 LDK378 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



# Elements of a graphics 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

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

# STRATOS Visualisation panel

“Visualization and the use of graphics can help at every stage of an analysis, from the planning and design of an experiment, the very first data explorations, through to the communication of conclusions and recommendations.

Visualization is more than "plotting data"; it can lead to a deeper understanding and inform next steps.

The role of the STRATOS visualization panel is to promote the use of good graphical principles for effective visual communication, providing guidance and recommendations covering all aspects from the design, implementation and review of statistical graphics.”

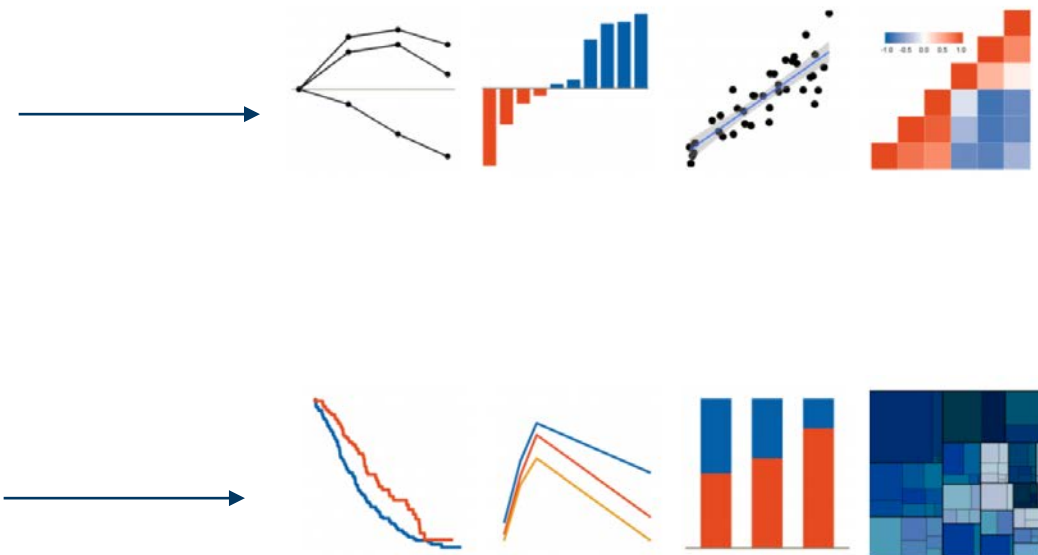
<http://www.stratos-initiative.org>



# Effective visualisation is important throughout the workflow

## Topic groups

1	Missing data
2	Selection of variables and functional forms in multivariable analysis
3	Initial data analysis
4	Measurement error and misclassification
5	Study design
6	Evaluating diagnostic tests and prediction models
7	Causal inference
8	Survival analysis
9	High-dimensional data





“Data Display is critical to data analysis. Graphs allow us to explore data to see overall patterns and to see detailed behavior; no other approach can compete in revealing the structure of data so thoroughly.

Graphs allow us to view complex mathematical models fitted to data, and they allow us to assess the validity of such models”

William Cleveland,  
The Elements of Graphing Data (1985)



