AMDS
Global Drug Development

Effective Visual Communication for Quantitative Scientists

Mark Baillie
September 10th, 2021
https://graphicsprinciples.github.io/

NOVARTIS | Reimagining Medicine

STRATOS Visualization panel

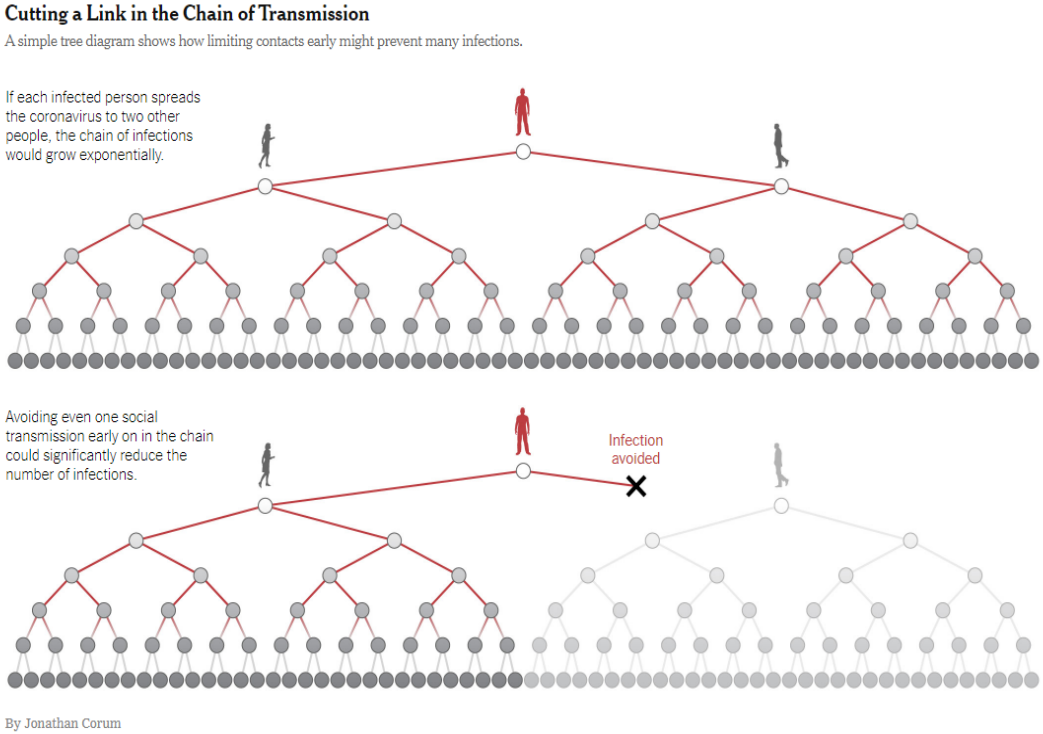
“Visualization and the use of graphics can help at every stage of an analysis, from the planning and design of an experiment, the very first data explorations, through to the communication of conclusions and recommendations. Visualization is more than "plotting data"; it can lead to a deeper understanding 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

STRATOS INITIATIVE

Effective visualisation is important

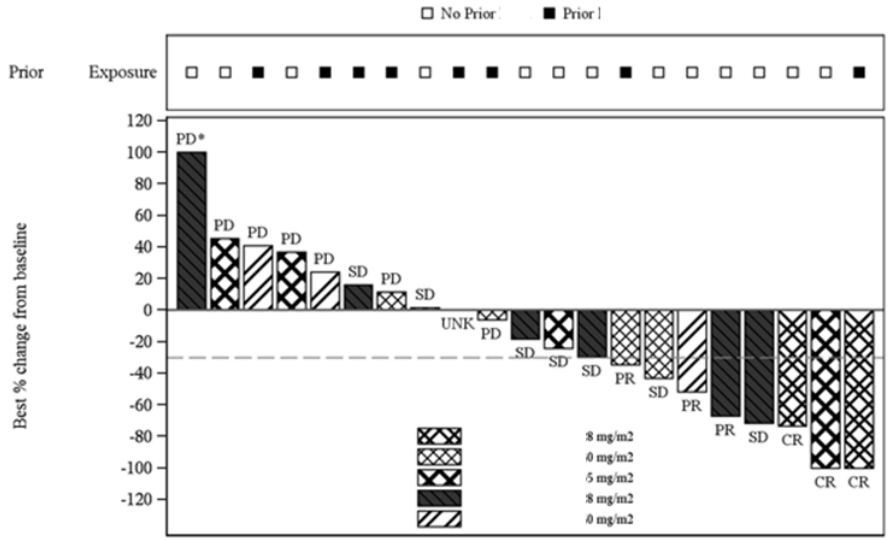


https://www.nytimes.com/2020/03/19/health/coronavirus-distancing-transmission.html

NOVARTIS | Reimagining Medicine

We are not always good at it

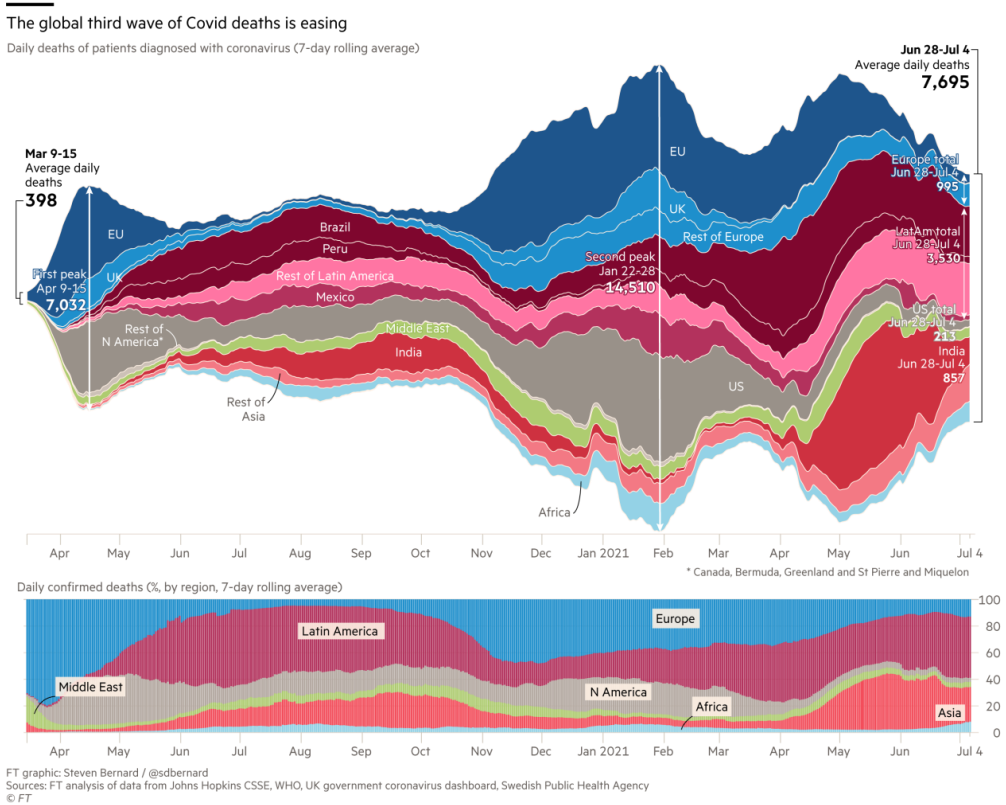
Figure 11-1 (Page 1 of 1)
Best percentage change from baseline in sum of longest diameters and best overall response as per investigator by prior treatment
(Full analysis set)



- * 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

NOVARTIS | Reimagining Medicine

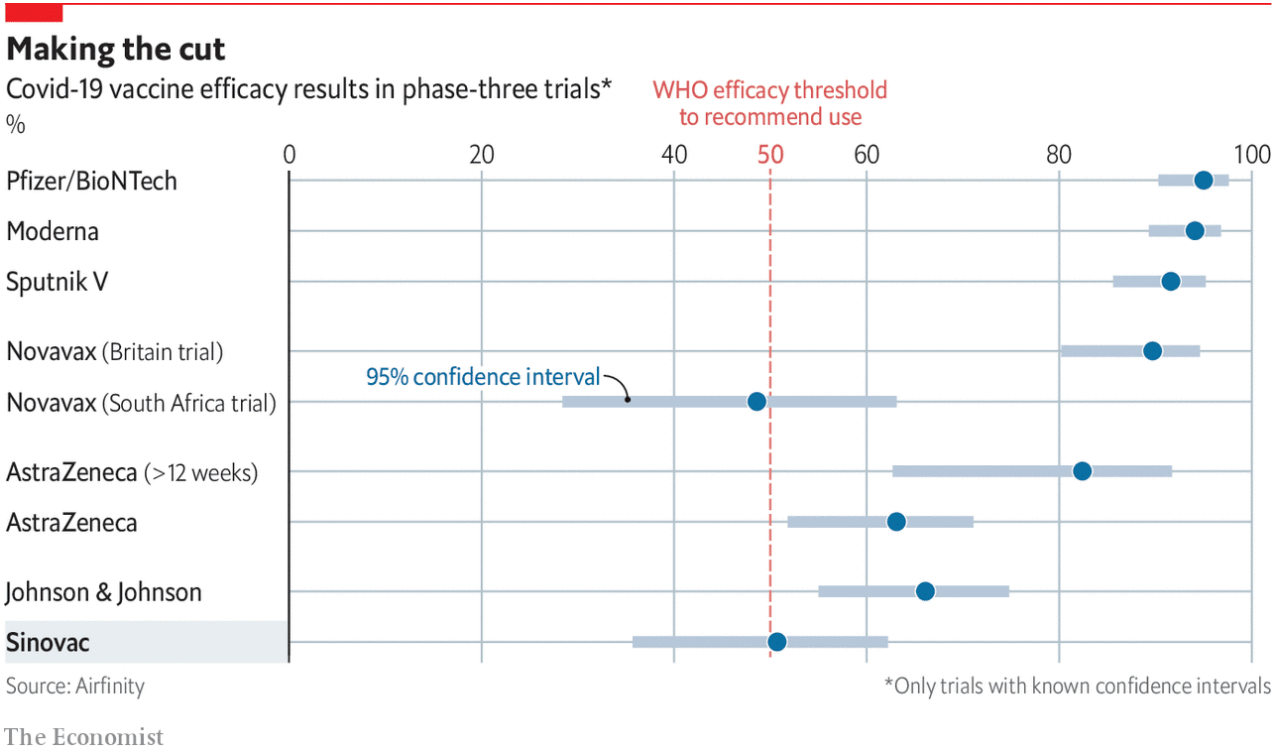
Beautiful but effective?



