AMDS/Analytics
Global Drug Development

STRATOS visualisation panel: virtual meeting update

Mark Baillie
November 9th, 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.”

Chair: Mark Baillie
Members: Tim Morris, Rachel Phillips, Willi Sauerbrei, Svenja Seide, Marc Vandemeulebroecke

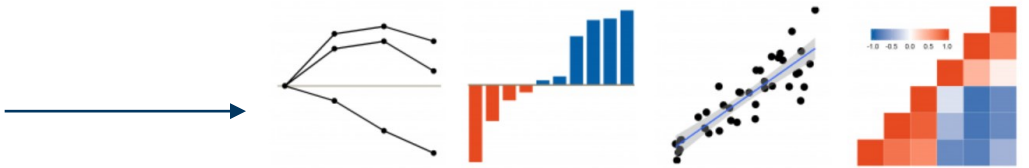
http://www.stratos-initiative.org

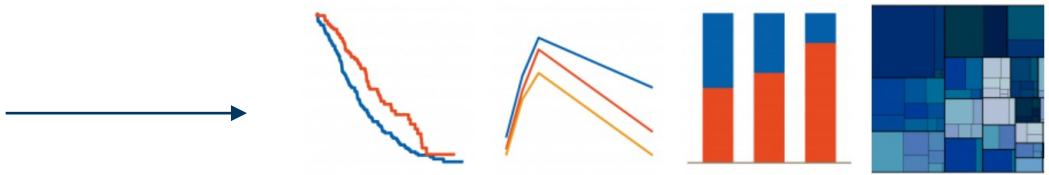
STRATOS INITIATIVE

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

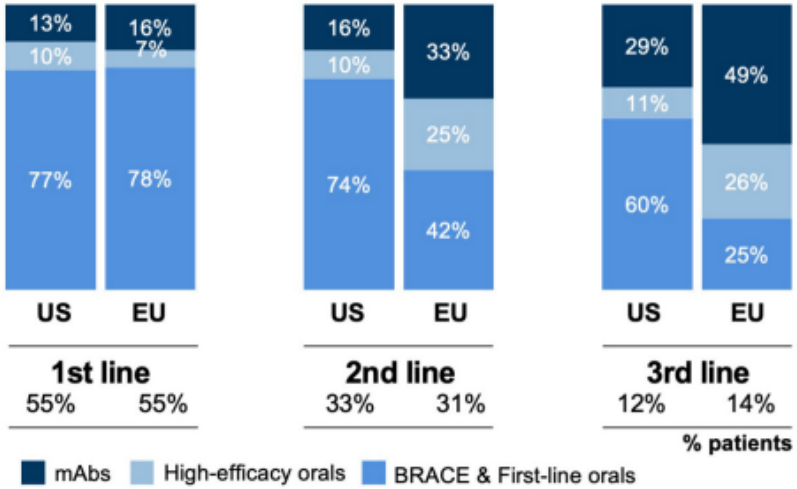




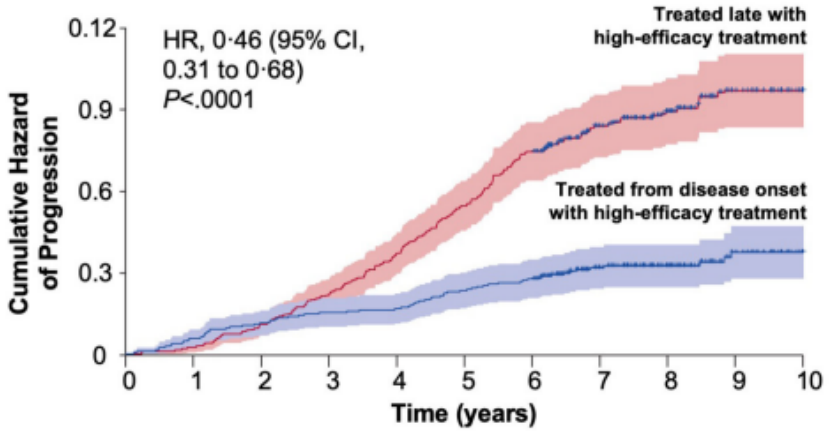
BRACE and first-line orals commonly used in early stages

Despite data showing high-efficacy treatments started early result in better outcomes

Use of disease-modifying treatments in MS[1]



Cumulative hazard of CDP in patients with RRMS treated from disease onset versus late with high-efficacy treatment[2]



MAbs: Ocrevus®, Lemtrada®, Tysabri®; **High-efficacy orals:** Gilenya®, Mayzent®, Mavenclad®; **BRACE & First-line orals:** Interferons, Copaxone®, GA Gx; Tecfidera®, Aubagio® *High efficacy DMTs may include orals and MAbs

1. Symphony APLD (Sep 2018-Aug 2019). , EU5 IPSOS Monitor 2019. 2. He A et al. Lancet Neurol. 2020;19(4):307–316. Retrospective analysis, measured from disease onset. Bold lines are cumulative hazard estimates and shaded areas are 95% CIs. CDP, confirmed disability progression; CI, confidence interval; HR, hazard ratio; RMS, relapsing multiple sclerosis.