# How do we get there?

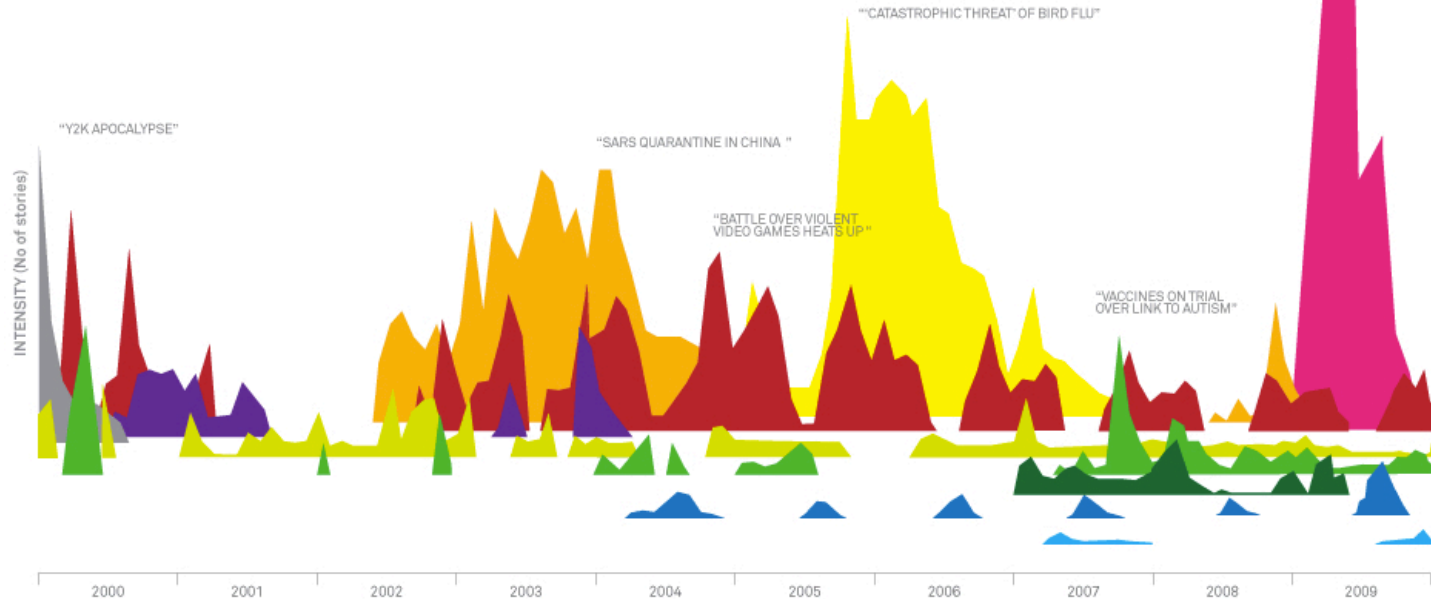
- How do we tell a good graph from a bad one?
- How do we ensure we design a good graph that is fit for purpose?
- Recall goal: enhance understanding and use of **good graphical principles**

# Effective data visualisation is effective visual communication

- Effective graphs...
  - are visually appealing, intuitive, legible
  - use the correct graph type and axis scales
  - use proximity & alignment to facilitate comparison
  - use labels and annotations to add clarity to the message
- Most importantly, effective use of visualisations
  - Enables clear and impactful communication
  - Elevates influence with stakeholders
  - Facilitates informed decision making

# Beautiful but effective?

"BRITAIN PREPARES FOR 85000 DEATHS FROM SWINE FLU"

