Modern graphical approaches

Paolo Eusebi, UCB

Subgroup SIG Webinar

Modern Approaches to subgroup identification

Wednesday 17° Novemebre 2021

Effective visualization: Why

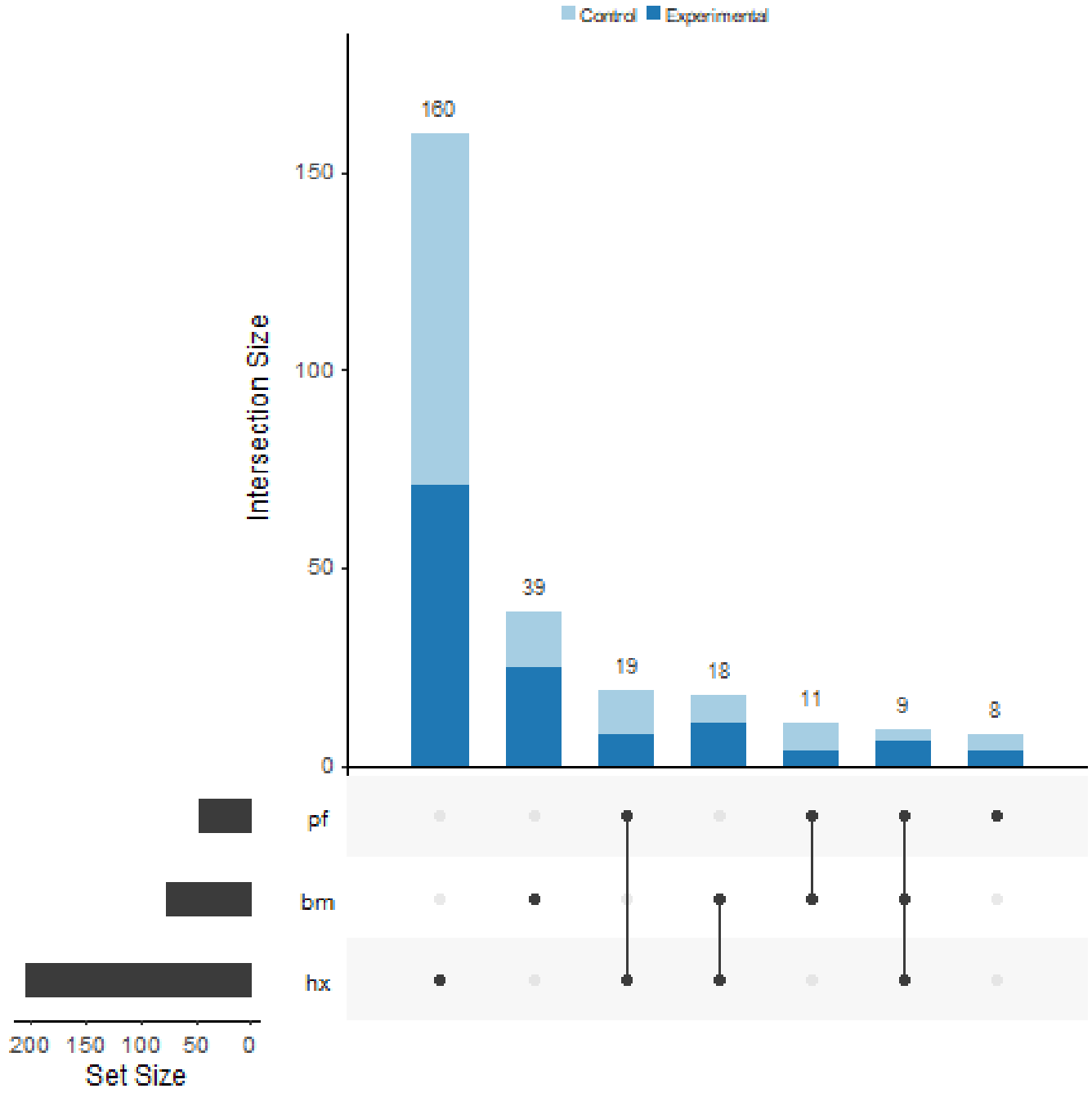
- “A picture is worth a thousand tables”
- To understand and communicate the results to a wider audience
- To leverage the use of animations and interactivity
- To educate our stakeholders
- More useful in data-driven scenarios for subgroups identification

Effective visualization: How

- Review of graphical approaches
- Interactivity/animations
- Multiple studies

Upset Plot Improved

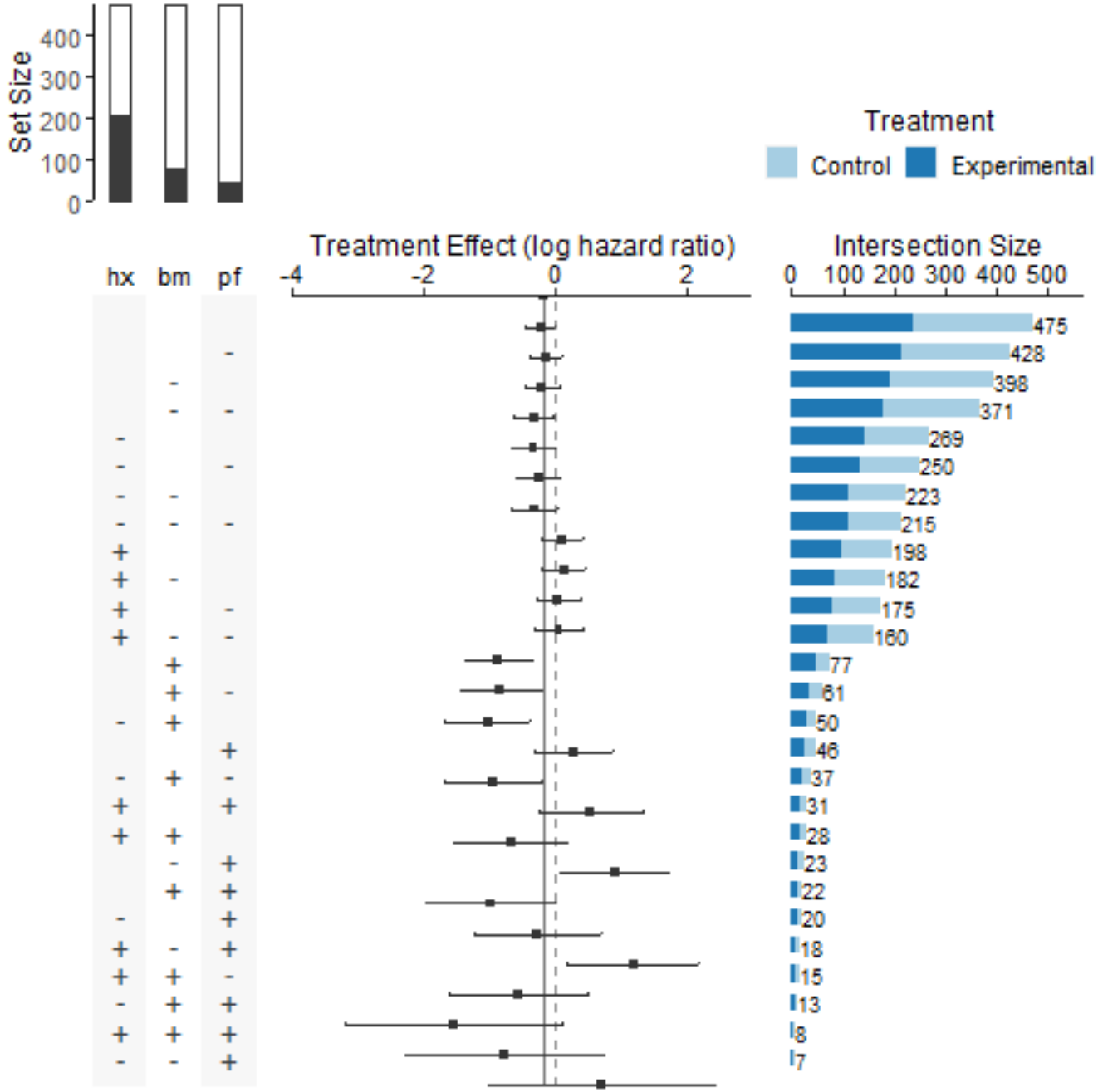
- Visualization of the subgroups formed by the intersection of all binary subgroup-defining covariates.
- The matrix layout on the bottom allows visualizing the composition of the subgroup by showing which sets are intersected.
- The main bar plot displays the sizes of the subgroups that are defined by the respective intersections.