NOVARTIS | Reimagining Medicine

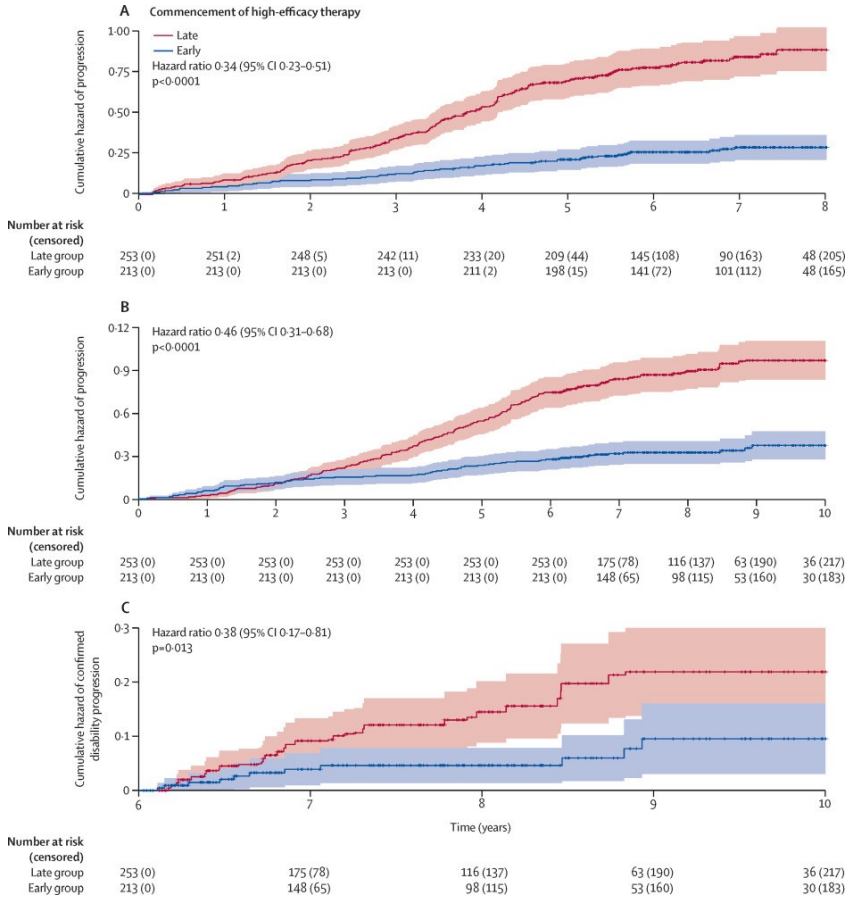
THE LANCET
Neurology
Volume 19, Issue 4, April 2020, Pages 307-316

Articles

Timing of high-efficacy therapy for multiple sclerosis: a retrospective observational cohort study

Anna He MBBS [a], Bernd Merkel PhD [a], James William L Brown PhD [a, b, c], Lana Zhovits Ryerson MD [d], Ilya Kister MD [d], Charles B Malpas PhD [a], Sifat Sharmin PhD [a], Dana Horakova MD [e, f], Eva Kubala Havrdova PhD [e, f], Tim Spelman PhD [g], Guillermo Izquierdo MD [i], Sara Eichau MD [i], Maria Trojano MD [j], Alessandra Lugaresi MD [k, l], Raymond Hupperts MD [m], Patrizia Sola MD [n], Diana Ferraro MD [n], Jan Lycke PhD [o], Francois Grand'Maison MD [p], Alexandre Prat MD [q, r], Marc Girard MD [q, r], Pierre Duquette MD [q, r], Catherine Larochelle MD [q, r], Anders Svenningsson MD [h], Thor Petersen MD [s], Pierre Grammond MD [t], Franco Granella MD [u], Vincent Van Pesch MD [v, w], Roberto Bergamaschi MD [x], Christopher McGuigan MD [y, z], Alasdair Coles PhD [b], Jan Hillert MD [g], Fredrik Piehl MD [g], Helmut Butzkueven PhD [aa, ab], Tomas Kalincik PhD [a, ac] ✉ on behalf of the MSBase study group [†]

https://www.sciencedirect.com/science/article/pii/S1474442220300673?via%3Dihub#fig3