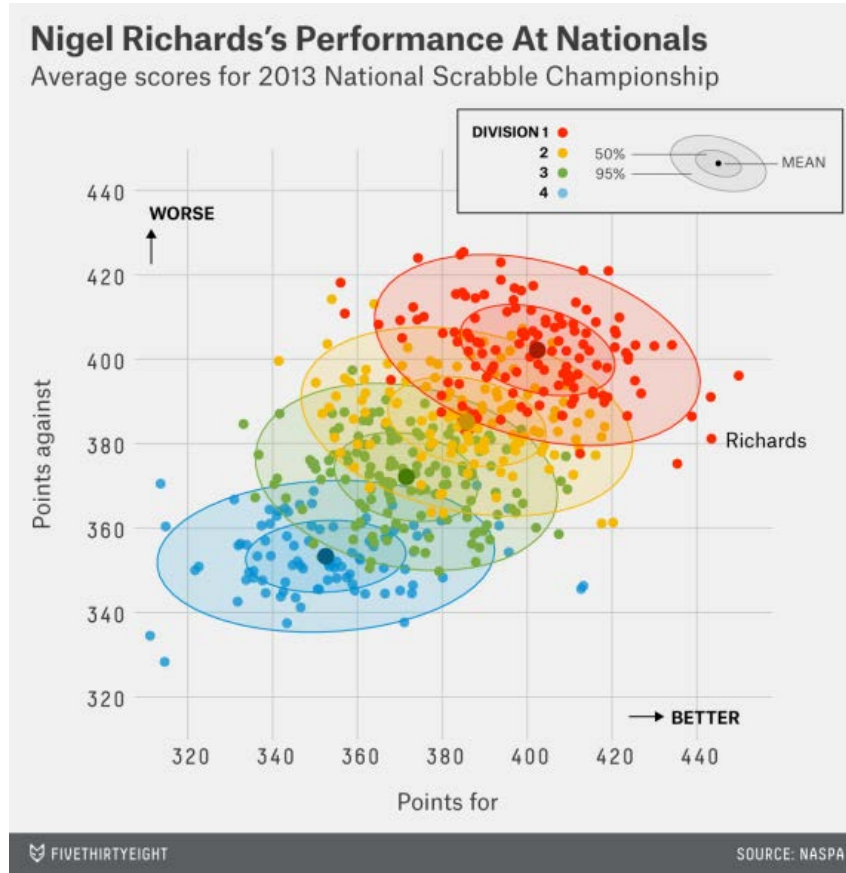


## Story (worldwide deaths)

- |                               |                           |                                 |                   |
|-------------------------------|---------------------------|---------------------------------|-------------------|
| ▲ Killer Wasps (1000)         | ▲ Autism Vaccinations (0) | ▲ Mad Cow Disease (204)         | ▲ Bird Flu (262)  |
| ▲ Killer Wifi (0)             | ▲ Asteroid Collision (0)  | ▲ Violent Video Games (Unknown) | ▲ Swine Flu (702) |
| ▲ Mobile Phones & Tumours (0) | ▲ Millenium Bug (0)       | ▲ SARS (774)                    |                   |

**Mountains Out of Molehills**  
A timeline of global media scare stories.

# Beautiful and effective



# Principles for effective visual communication

## Graphical Principles Cheat Sheet

Authors: Mark Bailille,<sup>1</sup> Alison Margolske,<sup>1</sup> Baldur Magnusson,<sup>1</sup> Andrew Wright,<sup>1</sup> Ruquan You,<sup>2</sup> Ivan-Toma Vranesic,<sup>1</sup> Marc Vandemeulebroecke<sup>1</sup>  
 Affiliations: <sup>1</sup>Novartis Pharma AG, Basel, Switzerland; <sup>2</sup>Novartis Institutes for Biomedical Research, Cambridge, MA, United States; <sup>3</sup>Novartis Institutes for Biomedical Research, Shanghai, China

### 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

**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).

### 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 & placements to create a viewing order

**Focus Points** – primary area of interest that immediately attracts the eye, emphasize the most important concept and make it your focal point. Use dark 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

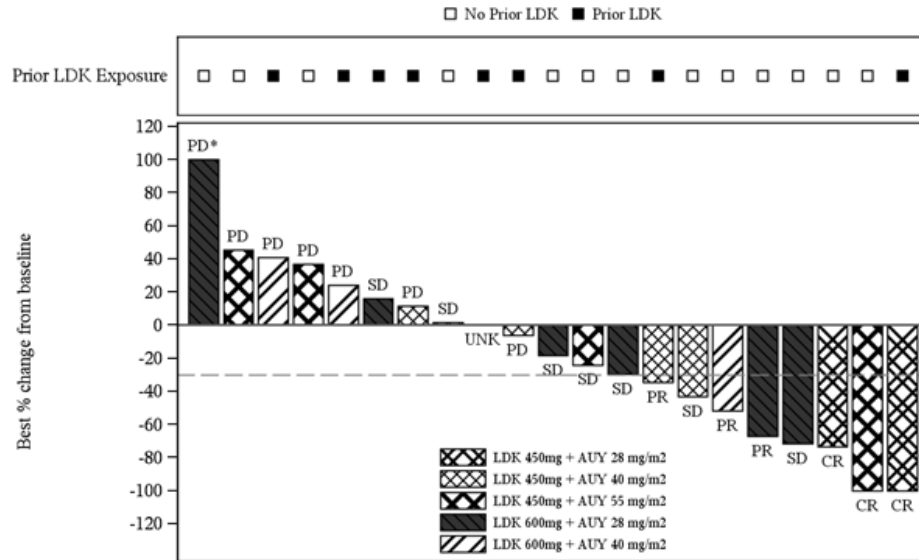
### Effectiveness Ranking

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	Slope or Angle	Length	Position on undivided scale	Position on common scale

# Use the cheat sheet for critical review

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 LDK378 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