Ballarini et al. "*A critical review of graphics for subgroup analyses in clinical trials*" Pharmaceutical Statistics. 2020

Upset Plot Improved

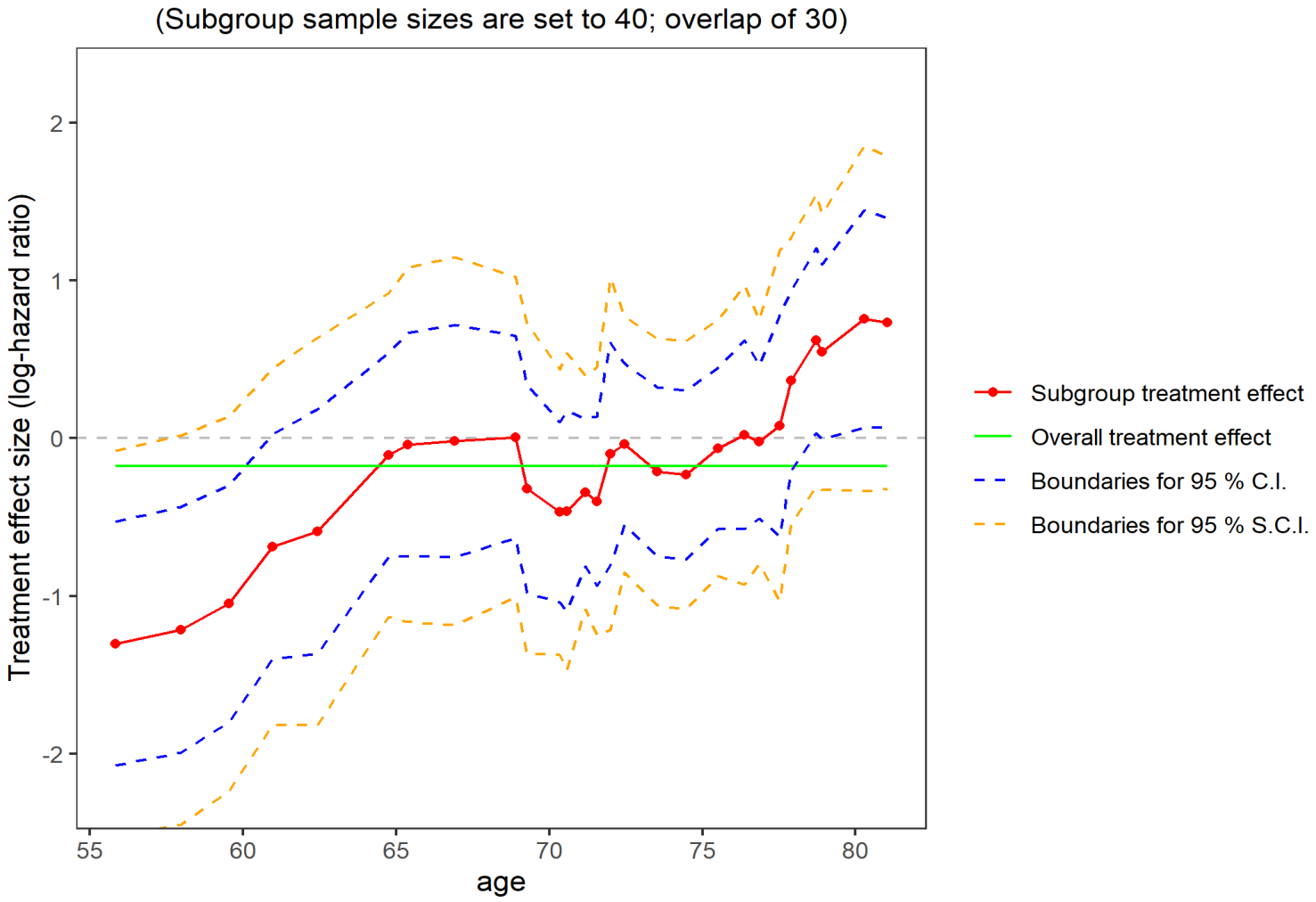
- Upset Plot included in the display of effect sizes.
- The log-hazard ratio and its confidence interval for each subgroup are shown as in a forest plot.
- The UpSet plot provides the advantage of displaying intersections of sets
- Intersections expanded to “either”



Ballarini et al. “*A critical review of graphics for subgroup analyses in clinical trials”* Pharmaceutical Statistics. 2020

STEPP plot for continuous covariates

- Subgroups with sample sizes of around 40 with an overlap of 30 subjects with immediately neighboring subgroups
- Orange dashed lines display 95% simultaneous confidence interval (SCI)
- Blue dashed lines display individual 95% CI (without multiplicity adjustment)
- Red line connects the point estimates of treatment effect (log-hazard ratio) for all subgroups
- Green line represents the log-hazard ratio estimate for the full patient population



Ballarini et al. "*A critical review of graphics for subgroup analyses in clinical trials*" Pharmaceutical Statistics. 2020

When a graphical display of subgroups is good

Checklist from Ballarini and colleagues:

- **Effect size -** Displays effect sizes for subgroups
- **Uncertainty -** Provides confidence intervals or standard errors of the treatment effect estimates
- **Sample size -** Exhibits subgroup sample sizes
- **Intersections -** Shows effect sizes for multivariate subgroups
- **Many covariates -** Applicable to many subgroup-defining covariates