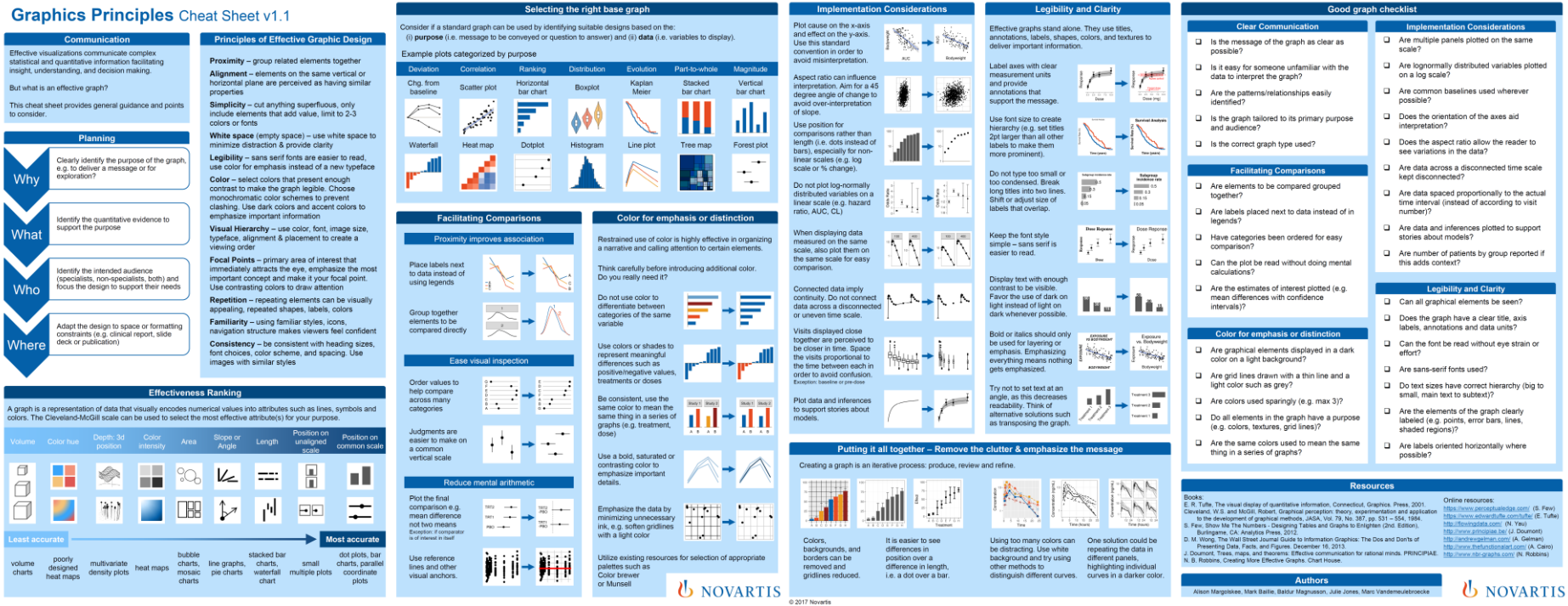
NOVARTIS | Reimagining Medicine

Visualisation Panel Updates

- Biometric Bulletin article 2020
- Support publication panel - provide review and support of STRATOS publications
- Presentation ROeS Sept 2021
- Collaboration with TG2-TG3
  - Regression without regrets (cross-sectional)
  - Leads: G. Heinze, M. Baillie, M. Huebner, a TG2-TG3 collaboration
- Workshops - M. Baillie and M. Vandemeulebroecke.
  - Developed hands on workshop and rolled out to 1,000 statisticians, clinicians and other quantitative scientists at Novartis.
  - Active learning – pen, paper and gray matter
  - Designed to be face to face, 3 hours, but now redeveloped to run virtually
  - Early 2021 invited presentation to ASCPT Pharmacometrics & Pharmacokinetics
  - 2022 Will present a version of the workshop at IBC
    - https://members.biometricsociety.org/ibc2022/events/showcase
- Related initiatives - R. Phillips & M. Baillie
  - Founding members of active PSI special interest group on visualization
  - https://www.psiweb.org/sigs-special-interest-groups/visualisation/welcome-to-wonderful-wednesdays

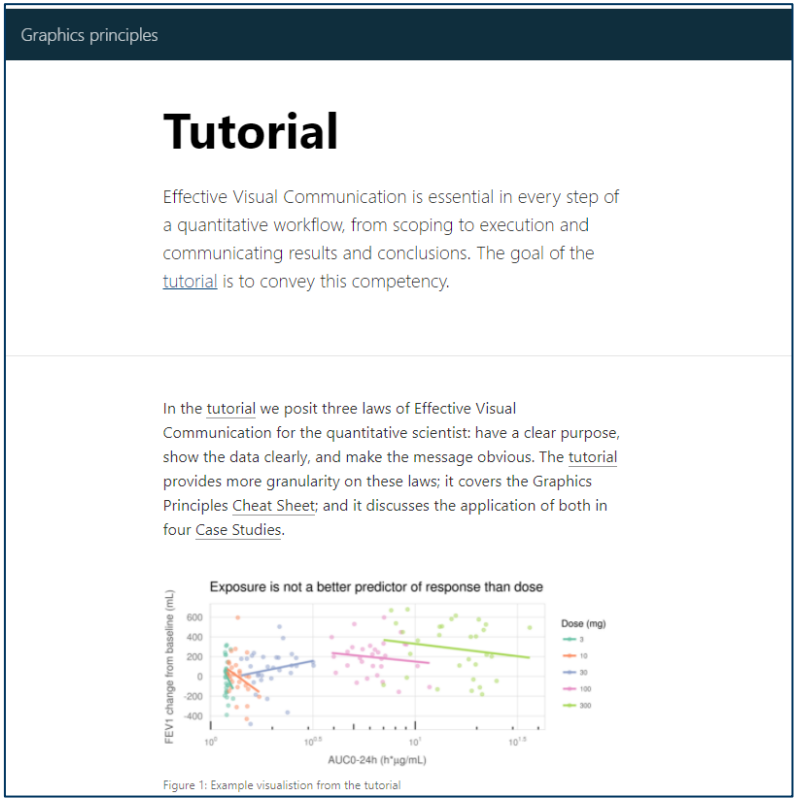
NOVARTIS | Reimagining Medicine

Principles for EVC



https://graphicsprinciples.github.io/cheatsheet.html

Resources, where to find out more?



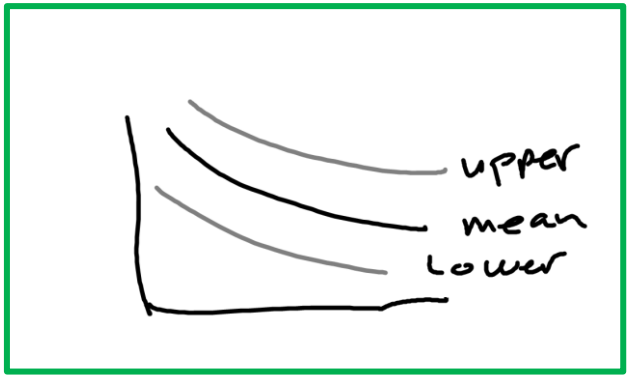
Additional resources:

- Tutorial: https://graphicsprinciples.github.io/
- Exploratory graphics: http://go/xGx
- Graph gallery: http://go/graphgallery
- Cheat sheet: GraphicsCheatSheet
- Paper on effective visual communication
- Video on the three principles
- Presentation checklist http://go/presentationchecklist
- EVC sharepoint
- Wonderful Wednesdays

NOVARTIS | Reimagining Medicine

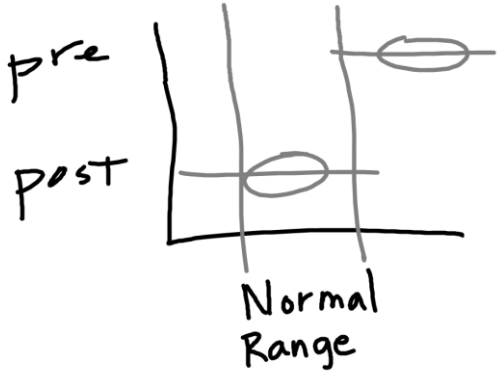
2nd principle - select the appropriate graph