https://www.ft.com/content/a2901ce8-5eb7-4633-b89c-cbdf5b386938

NOVARTIS | Reimagining Medicine

Beautiful and effective?



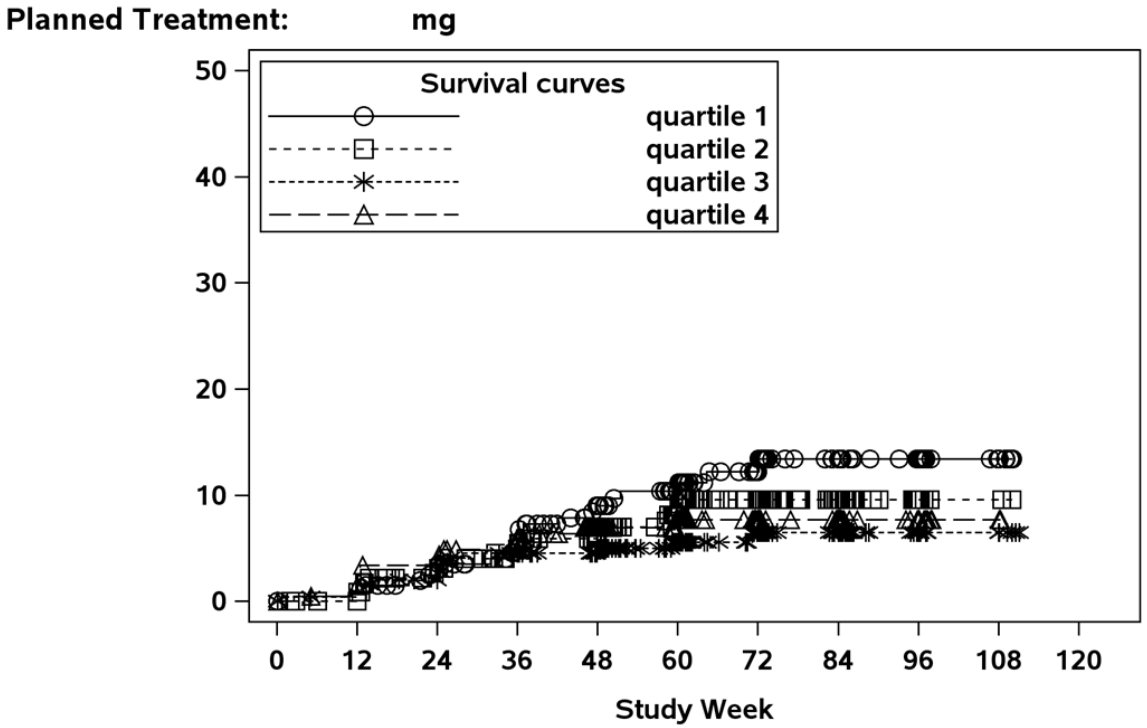
https://www.economist.com/graphic-detail/2021/04/15/in-clinical-and-real-world-trials-chinas-sinovac-underperforms

Principles for effective visual communication

https://graphicsprinciples.github.io/

NOVARTIS | Reimagining Medicine

This is a continual process



NOVARTIS | Reimagining Medicine

Three principles 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

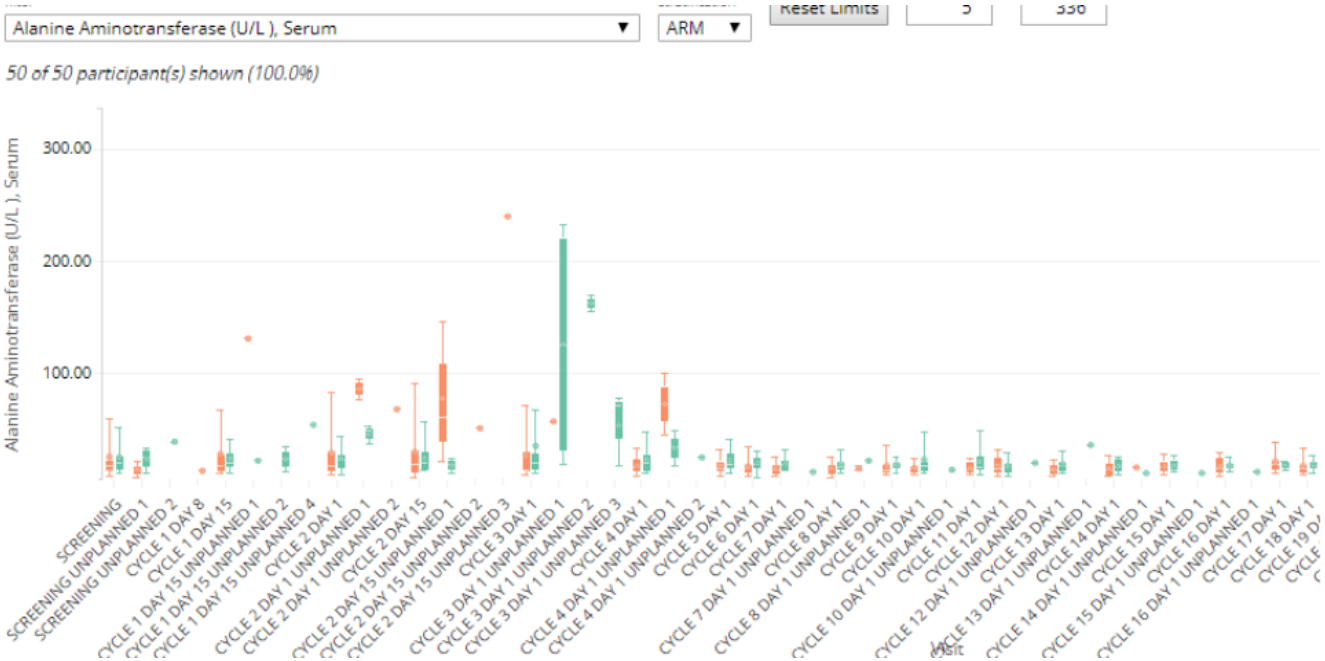
- Choose the appropriate graph type to display your data
- Avoid misrepresentation (use appropriate scales)
- Maximize data to ink ratio (reduce distraction, less is more)

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://ascpt.onlinelibrary.wiley.com/doi/full/10.1002/psp4.12455

This is a continual process...



NOVARTIS | Reimagining Medicine

Law 1
Have a clear purpose

The 4 areas for a clear purpose

What is the purpose of the visualization?

- What is the main objective of the visualization?
- List the (scientific) question(s) the visualization is trying to answer. Try to be specific.
- What is the key evidence that is available to answer the question?

Who is your audience?

- List the primary groups or individuals you will be communicating to.
- If you had to narrow that to a single person, who would that be?
- What does your audience care about?
- What action does your audience need to take?

What is the importance of this project?

- What are the benefits if your audience acts in the way that you want them to?
- What are the risks if they do not?

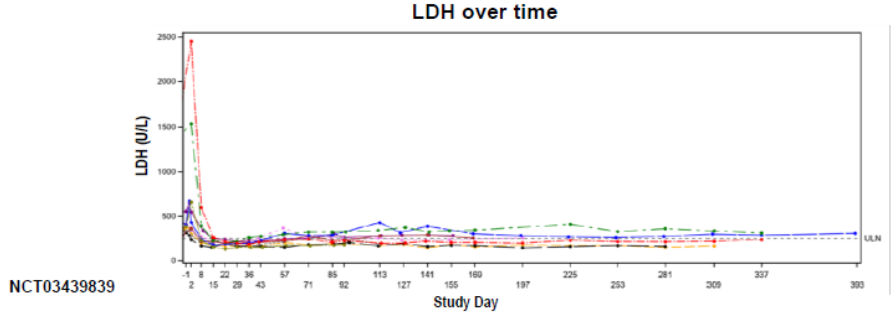
What is the key message (the so what?)