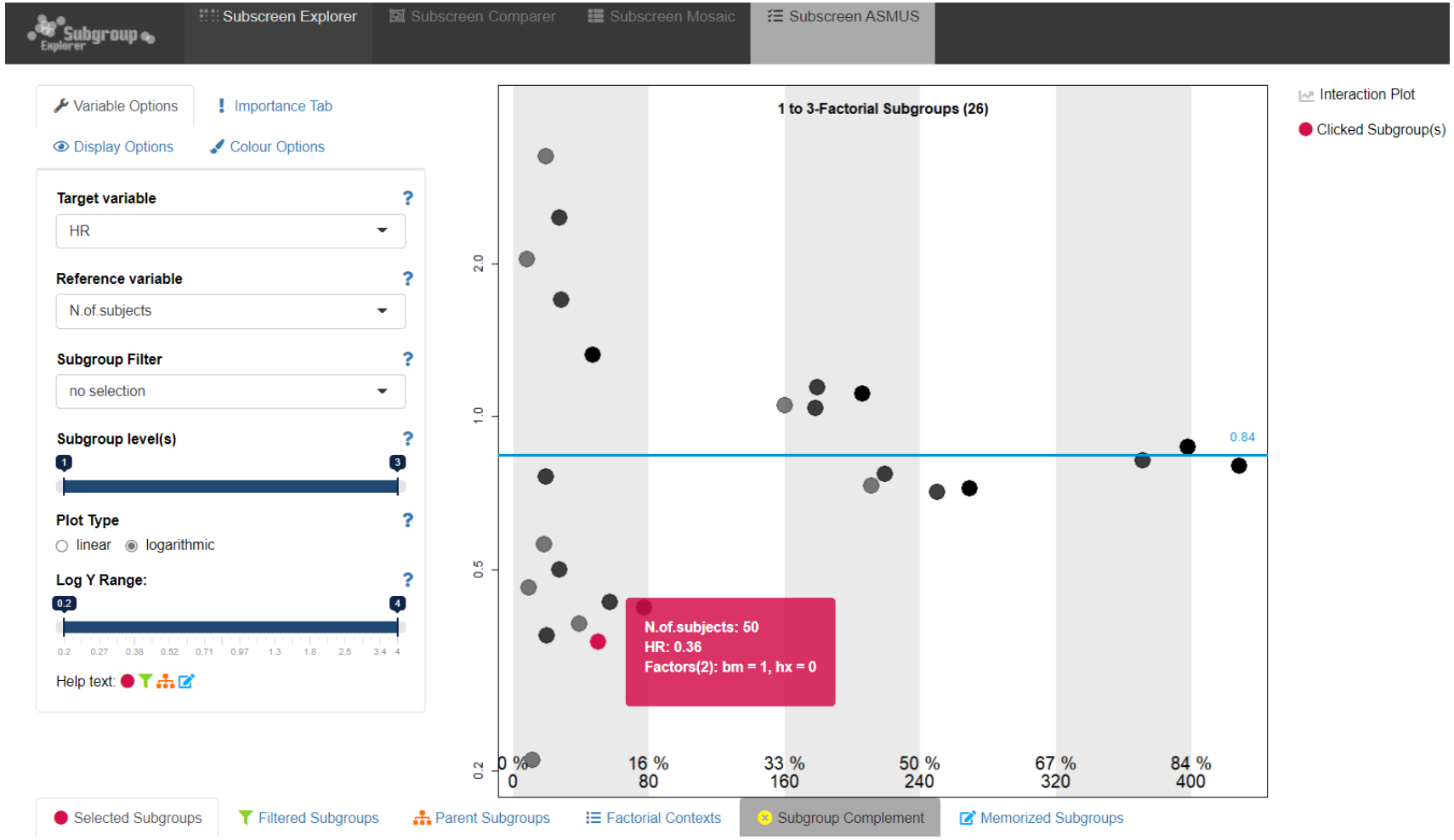
Ballarini et al. "*A critical review of graphics for subgroup analyses in clinical trials*" Pharmaceutical Statistics. 2020

Extension

- **Interactivity** – Allows interactivity and exploration
- **Many studies** – Allowing transparent pooling of different evidence
- **Dealing with biases** – Adjusting for prognostic effects, propensity to be treated.

Subgroup explorer (subscreen R package)

- All the combinations of subgroups in one funnel plot.
- P-values and CIs not displayed (guidance in UX)
- Suitable for any outcome, model needs to be coded before launching the app.



Muysers et al. "A systematic approach for post hoc subgroup analyses with applications in clinical case studies." TIRS 2020

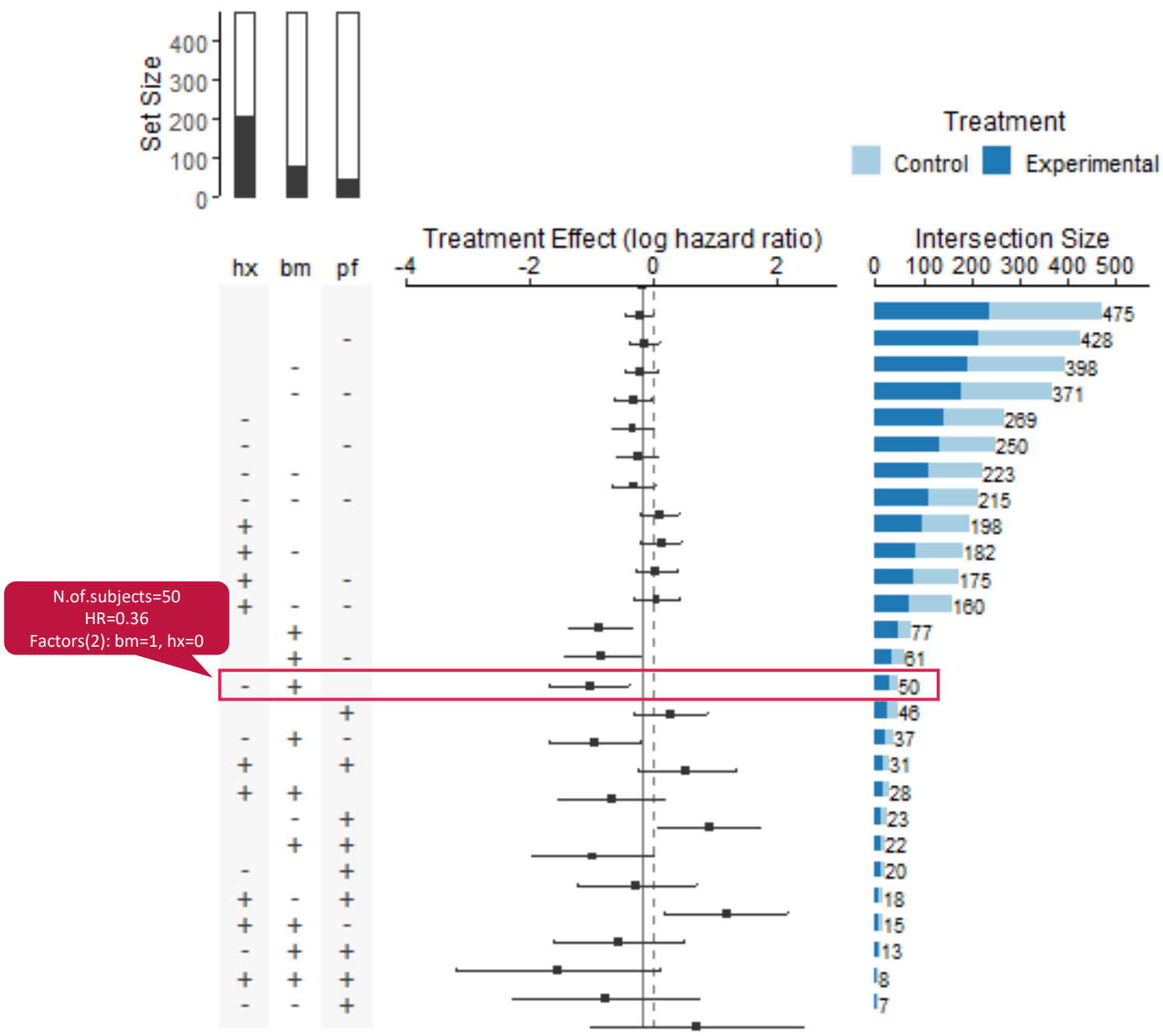
Subgroup explorer (subscreen R package)