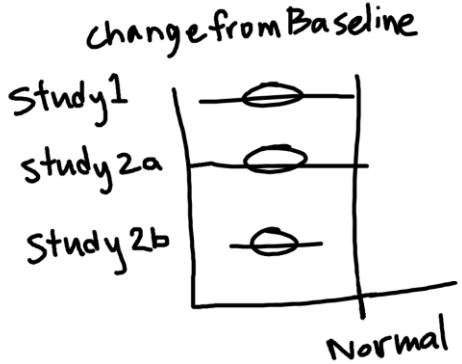
- Come up with several different ways to display the same information
- Display the key evidence in a way that supports the purpose



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





NOVARTIS | Reimagining Medicine

Play the EVC Cartoon

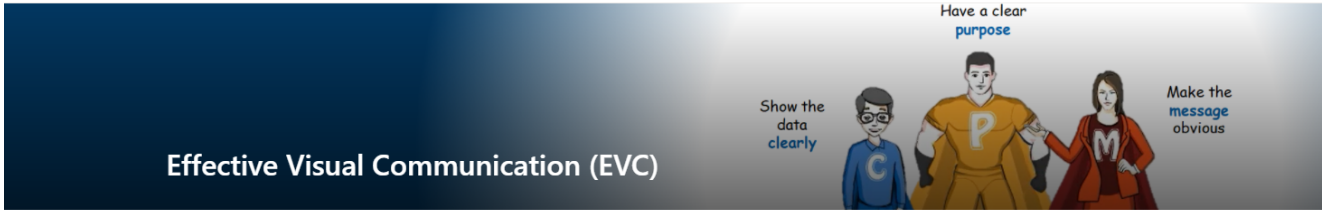
 http://go/graphvideo

Sharepoint: collection of resources

AVA Advanced Visual Analytics M365 Business Use Only (Internal)

Home AVA Outreach Forum Shiny Developer Working Group (SDWG) Analytics

Send to Immersive Reader



Scientists are generating more and more data every day. Our job as quantitative scientists is not only to analyze this data, but also to communicate our analyses in an efficient, effective, and impactful way. As Chambers et al. so elegantly stated:

"There is no single statistical tool that is as powerful as a well-chosen graph" - Chambers et al 1983

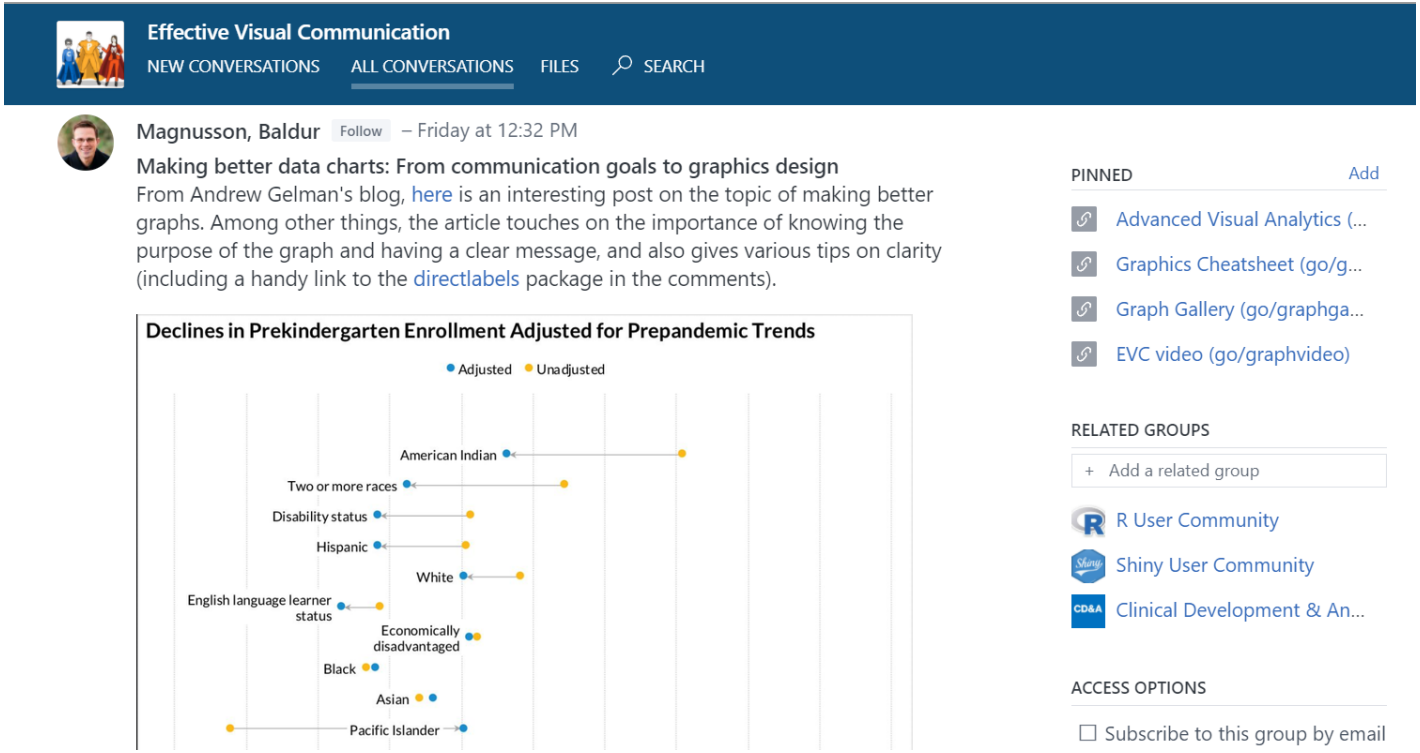
Effective visualizations may appear obvious, but creating them **requires skill**, patience, trial and error. Whether you are just beginning your career, or have many years of experience under your belt, **the following resources can help you** to develop effective data visualizations.