- Write out in a single sentence the key message

Early Phase 2 data support advancing LNP023 as a front-line treatment for PNH

In a Phase 2 PNH trial, LNP023 add-on to eculizumab in patients with hemolysis delivered consistent LDH normalization and transfusion-free hemoglobin increase in all patients

The ongoing LNP023 monotherapy trial in eculizumab-naive PNH patients shows early efficacy (LDH↓)



LNP023 included from the public 2019 R&D day deck
https://www.novartis.com/sites/www.novartis.com/files/2019-12-05-novartis-r-d-day-investor-presentation.pdf

Purpose

What is the purpose of the visualization?

What is the main objective of the visualization?

The visualization is to display supporting evidence that LNP023 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 LNP023 dose as a mono and combo therapy?*
- *Does LNP023 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.

What type of graph do I want to create?

EXPLORATORY

"I want to dig into the data"

"I want to get familiar with the data"

"I want to find the story in my data"

The audience is:
YOU

EXPLANATORY

"I want to communicate the results"

"I want to tell the story behind the data"

The audience is:
SOMEONE ELSE

Credit Andrew Wright, Novartis

Do you want your audience to play ‘Where’s Wally?’

Credit Andrew Wright, Novartis

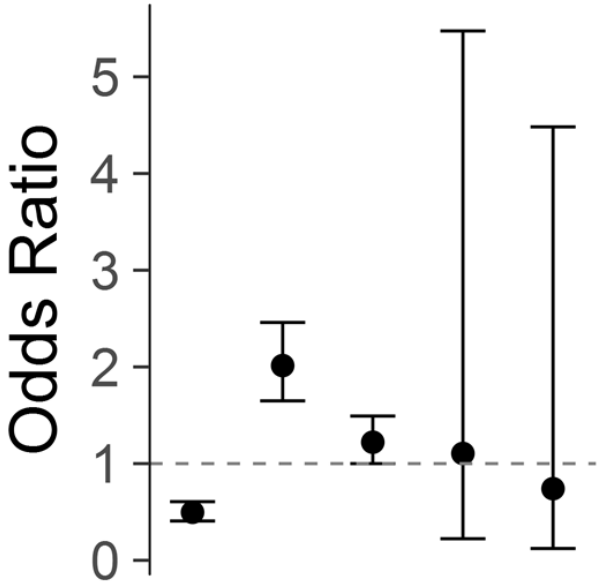
Law 2
Show the data clearly

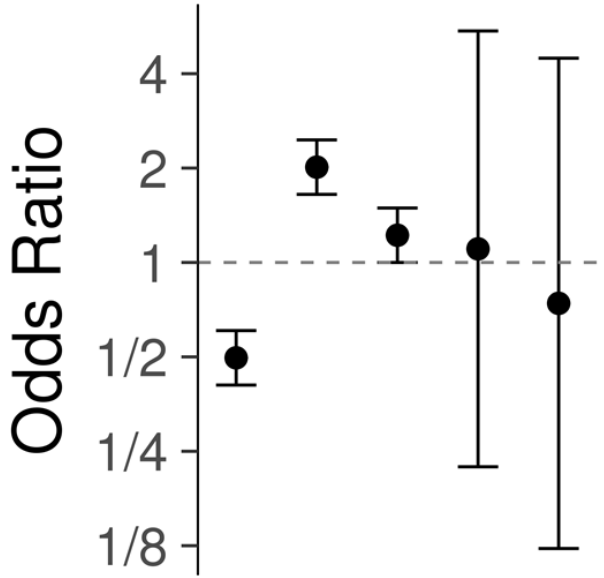
Show the data clearly

https://www.theguardian.com/world/2019/sep/04/trump-hurricane-dorian-alabama-sharpie-map

Choose the right scale for your data

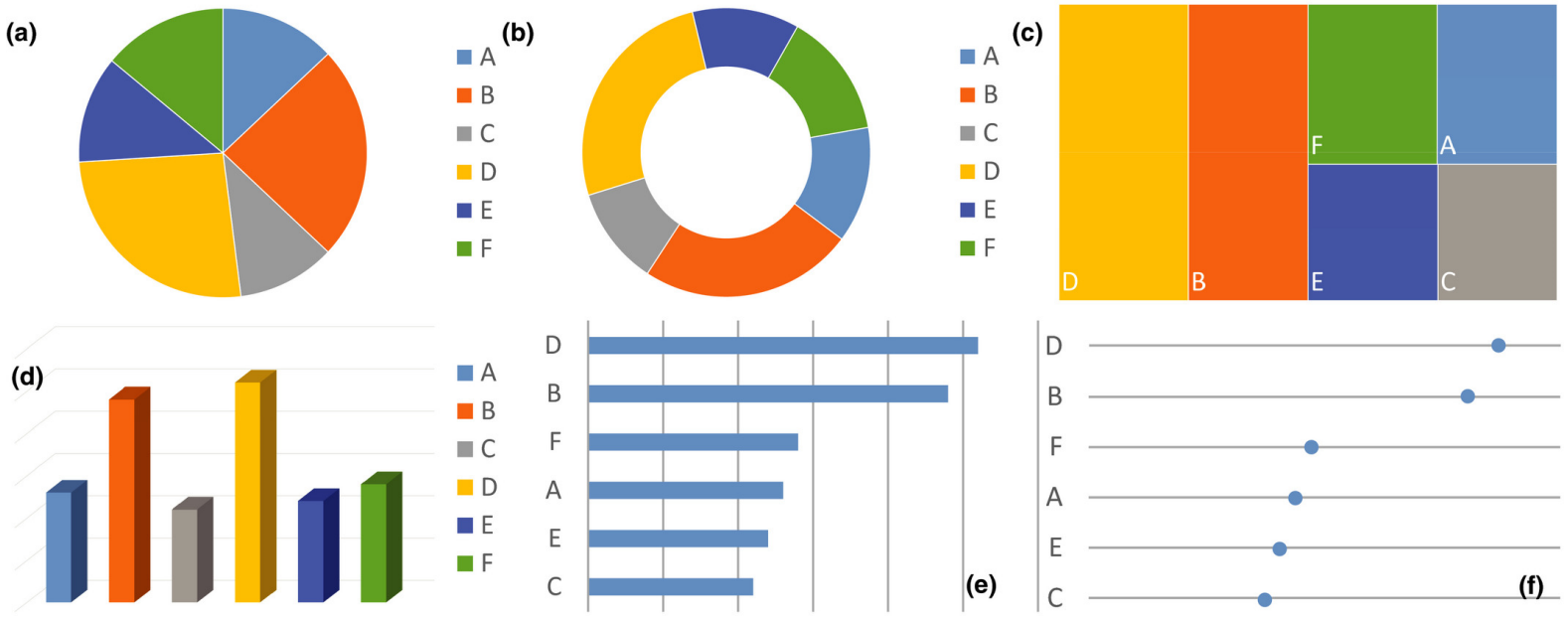
Avoid plotting log-normally distributed variables on a linear scale
(e.g. hazard ratio, AUC, CL)





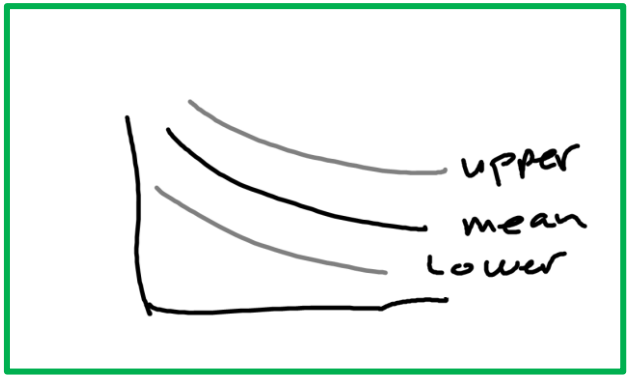
https://graphicsprinciples.github.io/

Choosing the Correct Graph Type Aids in interpretation



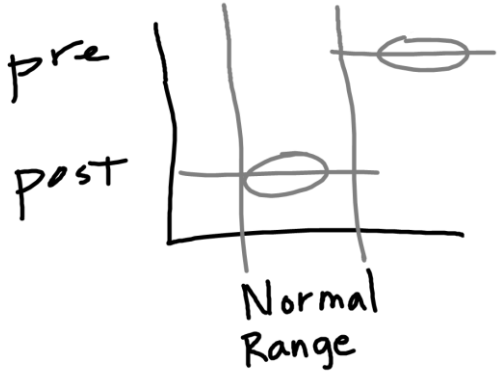
2nd principle - select the appropriate graph

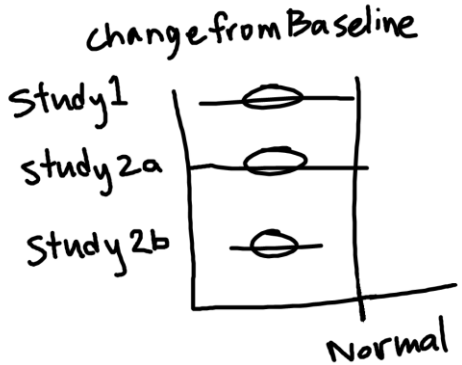
- Come up with several different ways to display the same information
- Display the key evidence in a way that supports the purpose