- A ML algorithm highlights subgroups (pink dots) driven by the most relevant covariates.

Muysers et al. "A systematic approach for post hoc subgroup analyses with applications in clinical case studies." TIRS 2020

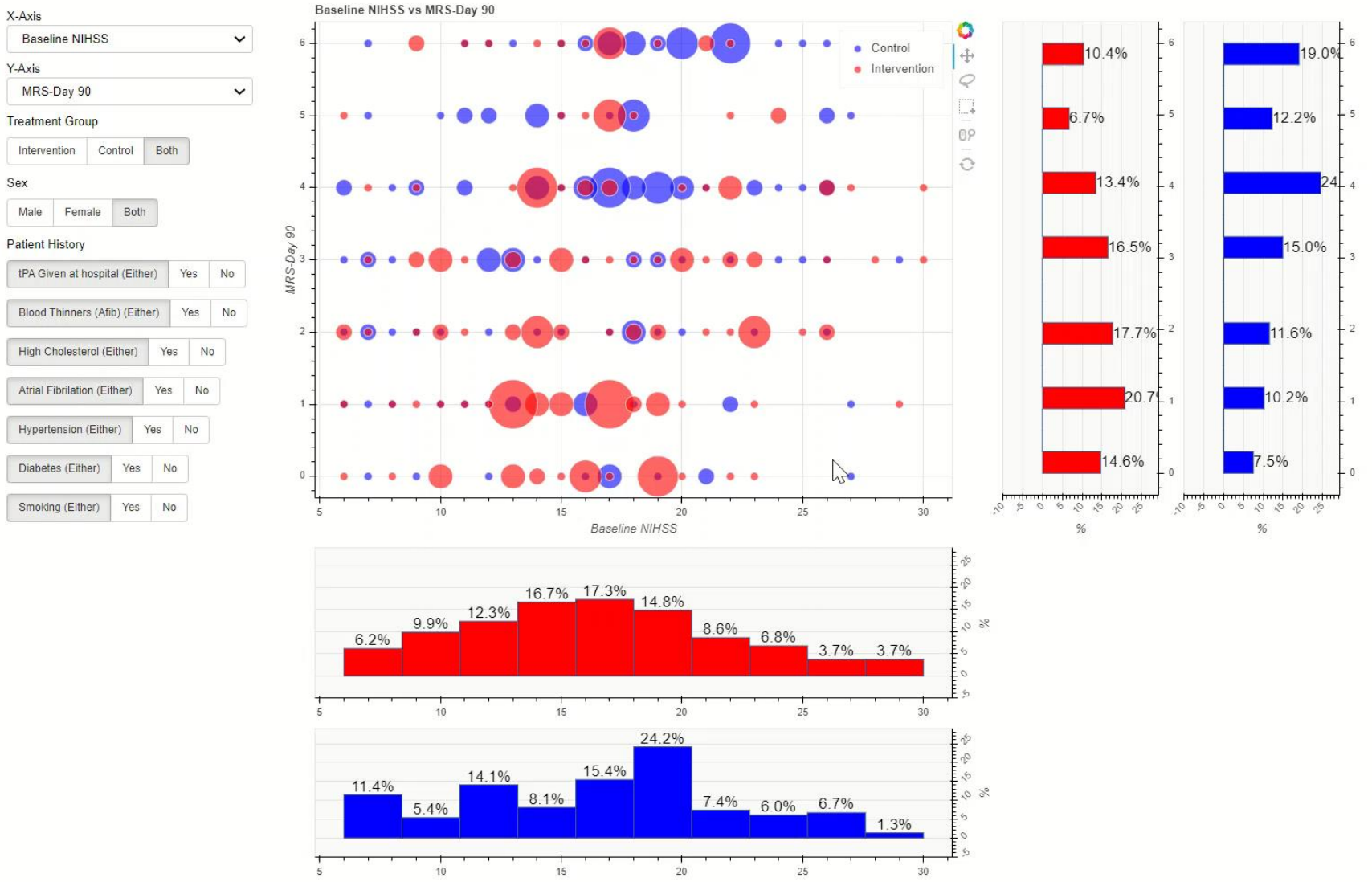
Upset Plot Improved

- Upset Plot included in the display of effect sizes.
- The log-hazard ratio and its confidence interval for each subgroup are shown as in a forest plot.
- The UpSet plot provides the advantage of displaying intersections of sets
- Intersections expanded to “either”



Ballarini et al. “*A critical review of graphics for subgroup analyses in clinical trials”* Pharmaceutical Statistics. 2020

A case study from a study on stroke



Brigdan et al. Stroke 2018 "*Novel Interactive Data Visualization: Exploration of the ESCAPE Trial*" Stroke 2018
http://escapevisualization.herokuapp.com/novel_vis

Exploring robustness of subgroups between studies

Why a 2-stage (meta-analytical) approach?