Internal Resources

- **Graph Gallery** (go/graphgallery) - dynamic resource for graphics and corresponding code, displayed in a web browser. Users can browse through samples or search for certain tags (e.g. SAS, R) or keywords (e.g. "Survival","Safety"). Registered users can also rate and comment on individual graphs, participate in discussion forums, upload new graphs (along with code), and request help from graphical experts.
- **Graphics Principles Cheat Sheet** (go/graphcheatsheet) - reference sheet summarizing essential graphical principles on a single page, from selecting the right base graph, to facilitating comparisons (proximity, alignment), effective use of color (minimal, only for

NOVARTIS | Reimagining Medicine

Yammer: community of practice



NOVARTIS | Reimagining Medicine

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

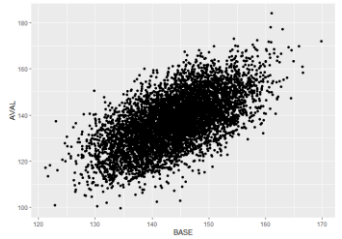
Meta-analysis example data set

- 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

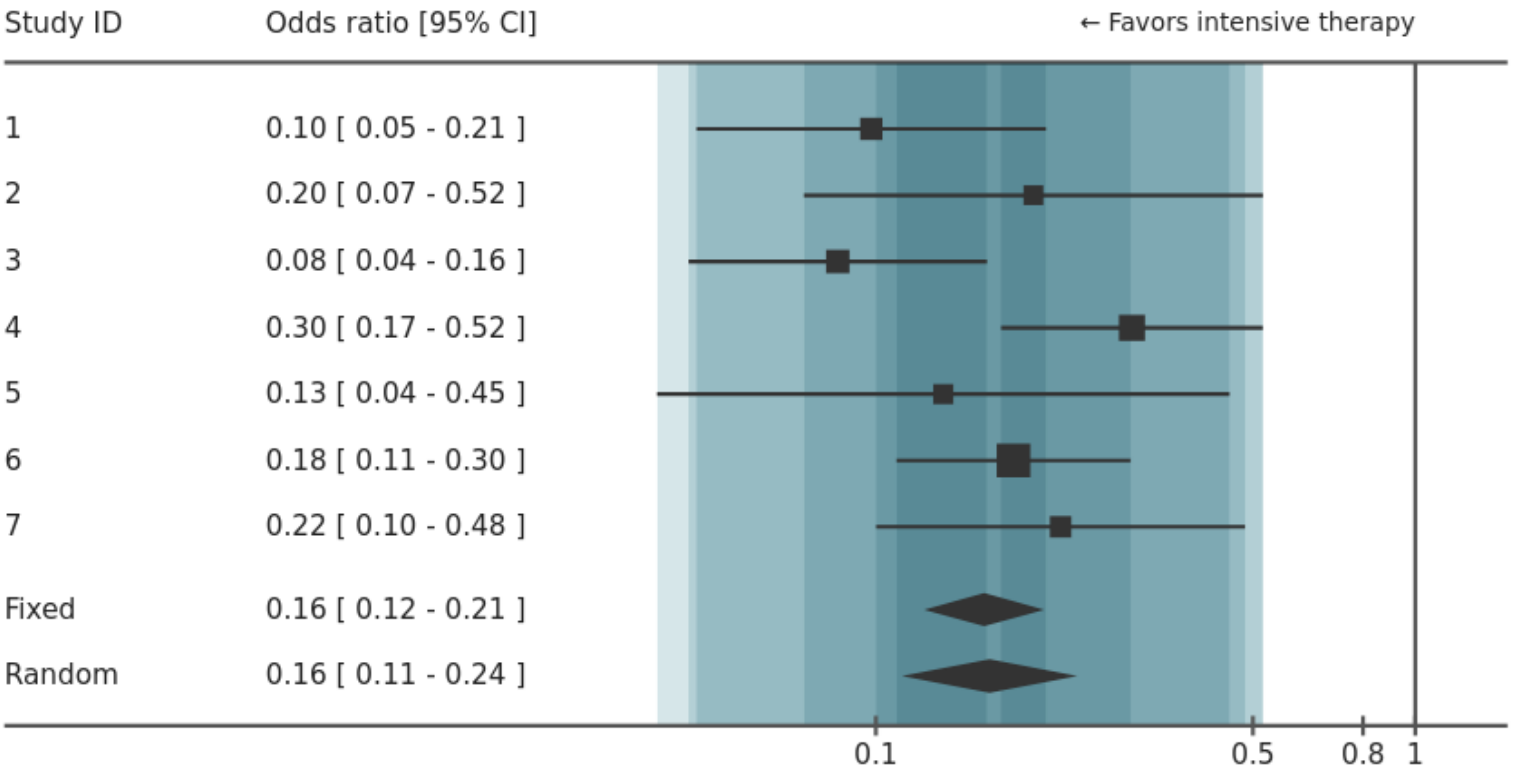
AGE	Age (years)
AGECAT1	Age Group 3: 75 years and older = TRUE
AGECAT1C	Age Group 3: 75 years and older = TRUE
AGECAT1N	Age Group 3: 75 years and older = TRUE
ALBSI	Albumin (g/L)
BASE	Mean systolic blood pressure (mm Hg) measured at baseline
BASOSI	Basophils (Absolute) (10E9/L)
BICARSI	Bicarbonate (mmol/L)
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)