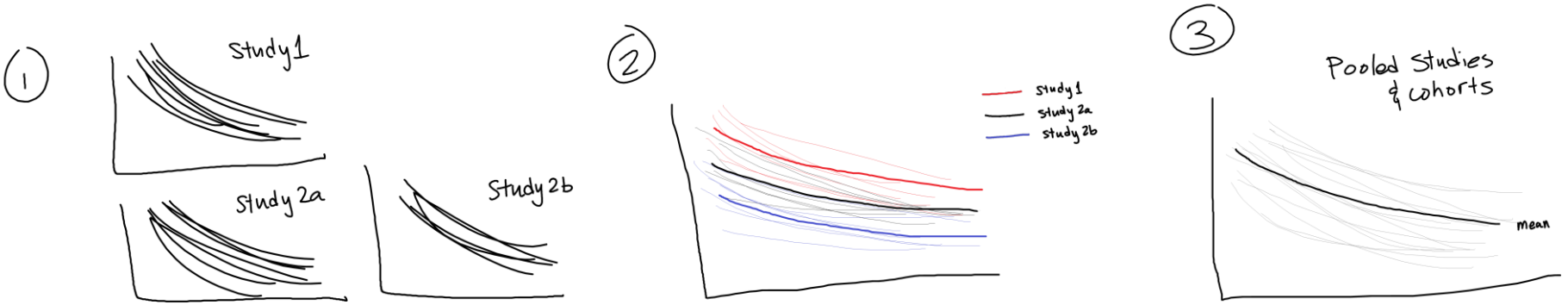
Continuing with the LNP example

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





2nd principle – Iterate and eliminate clutter

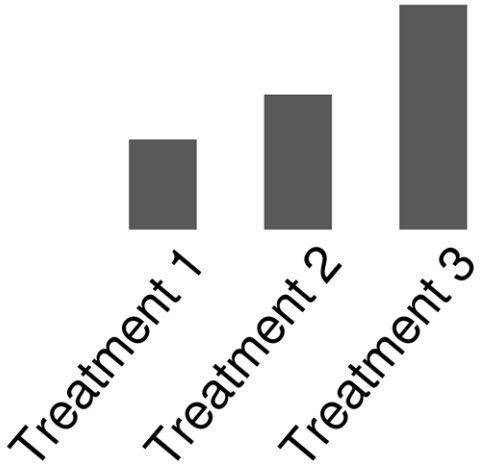


Law 3
Make the message obvious

NOVARTIS | Reimagining Medicine

Try not to set text at an angle

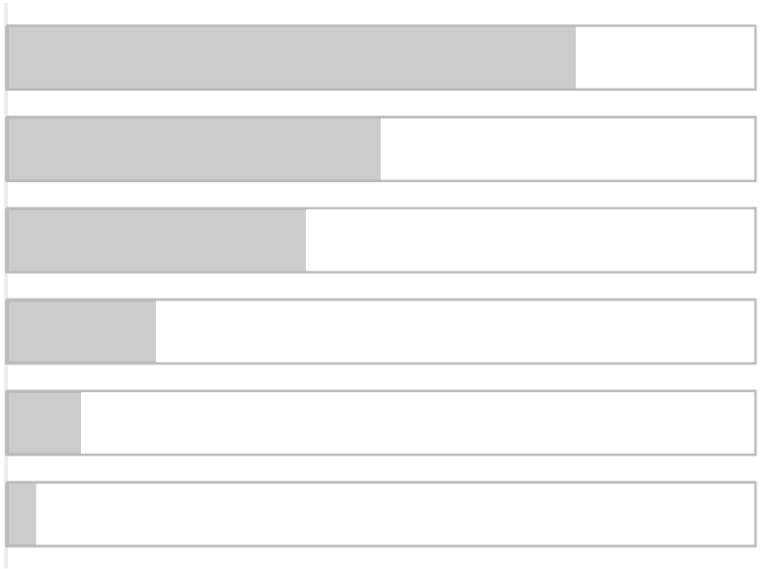
Think of alternatives such as transposing the graph

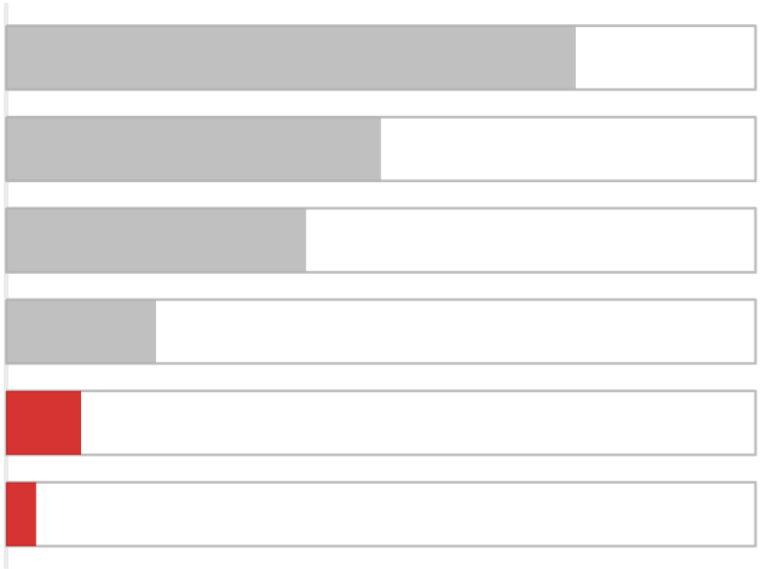


https://graphicsprinciples.github.io/

Only use color when it adds value

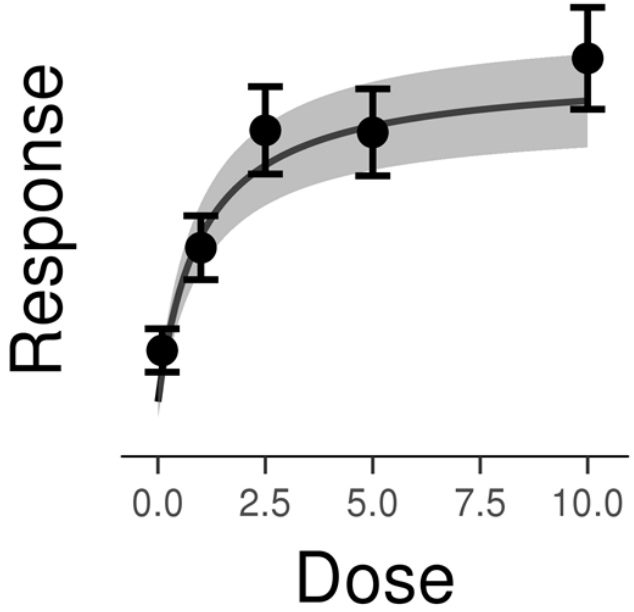
Use a bold, saturated or contrasting color to emphasize important details