- Predictive/prognostic variables
- Several studies conducted for testing safety and efficacy of one treatment (RWE/RCT)
- To explore heterogeneity at multiple levels

A case study from a study on stroke

- 7 studies
- 3 small studies (n=100) and 4 big studies (n=500)
- 5 laboratory markers ~ N(0,1)
- 5 genetic markers (Yes/No)
- 2 treatments
- Prognostic: Lab Marker A, Lab Marker B and Gene A
- Predictive: Lab Marker A and Gene A

A Three Step Approach

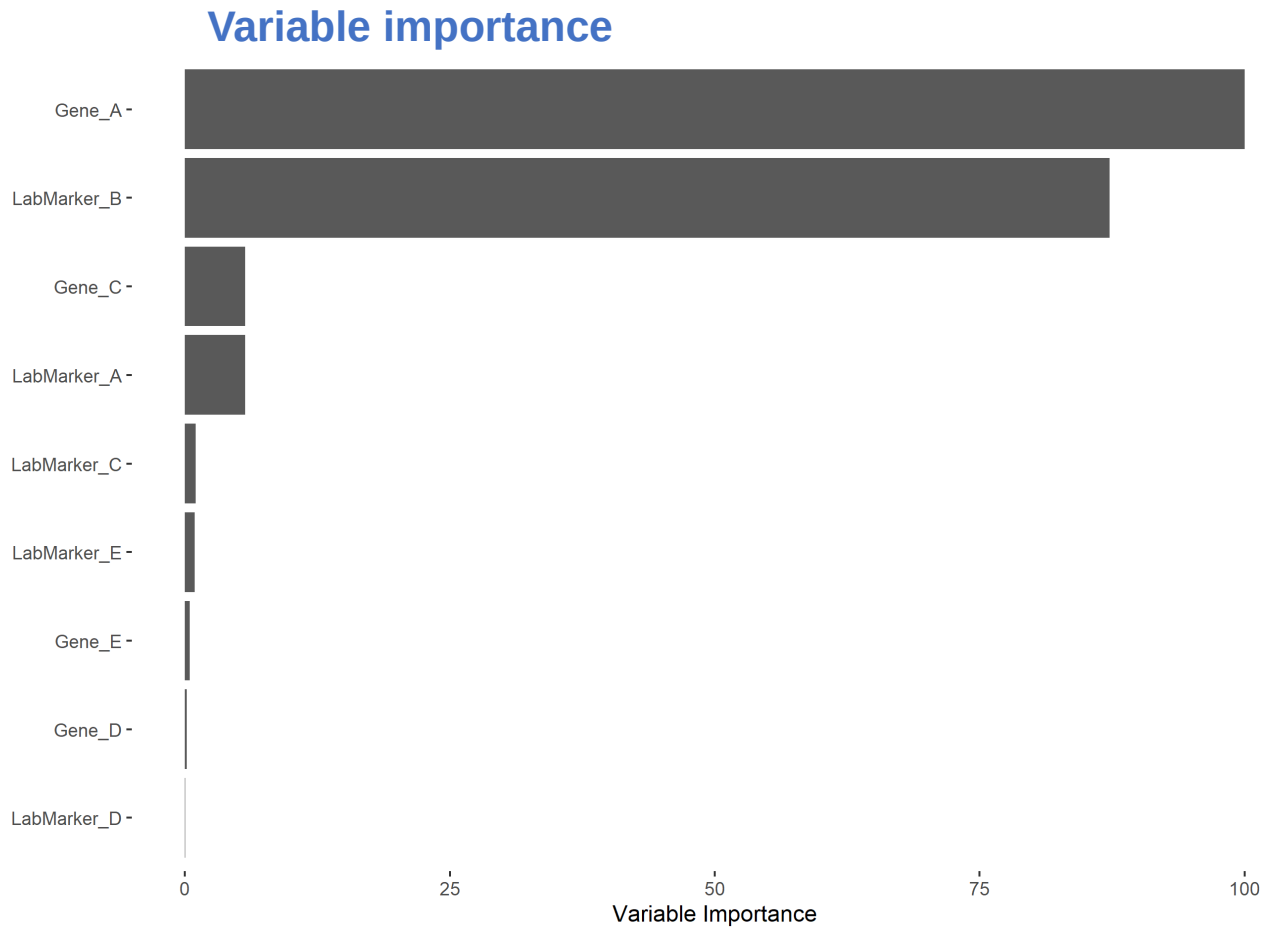
- **Step I** Developing / validating the prognostic model in the Standard of Care arm
- **Step II** Meta-analysis of interaction between baseline variables and treatment for predicting the outcome of interest
- **Step III** Graphic Display of Heterogeneity (GOSH) Diagnostics

I - Developing and validating the prognostic score

- Lasso regression fitted on the control arm.
- Repeated (n=100) ten-fold cross validation.
- The prognostic score is validated on the testing set (20%) and added to the dataset.

Confusion matrix (validation set)

	Observed=1	Observed=0
Predicted=0	92	12
Predicted=1	13	170



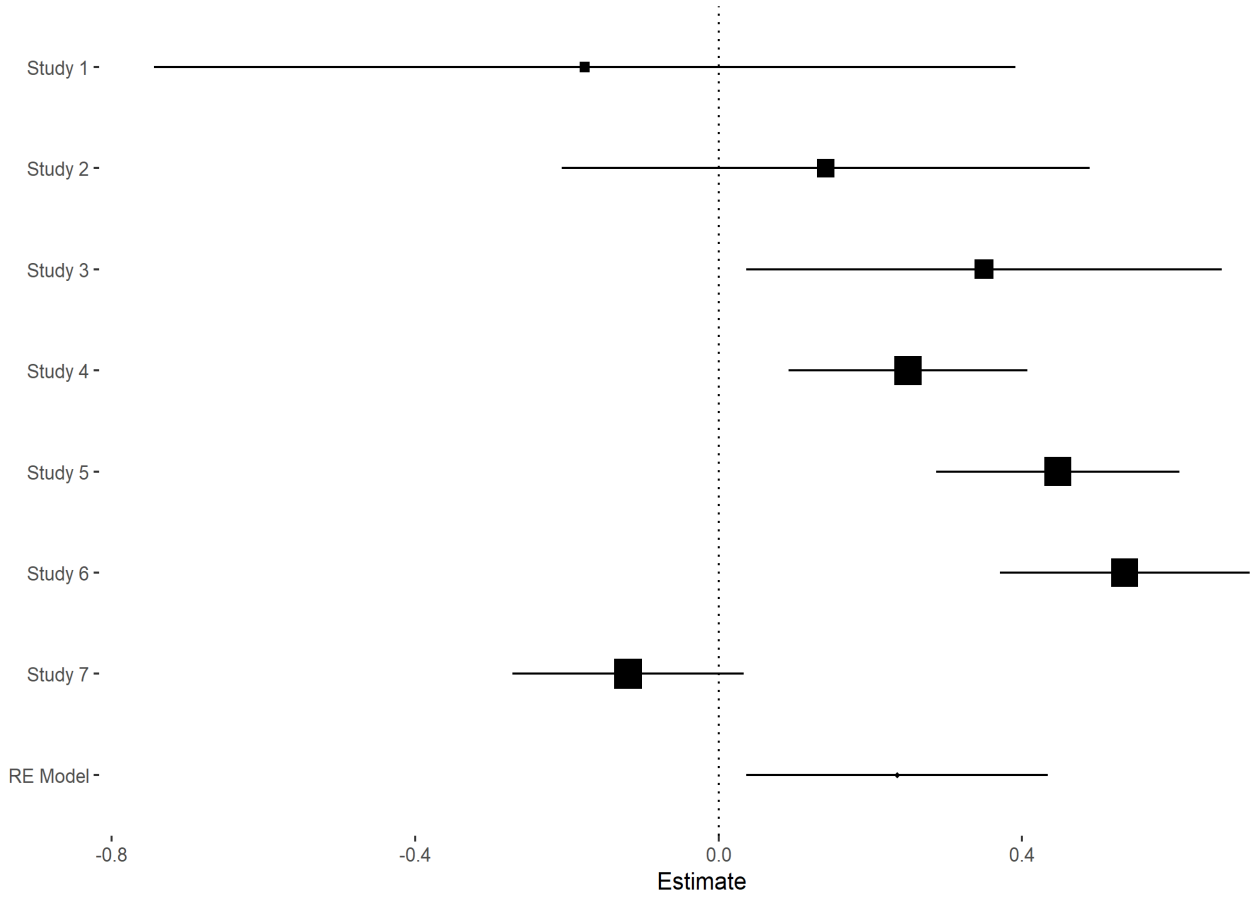
II - Meta-analysis of interaction

- Fitted models with interaction between LabMarker_A and Treatment including the prognostic score in the main effects.