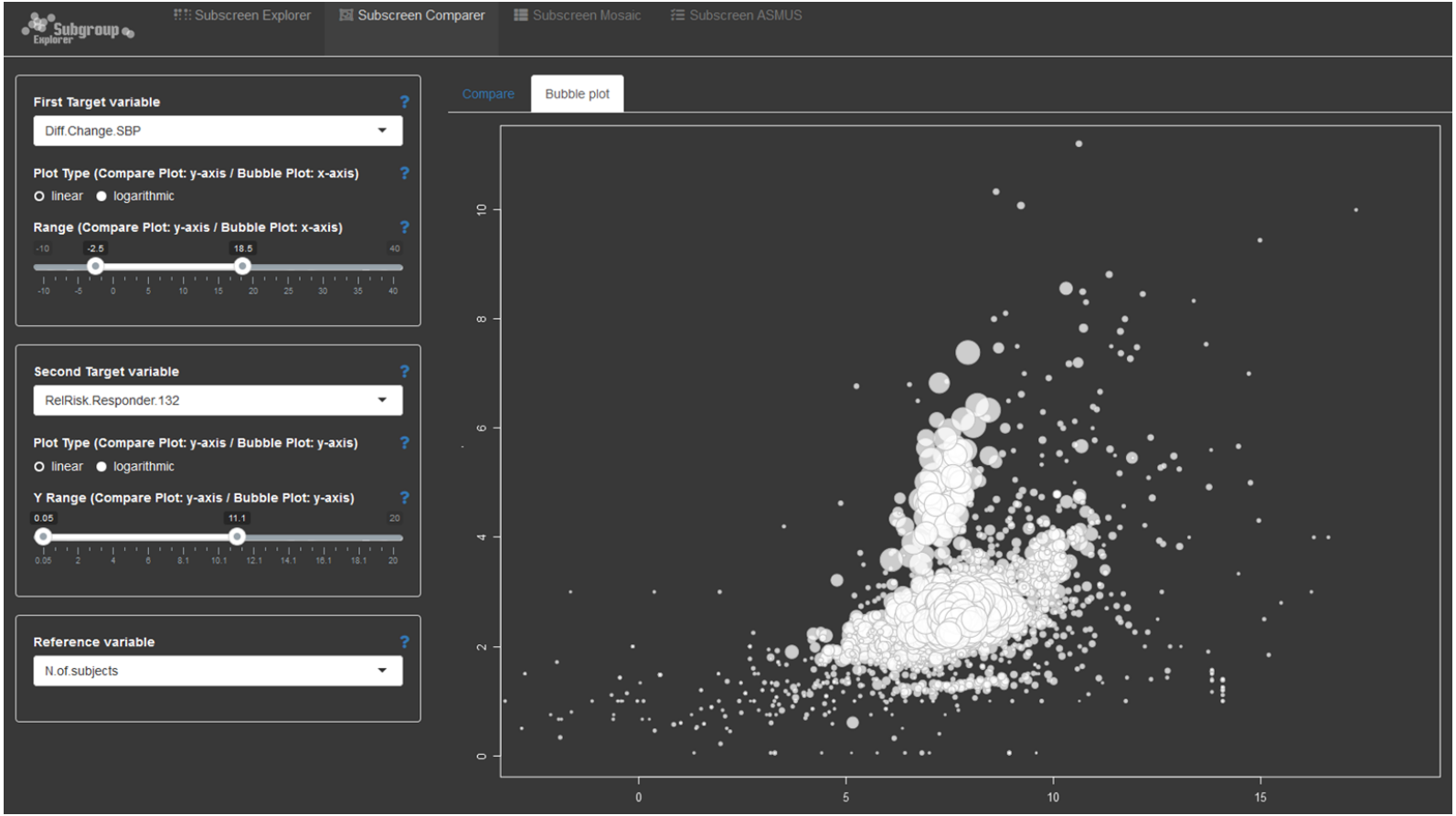
Intensive antihypertensive therapy versus standard of care