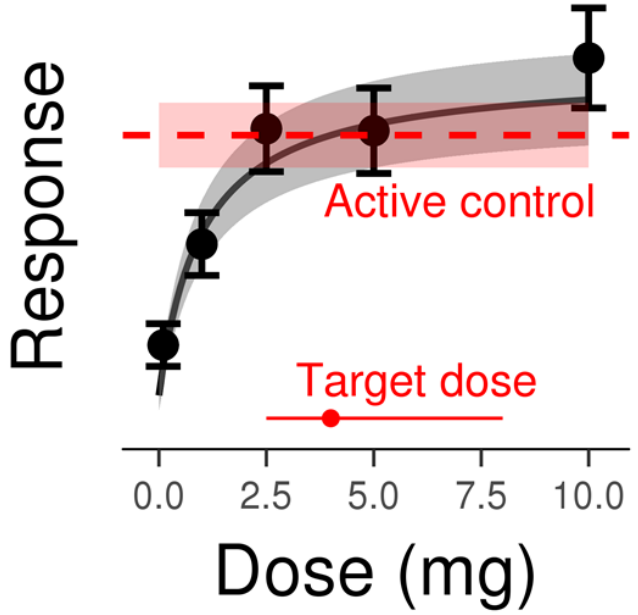




https://graphicsprinciples.github.io/

Use informative labels and annotations to support the message

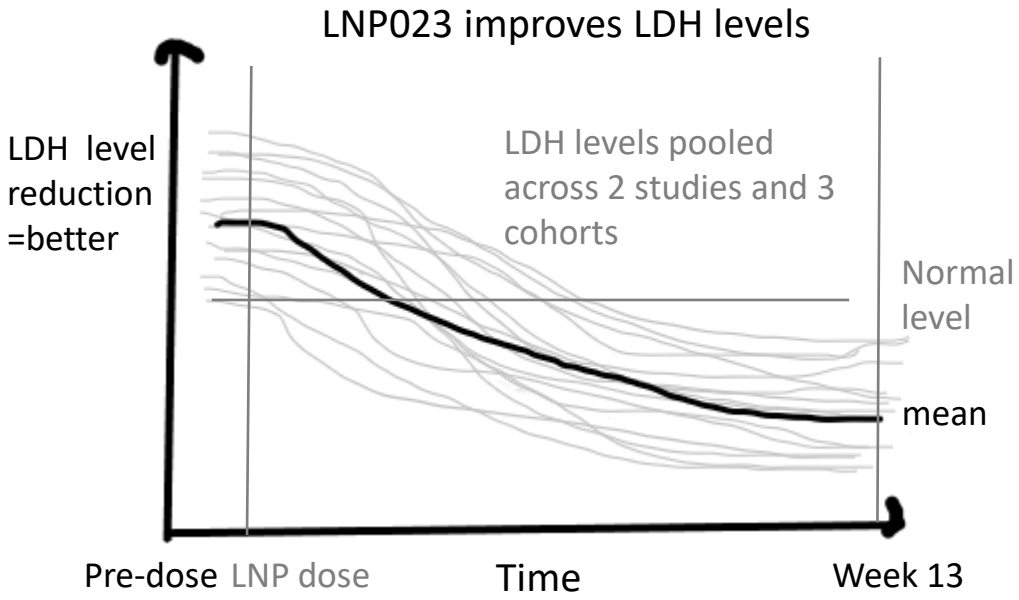




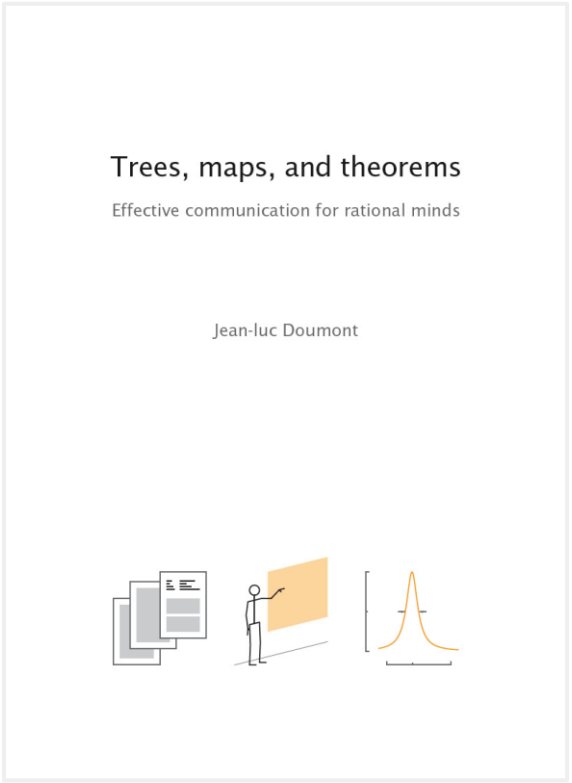
https://graphicsprinciples.github.io/

3rd principle – draw attention

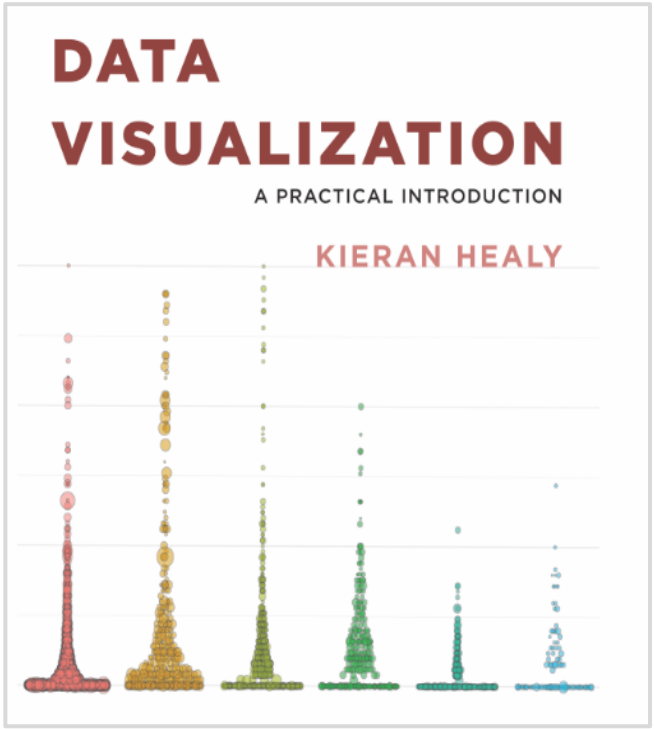
- Draw the viewer's attention to points of interest
- Use arrows, labels, reference lines to drive home the message
- Make sure to have clear axis labels and informative titles



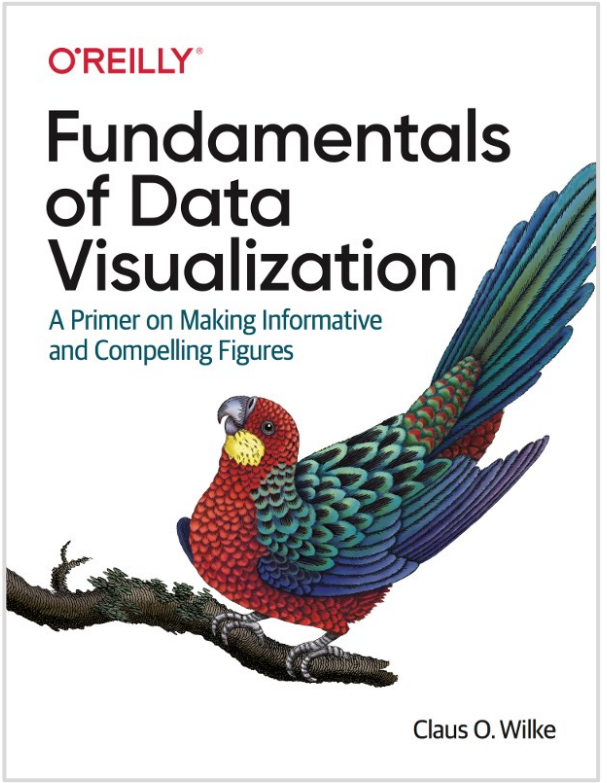
Where to find to out more?



https://www.principiae.be/book/



https://socviz.co/



https://serialmentor.com/dataviz/

Wonderful Wednesdays 10

08-Dec-2020

EFSPI
EUROPEAN FEDERATION OF STATISTICIANS IN THE PHARMACEUTICAL INDUSTRY
Representing Statistical Associations in Europe

An EFSPI/PSI VIS SIG initiative

NOVARTIS | Reimagining Medicine

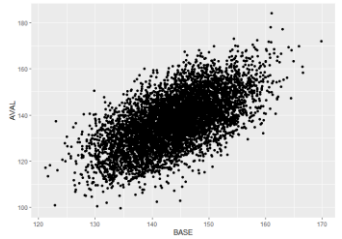
Meta-analysis example data set

- The example simulated data set is based on **seven** phase III studies in Hypertension.
- A wide collection of baseline measurements are also included which can be explored to understand the patient populations within each trial, to search for potential subgroups or differential treatment effects, or even to develop prognostic or predictive risk models.
- For a detailed overview of the data set, please refer to the data dictionary provided:

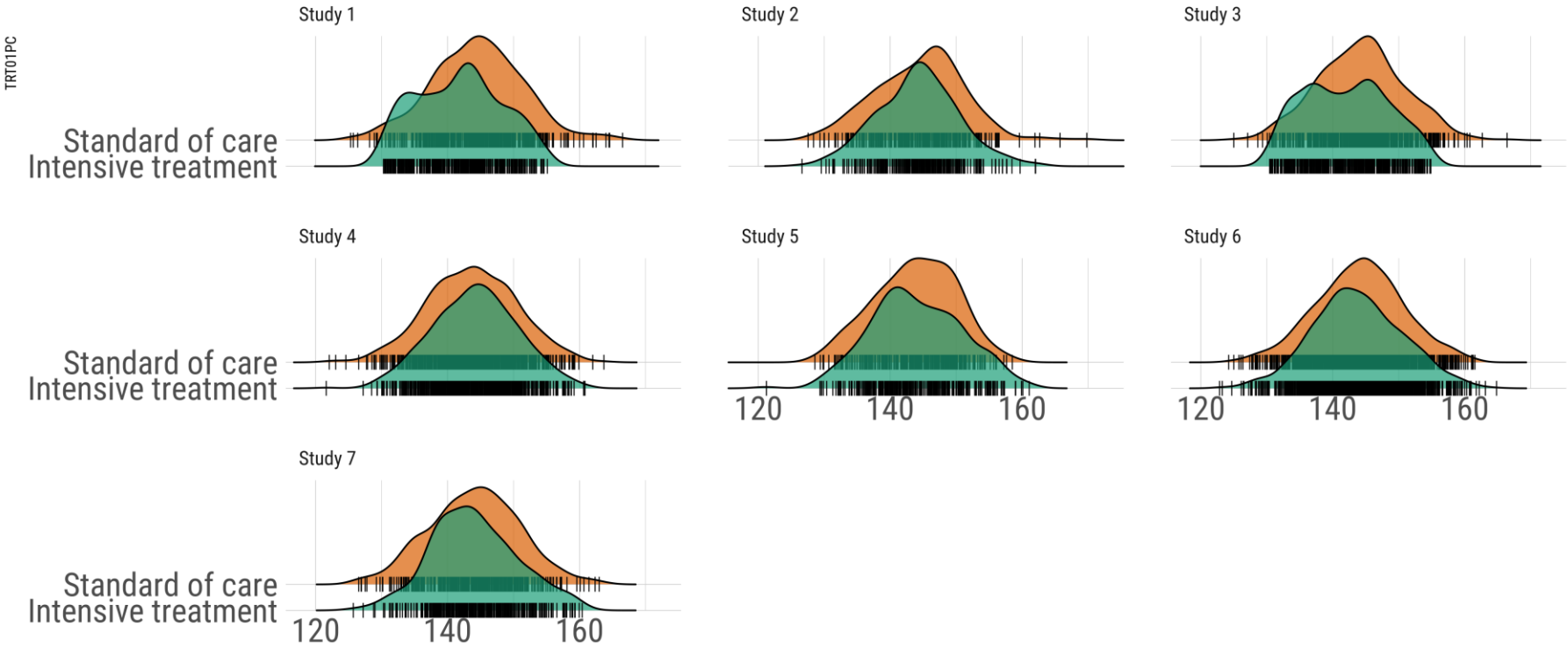
https://github.com/VIS-SIG/Wonderful-Wednesdays/tree/master/data/2020/2020-11-11