### Facilitating Comparisons

**Proximity improves association**

Place labels next to data instead of using legends

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

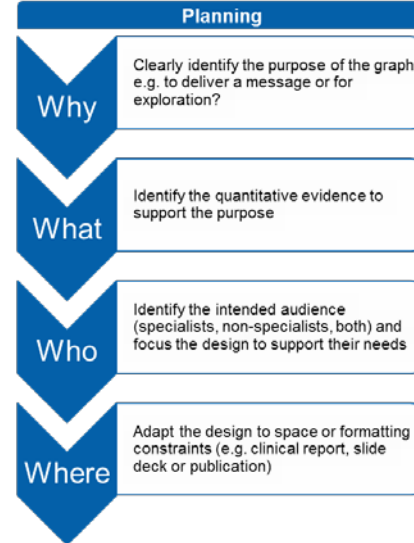
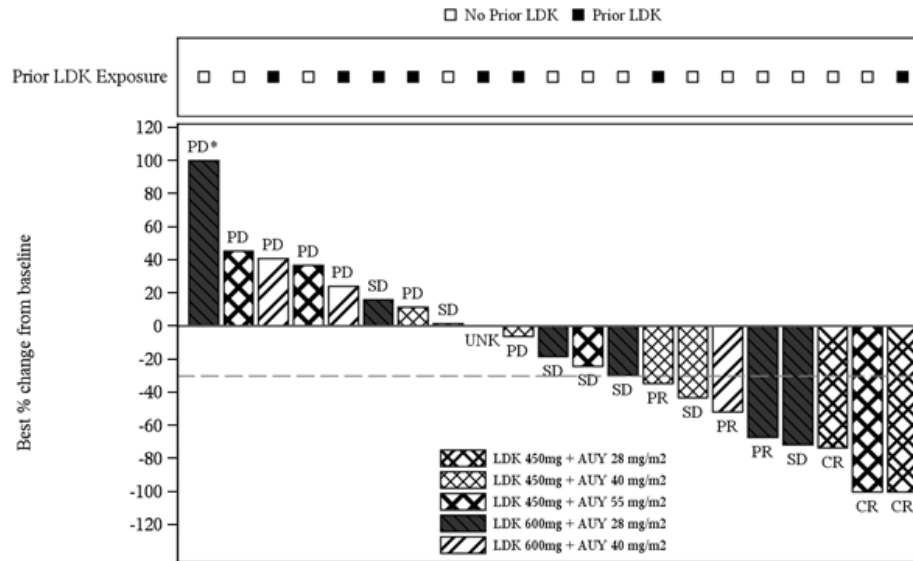
Exception: if comparator is of interest itself

Use reference lines and other visual anchors.



# Use the cheat sheet for critical review

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 LDK378 treatment  
 (Full analysis set)



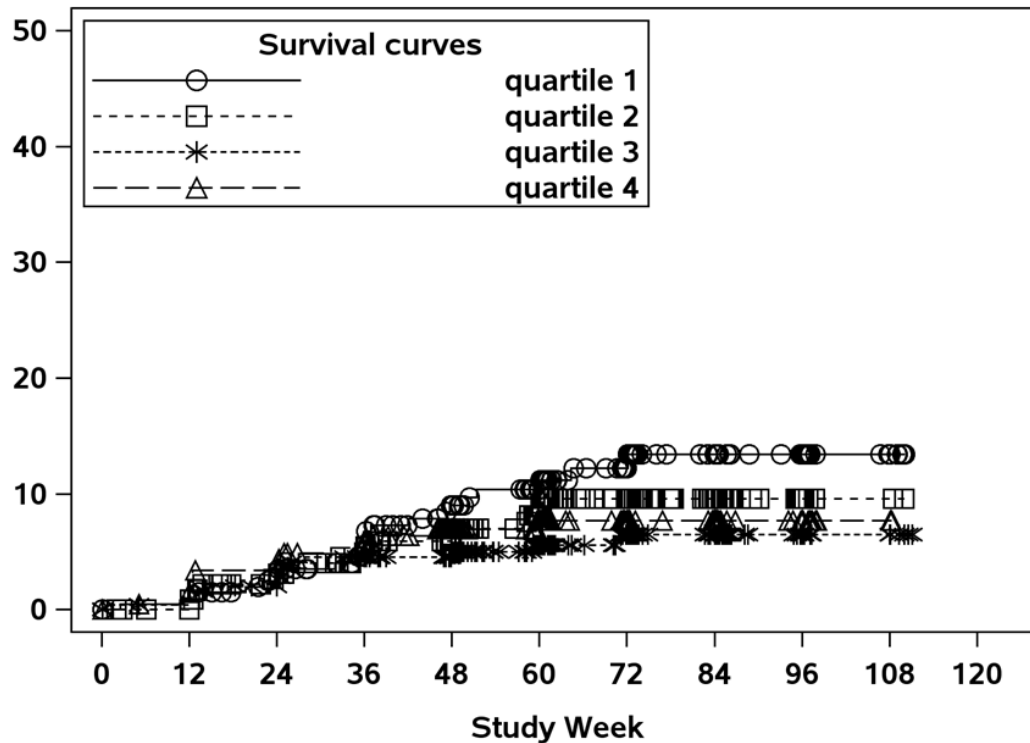
- \* Denotes the percentage change from baseline greater than 100.  
 Source: Table 11-4, Listing 14.2.1.2 and Listing 16.2.4.1.5