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



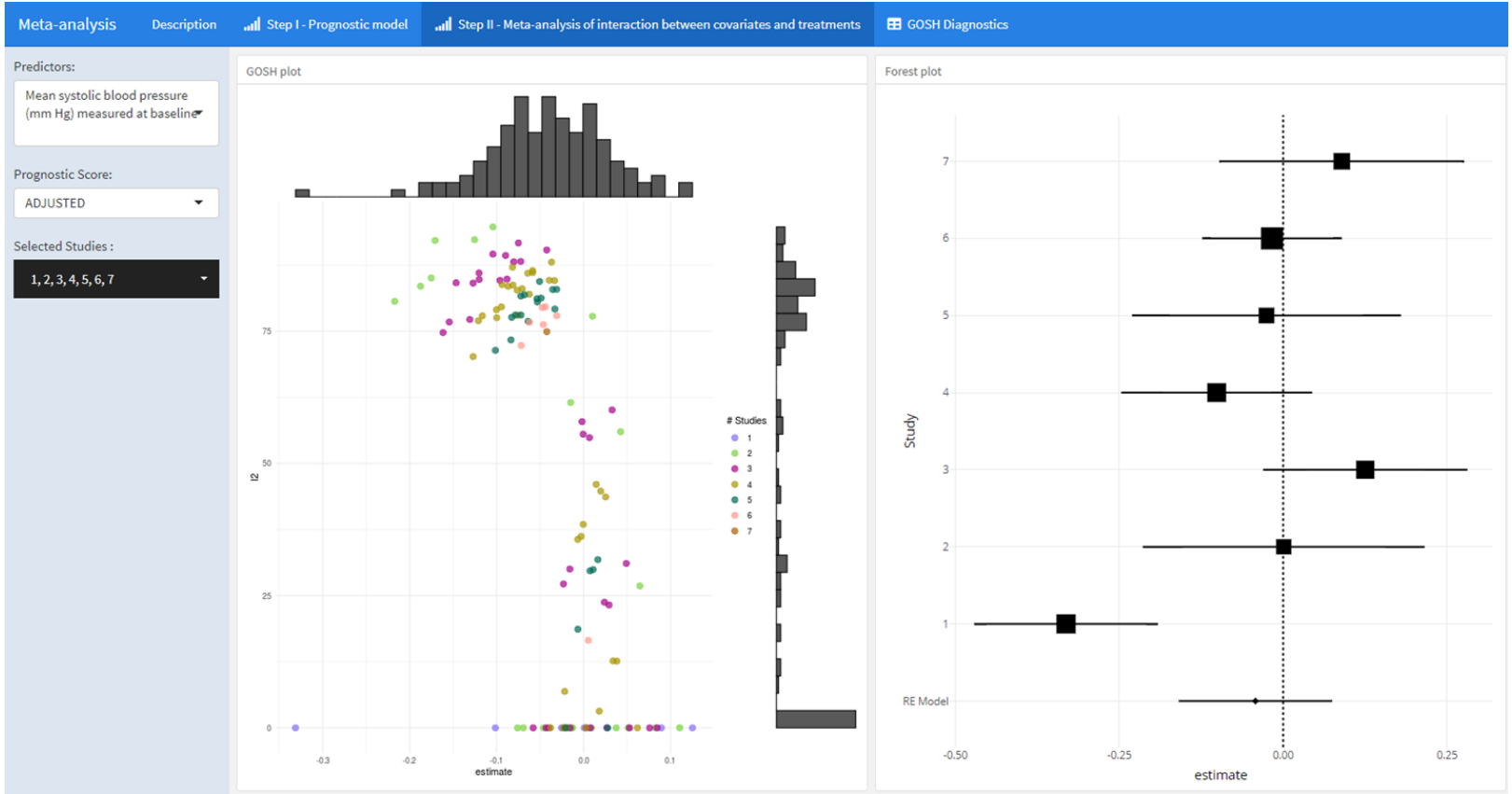
NOVARTIS | Reimagining Medicine

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/

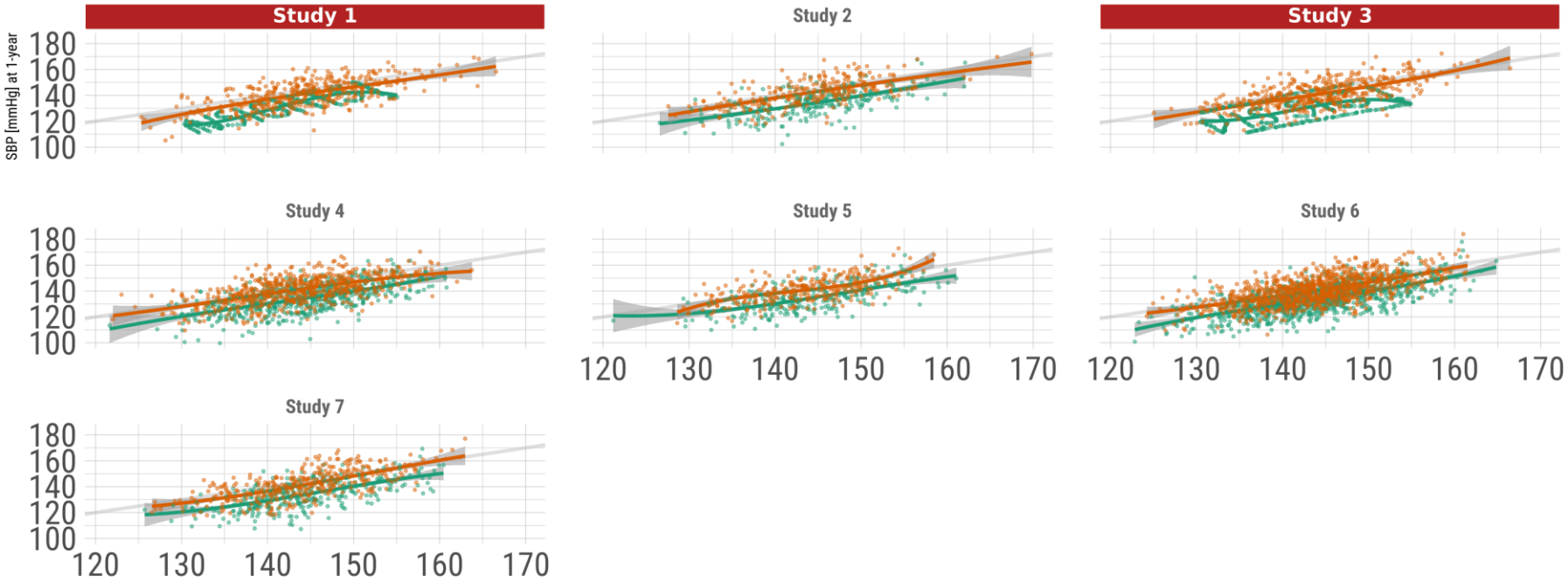
NOVARTIS | Reimagining Medicine

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

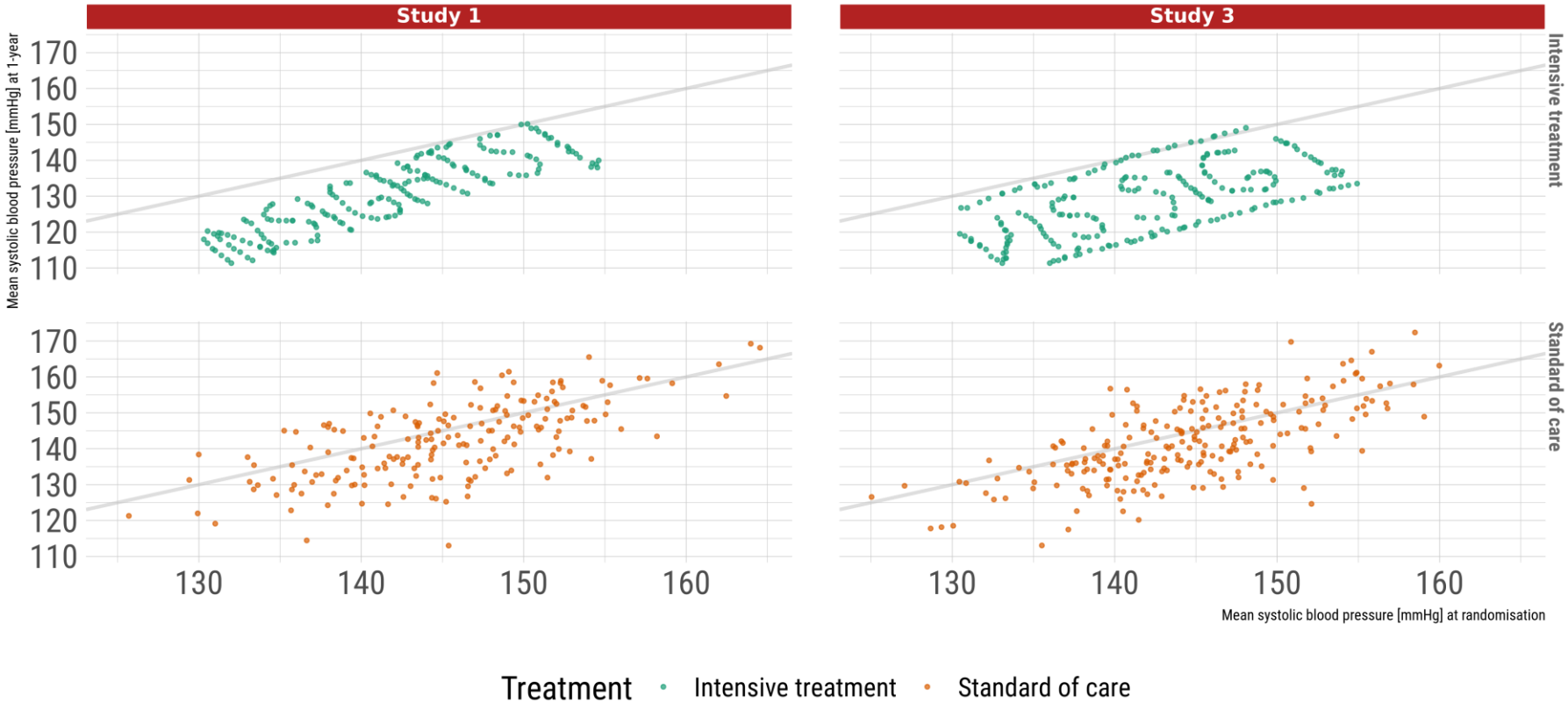
Study 1 and 3 may have data quality issues - further investigation required