AGE	Age (years)
AGECAT1	Age Group 3: 75 years and older = TRUE
AGECAT1C	Age Group 3: 75 years and older = TRUE
AGECAT1N	Age Group 3: 75 years and older = TRUE
ALBSI	Albumin (g/L)
BASE	Mean systolic blood pressure (mm Hg) measured at baseline
BASOSI	Basophils (Absolute) (10E9/L)
BICARSI	Bicarbonate (mmol/L)
BILISI	Bilirubin (umol/L)
BMI	BMI
BUNSI	Blood Urea Nitrogen (mmol/L)
CASI	Calcium (mmol/L)
CHD10R1	10-year Coronary heart disease (CHD) risk category (High (>20%) , Me
CHD10R1N	10-year CHD risk category (Numeric, 1 = Low, 2 = Medium, 3 = High)
CHOL_HDL	Ratio of Total Cholesterol / HDL
CHOLSI	Cholesterol (mmol/L)
COUNTRY	Country indicator
CREATSI	Creatinine (umol/L)
EOSLESI	Eosinophils/Leukocytes (%)
EOSSI	Eosinophils (Absolute) (10E9/L)
ETHNIC	Ethnicity
GGTSI	Gamma Glutamyl Transferase (U/L)
GLUCPSI	Glucose, Plasma, Fasting (mmol/L)

GGTSI	Gamma Glutamyl Transferase (U/L)
GLUCPSI	Glucose, Plasma, Fasting (mmol/L)
GREGGR1O	Regional stratification group
HCT	Hematocrit
HDLSI	HDL Cholesterol (mmol/L)
HDT	Phosphate (mmol/L)
HEIGHT	Height (cm)
HGBSI	Hemoglobin (g/L)
KSI	Potassium (mmol/L)
LDLSI	LDL Cholesterol (Assayed) (mmol/L)
LPASI	Lipoprotein-A Protein (g/L)
LYMLESI	Lymphocytes/Leukocytes (%)
LYMSI	Lymphocytes (Absolute) (10E9/L)
MONOLSI	Monocytes/Leukocytes (%)
RACE	Race
SBPCAT1C	Mean systolic blood pressure (mm Hg) at baseline (Category)
SBPCAT1N	Mean systolic blood pressure (mm Hg) at baseline (Numeric)
SEX	Sex
TRIGFSI	Triglycerides (Fasting) (mmol/L)
URATESI	Uric Acid (umol/L)
WBCSI	Leukocytes (10E9/L)
WEIGHT	Weight (kg)