# This is a continual process

Planned Treatment:

mg



# Three laws for improving visual communication

## Have a clear purpose

- Know the purpose of creating the graph
- Identify the quantitative evidence to support the purpose
- Identify the audience and focus the design to support their needs

## 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)

## Make the message obvious

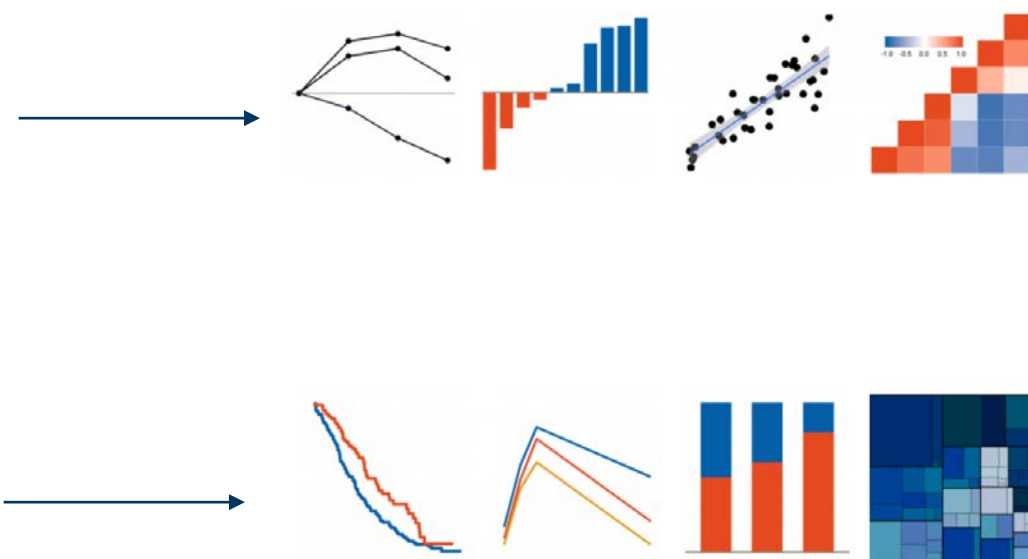
- Use proximity and alignment to aid in comparisons
- Minimize mental arithmetic (e.g. plot the difference)
- Use colors and annotations to highlight important details

<https://arxiv.org/abs/1903.09512>

# Elements of a STRATOS VP initiative

## Topic groups

1	Missing data
2	Selection of variables and functional forms in multivariable analysis
3	Initial data analysis
4	Measurement error and misclassification
5	Study design
6	Evaluating diagnostic tests and prediction models
7	Causal inference
8	Survival analysis
9	High-dimensional data



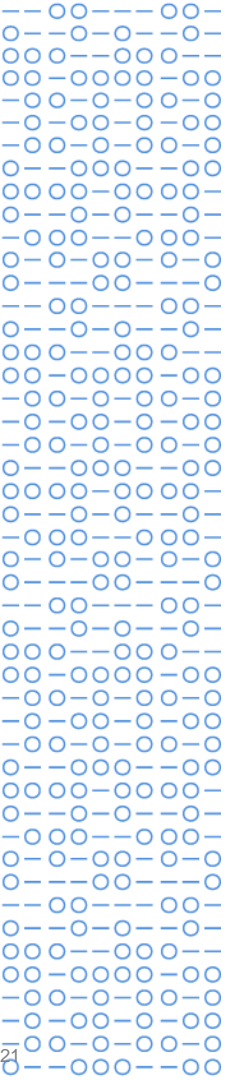
# STRATOS Visualisation panel

“Visualization and the use of graphics can help at every stage of an analysis, from the planning and design of an experiment, the very first data explorations, through to the communication of conclusions and recommendations.

Visualization is more than "plotting data"; it can lead to a deeper understanding and inform next steps.

The role of the STRATOS visualization panel is to promote the use of good graphical principles for effective visual communication, providing guidance and recommendations covering all aspects from the design, implementation and review of statistical graphics.”





Thank you

# Acknowledgements

- Alison Margolskee
- Marc Vandemeulebroecke
- David Carr
- Andrew Wright
- Jürgen Löffler
- Linda Kanitra
- Baldur Magnusson
- Matthew Brierley
- David Ohlssen
- Shaun Butcher
- Julie Jones
- Walter Hufford
- Ruquan You
- Ivan-Toma Vranesic
- Ian Rees
- Nicolas Guerro
- Keo Chanthavinout
- Frank Bretz