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

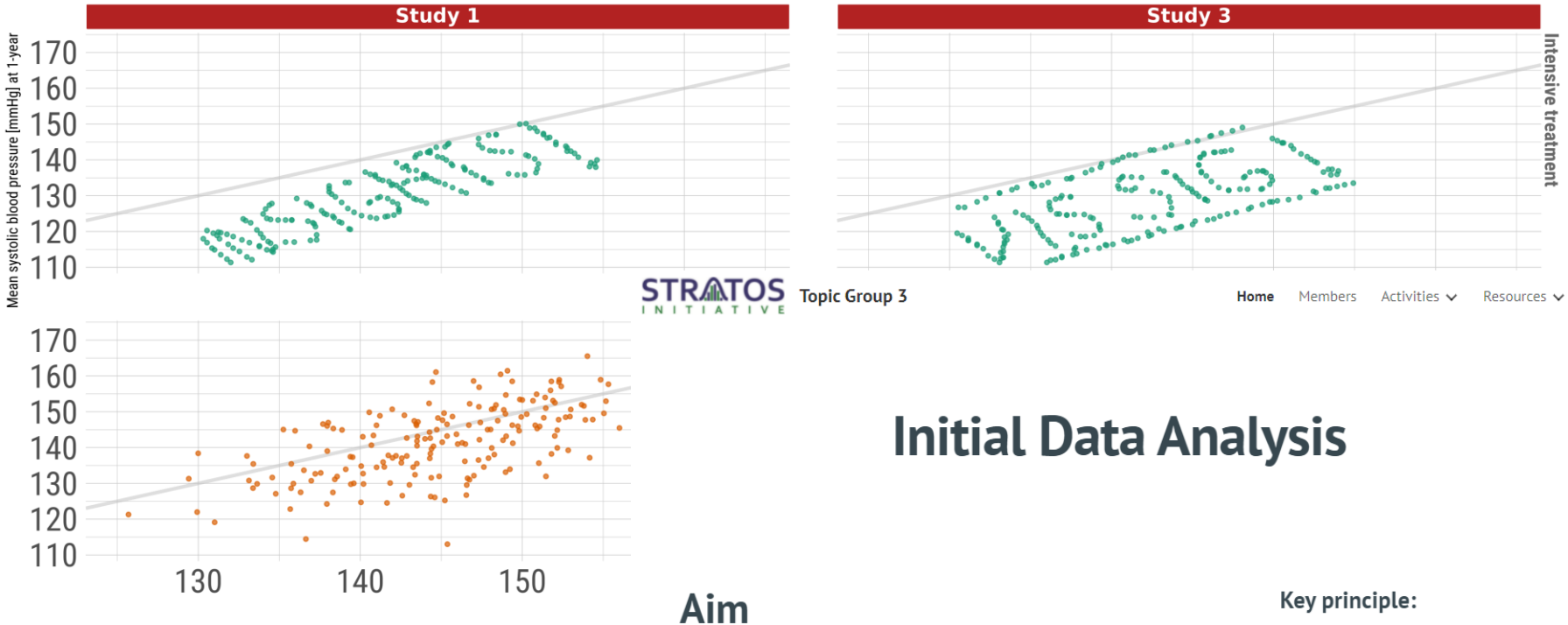
It is always important to plot data many ways



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.

Discussion

- Would like to identify co-chairs, help with activating the panel
- General thoughts or input from the wider group on future direction
- Questions?

NOVARTIS | Reimagining Medicine