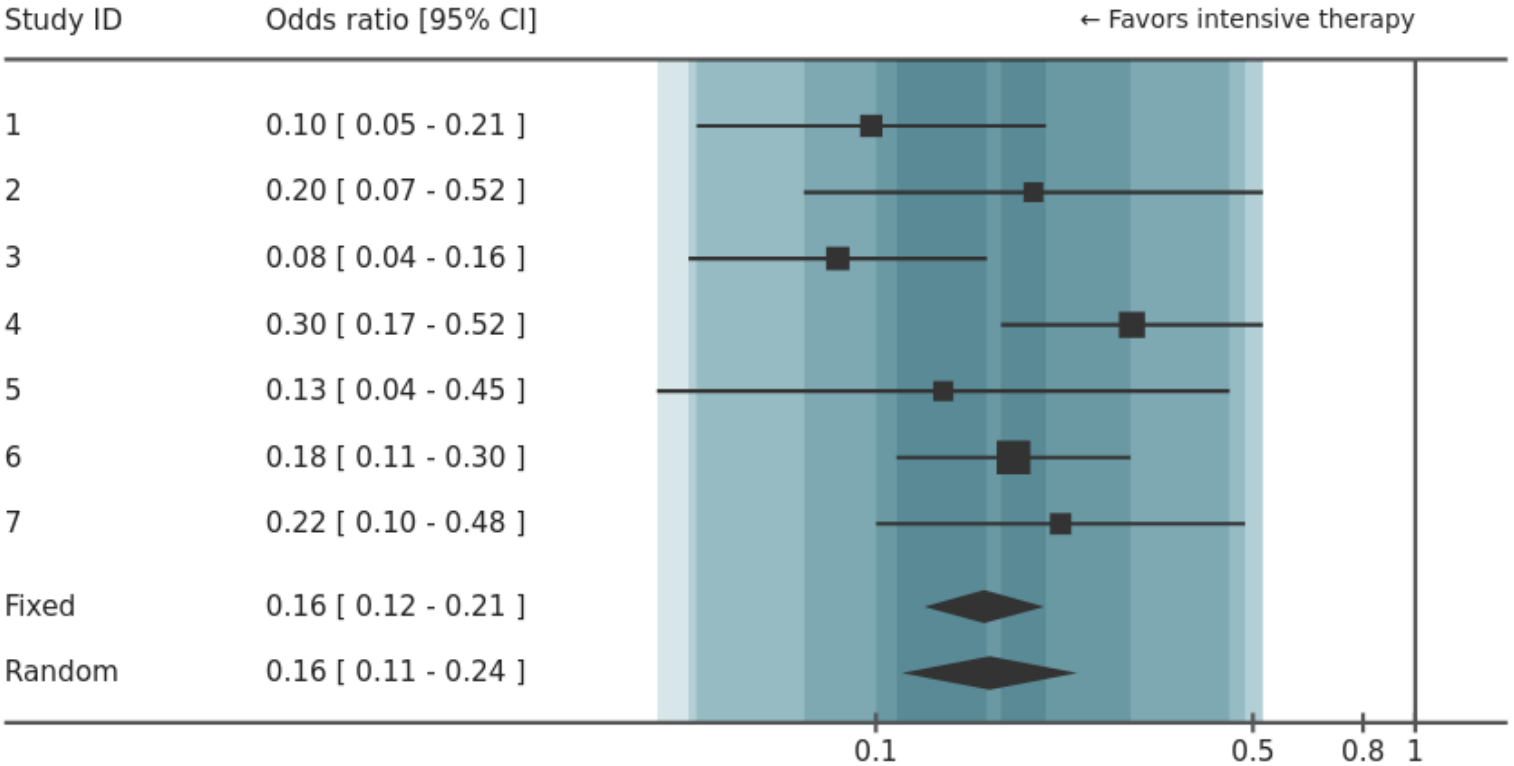


Comparison of mean systolic blood pressure measured at baseline by study

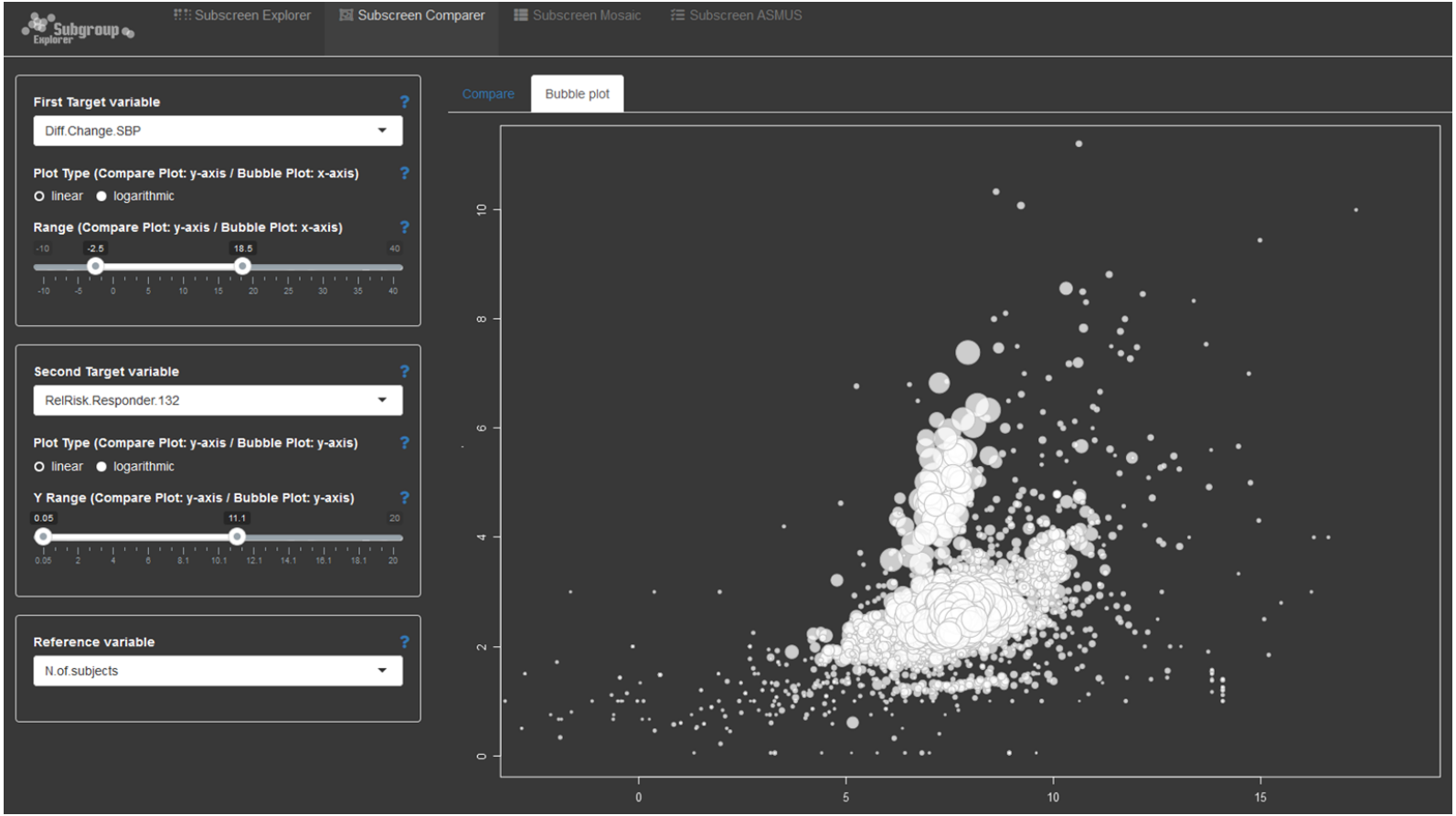


Intensive antihypertensive therapy versus standard of care

Responder analysis - patients with controlled systolic blood pressure at 1 year
(≤ 120 mmHg)

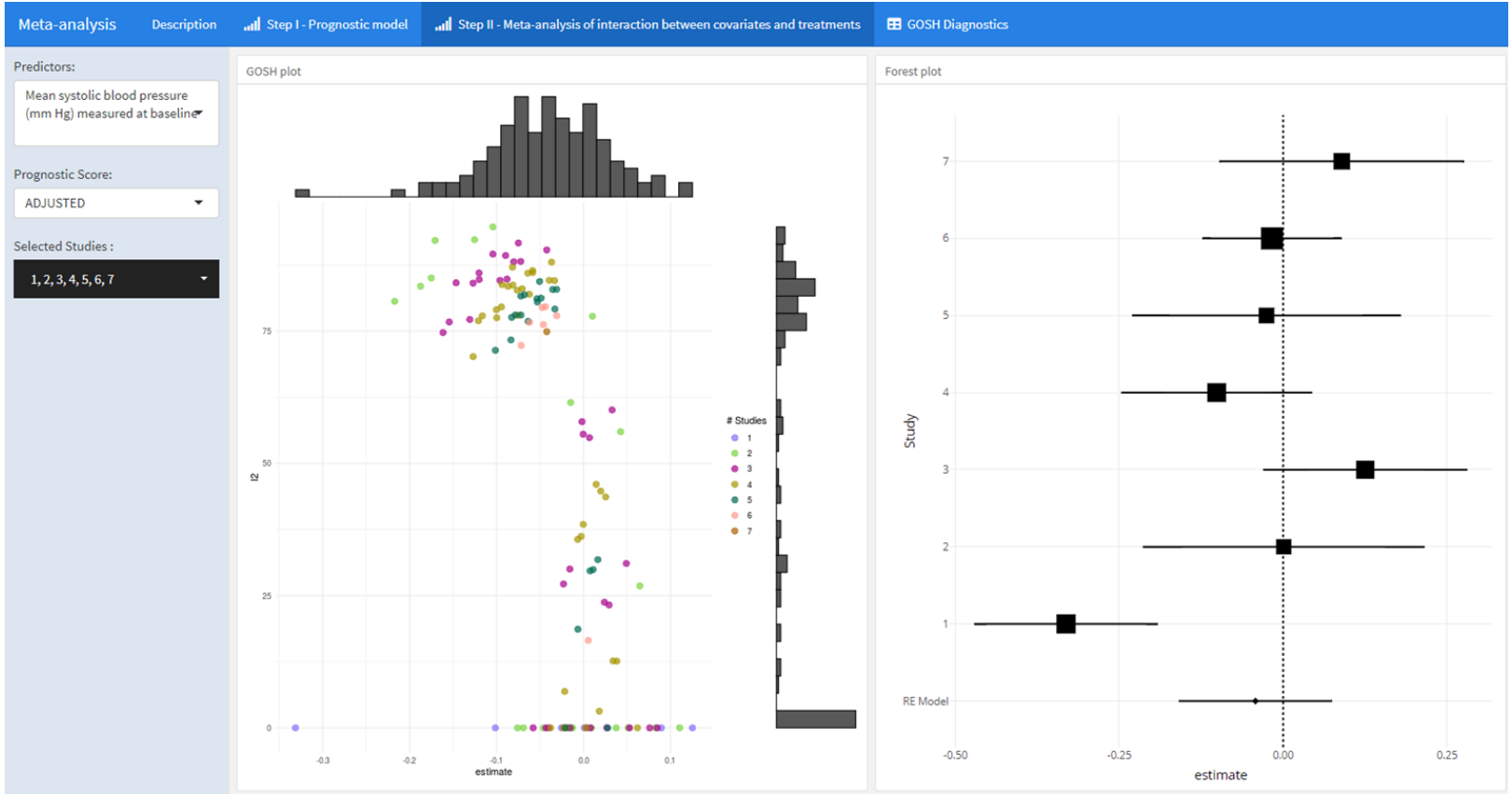


https://waseem-medhat.shinyapps.io/forest_plot_with_bands/



NOVARTIS | Reimagining Medicine

https://cran.r-project.org/web/packages/subscreen/index.html



https://figplot.shinyapps.io/WW20201209/

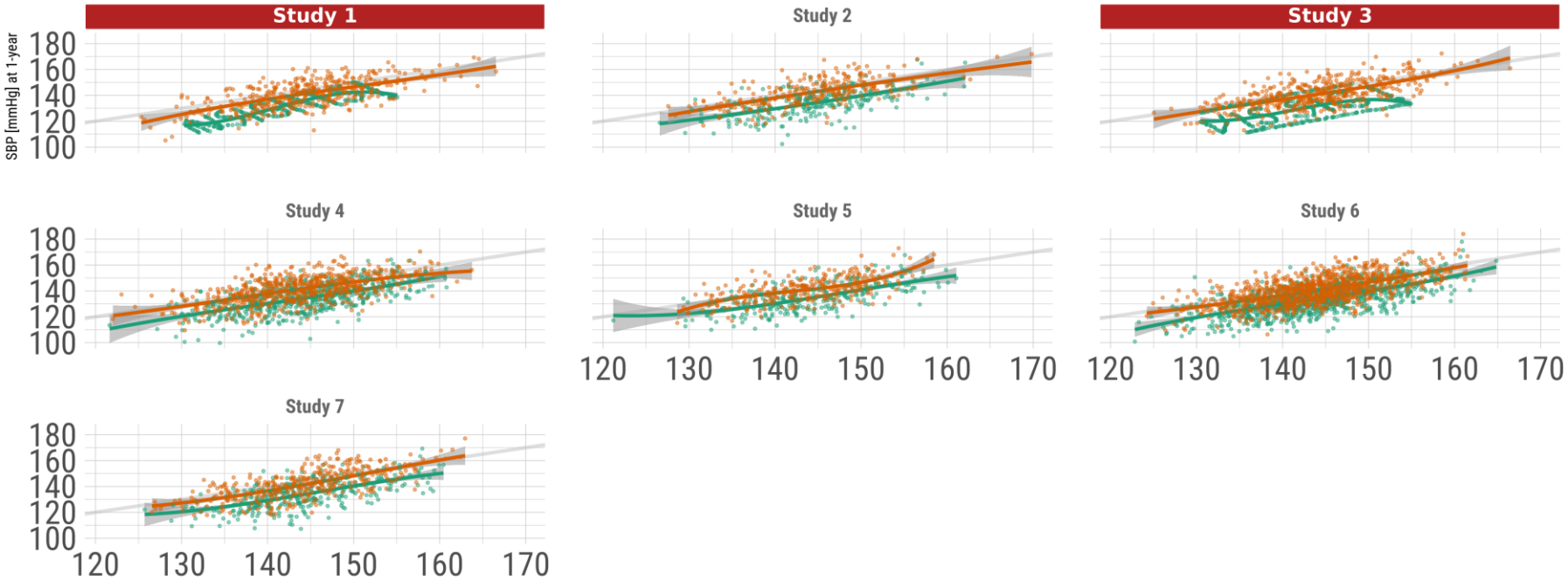
NOVARTIS | Reimagining Medicine

Meta-analysis example data set

- How data visualisation can be deployed to understand integrated data?
- **Key issues where data visualisation can help are around the investigation of whether studies can be combined due to study heterogeneity**
- This throws up questions such as:
  - What graphical tools can be used to assess heterogeneity?
  - What variables are prognostic or predictive of outcome?
  - Where can graphical methods provide general recommendations?