$$Y = b_0 + b_1 \cdot Score + b_2 \cdot (T \times LabMarker\_A)$$

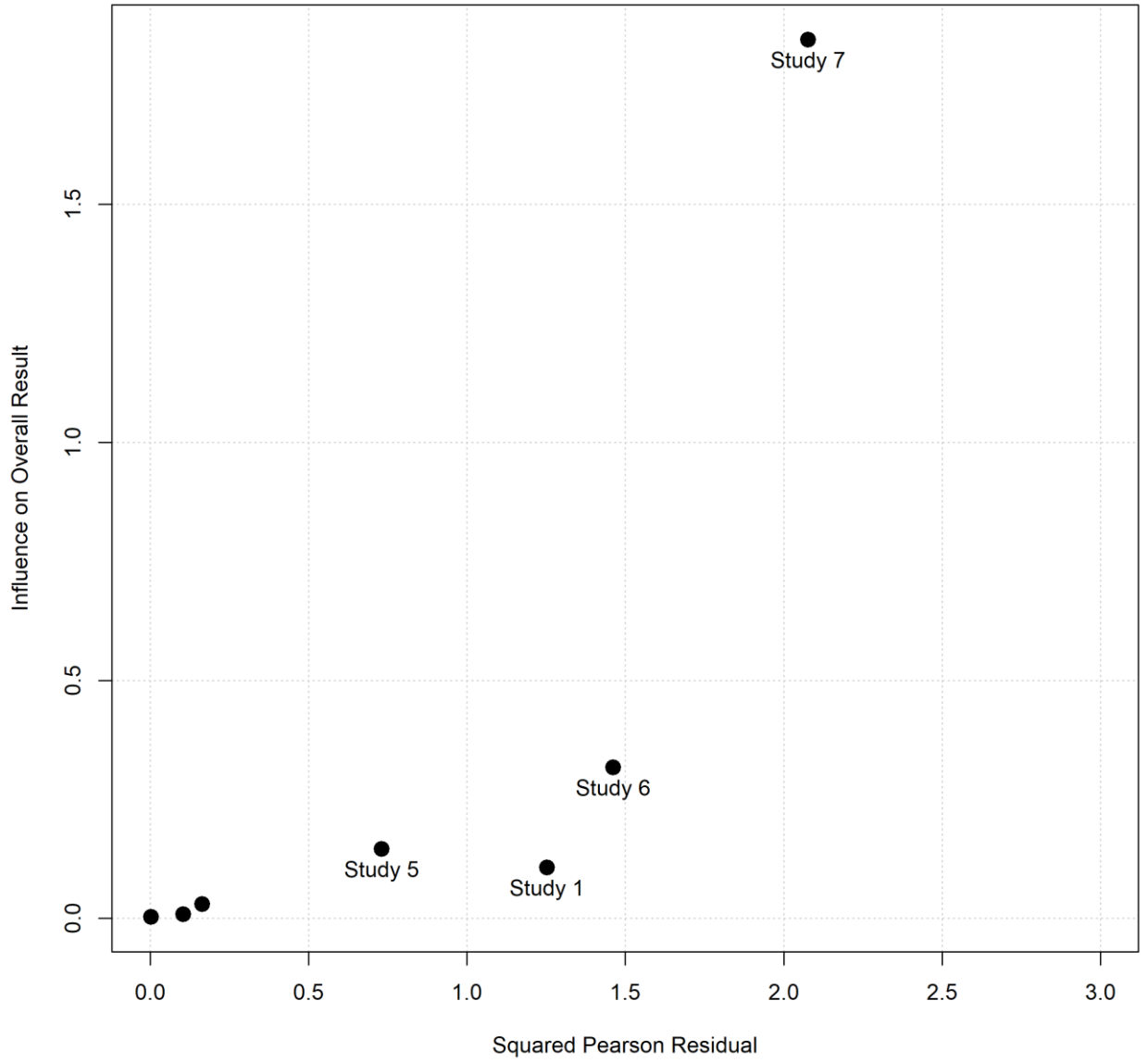
- Interaction terms b_2 meta-analyzed with a random-effect model.

Forest plot of interaction between LabMarker_A and Trt (b_2)



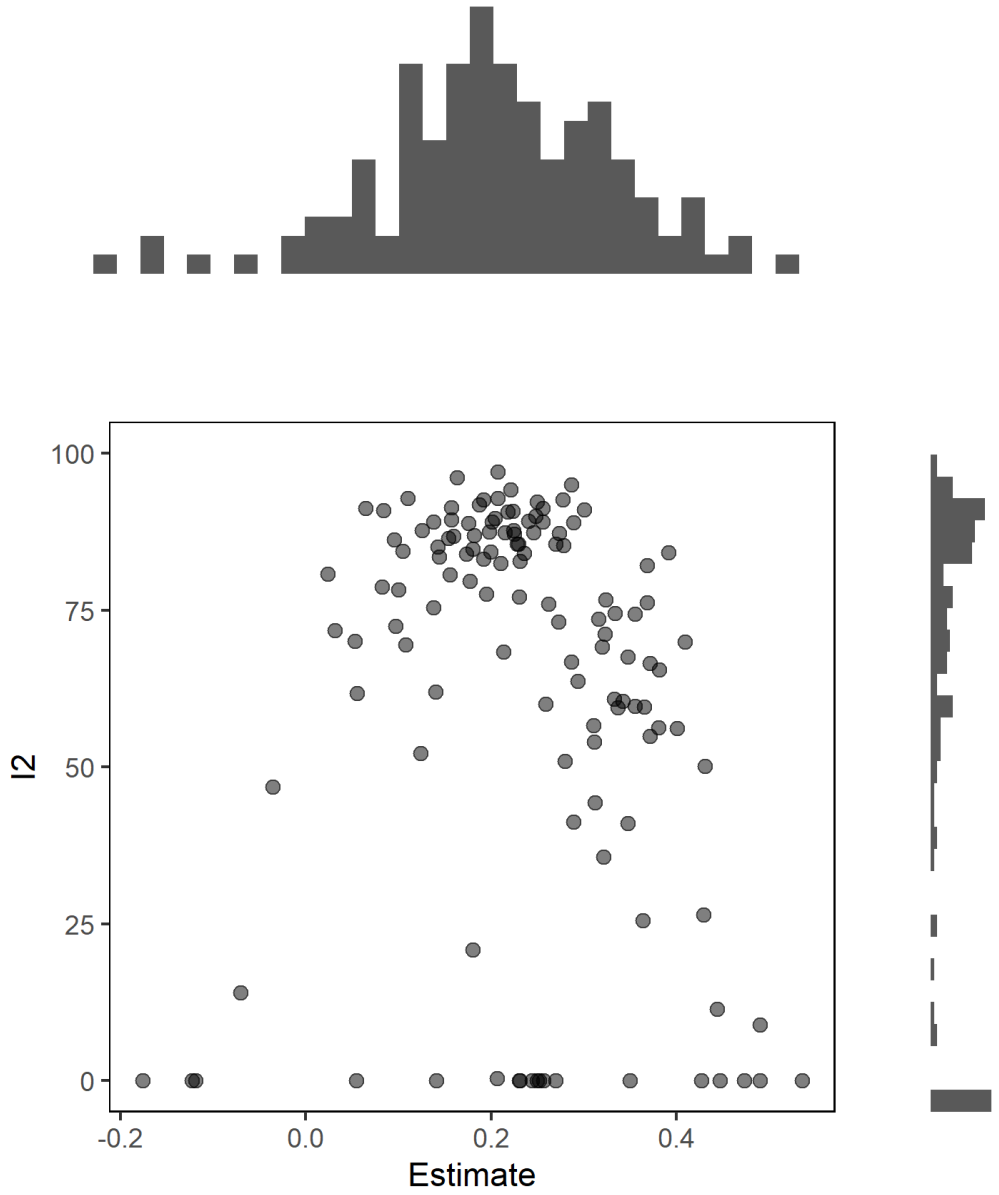
III - Graphic Display of Heterogeneity Baujat plot