Comparison of pre-post mean systolic blood pressure (SBP) measured at baseline and 1-year

Study 1 and 3 may have data quality issues - further investigation required

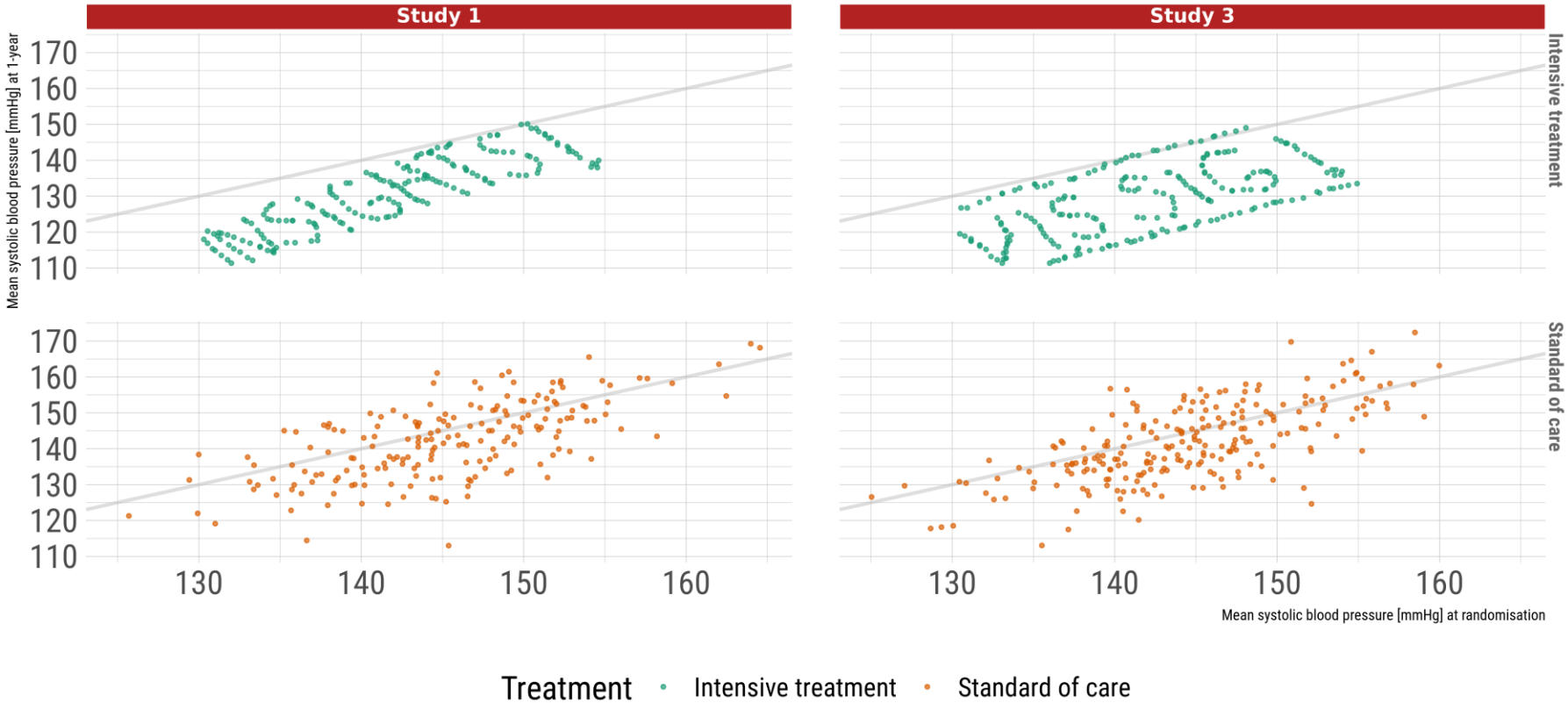


Treatment — Intensive treatment — Standard of care

The by-treatment relationship also disaplyed using a cubic splines.
y = x reference line also displayed.
Data: BIG_DATA_PSI_WW_DEC2020.csv

The intensive treatment arm for study 1 and 3 displayed patterns of interest

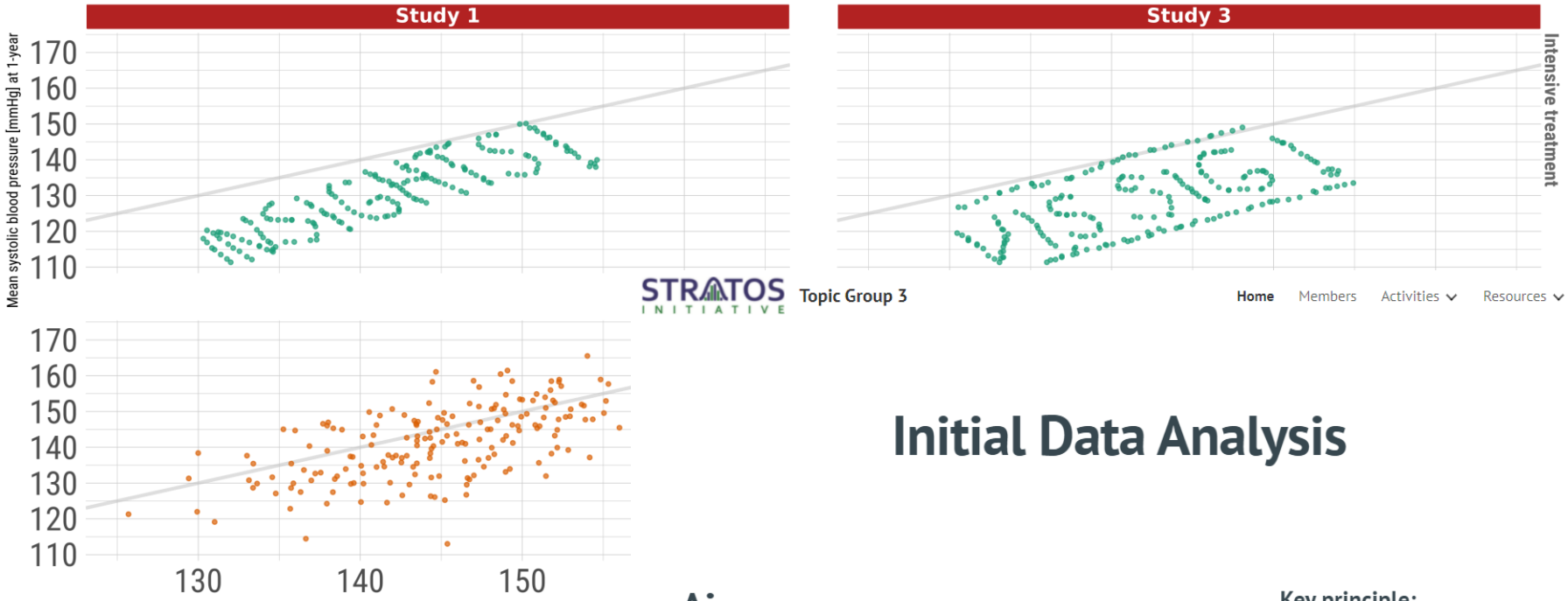
It is always important to plot data many ways



http://robertgrantstats.co.uk/drawmydata.html
Data: BIG_DATA_PSI_WW_DEC2020.csv

The intensive treatment arm for study 1 and 3 displayed patterns of interest

It is always important to plot data many ways



STRATOS INITIATIVE Topic Group 3

Home Members Activities Resources

Initial Data Analysis

Aim

To improve awareness of Initial Data Analysis (IDA) as an important part of the research process and to provide guidance on conducting IDA in a systematic and reproducible manner.

Key principle:

IDA should not touch the research question.

Our group promotes initial data analysis (IDA) as a highly structured step in the data analysis process. For this purpose, we developed a framework for IDA and are creating tools to facilitate the IDA process.

Effective data visualisation is effective communication

Effective visualisations

- enable clear and impactful communication,
- elevate influence with stakeholders,
- facilitate informed decision making.

To help design effective visualisations, remember three principles:

- Purpose,
- Clarity
- Message

More here

- http://www.stratos-initiative.org
- https://graphicsprinciples.github.io/

STRATOS INITIATIVE

http://www.stratos-initiative.org

Thank you

NOVARTIS | Reimagining Medicine