- Influence on overall results versus contribution to heterogeneity

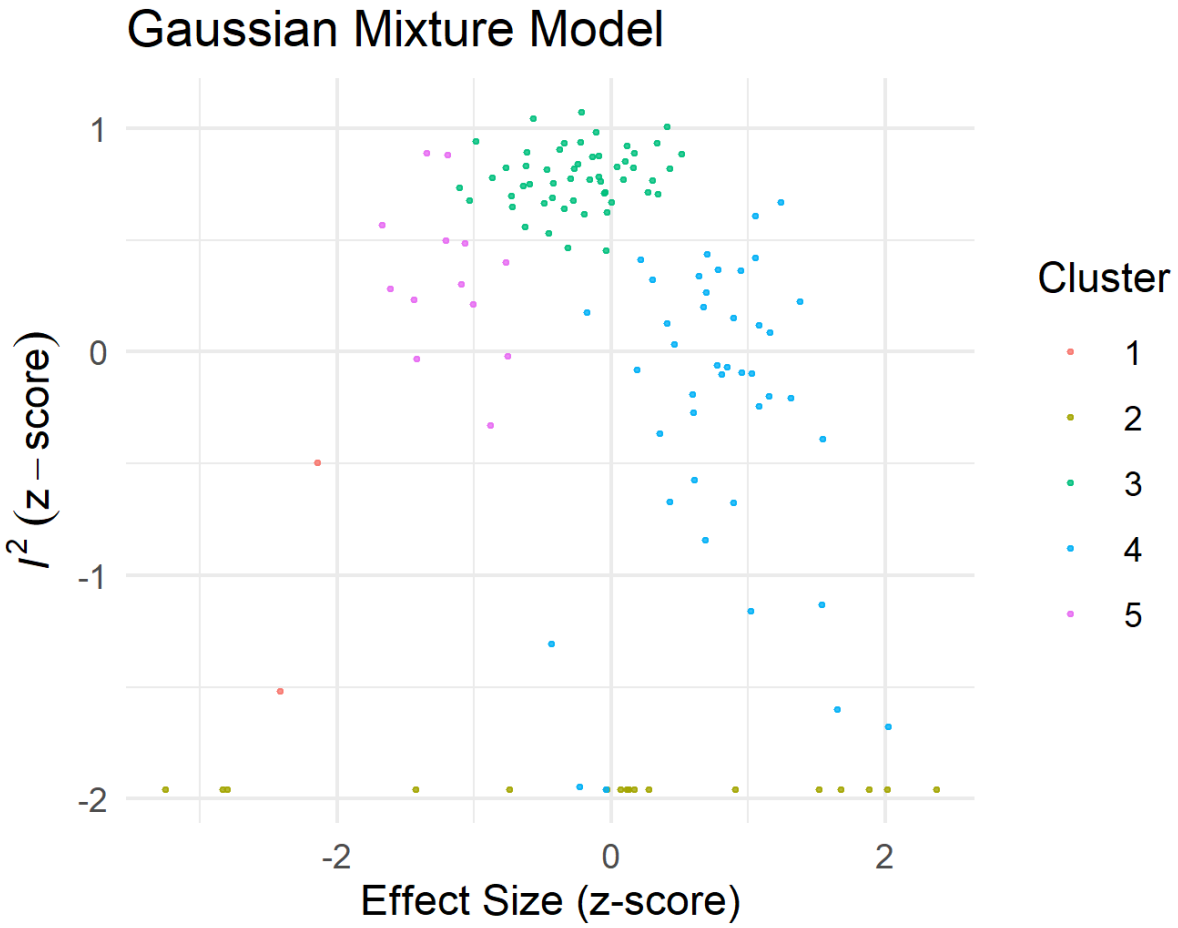


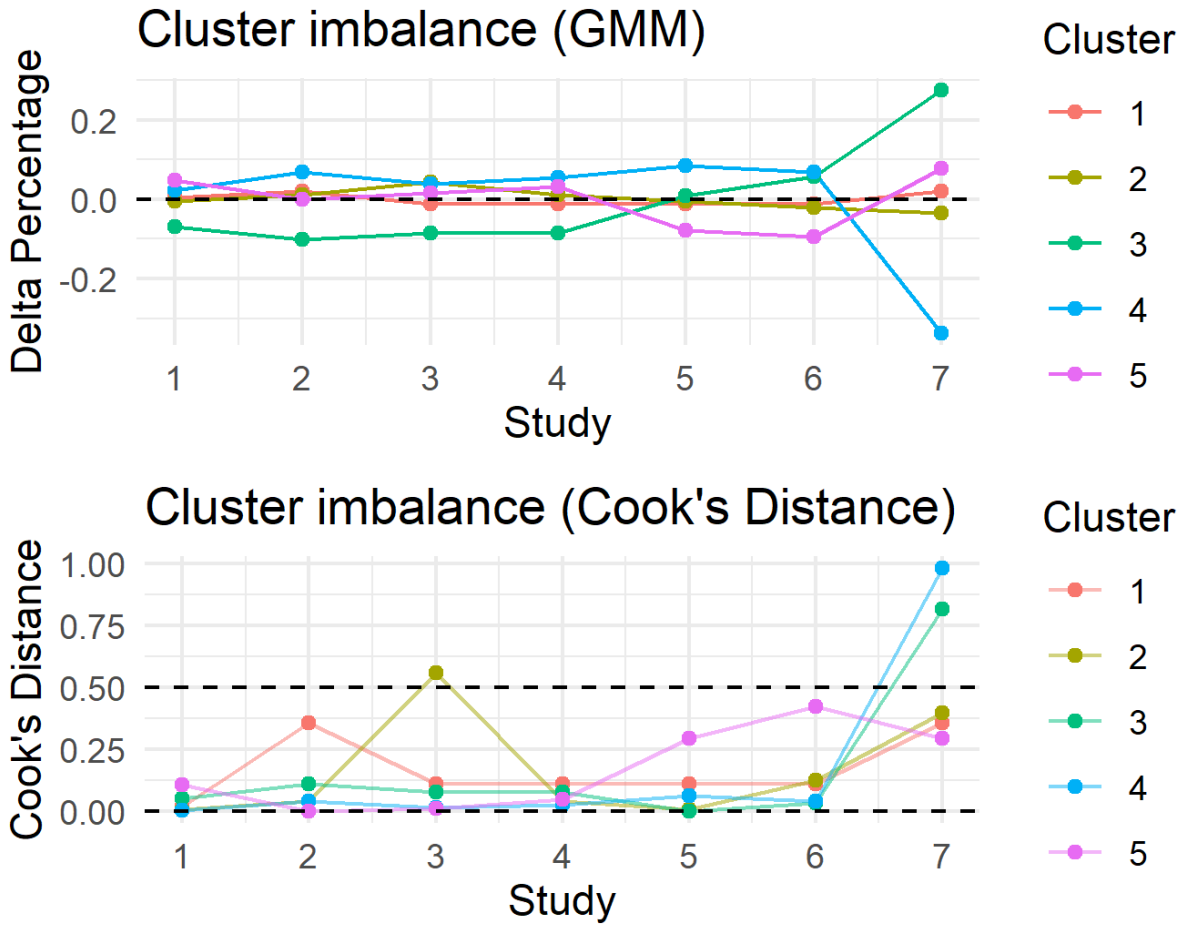
III - Graphic Display of Heterogeneity GOSH plot

- Graphical Display of Heterogeneity for investigating patterns of the model Estimate and Heterogeneity (I^2) across all combination of studies.
- A new dataset with 127 points



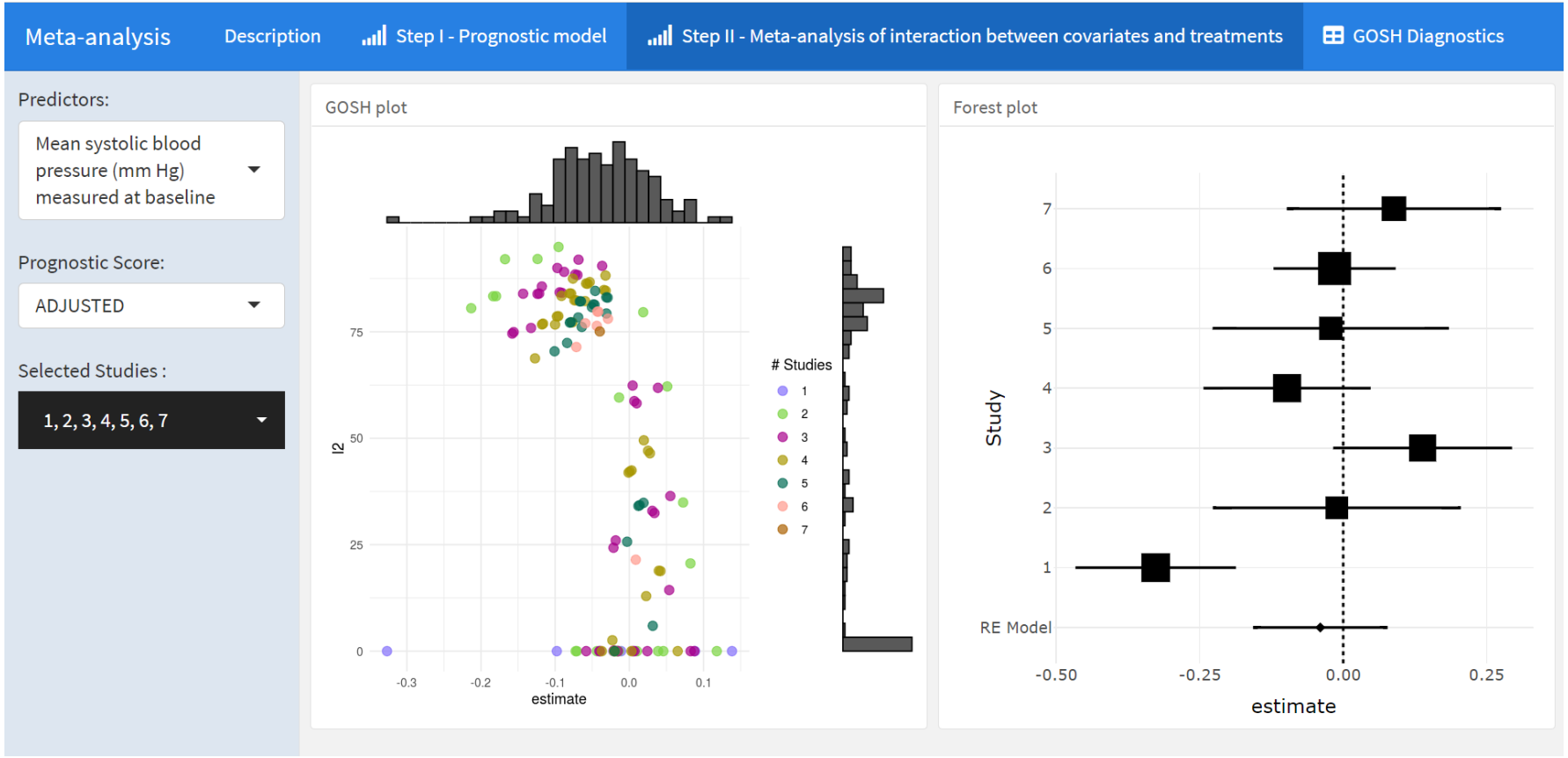
III - GOSH Diagnostics





A Shiny App

- Demo developed for a Wonderful Wednesday session (VIS PSI SIG).
- The purpose of the challenge was to explore how data visualization can be deployed to find insights when faced with big data.



https://vis-sig.github.io/blog/posts/2020-12-03-wonderful-wednesdays-december-2020/

Conclusions

- Visualization is an important component of subgroup exploration/identification
- Interactivity could play an important role
- Important principles to take into account in terms of UX design and statistical rigor
- Interactive exploration of safety data with regulators is already happening

Aknowledgments

- David Svensson (event / inputs for the presentation / Subgroup SIG)

- Alexander Schacht (ideas / discussions)
- Gakava Lovemore (programming work)

Thanks!

References / Resources (1)

•Ballarini, Nicolás M., et al. "A critical review of graphics for subgroup analyses in clinical trials." Pharmaceutical statistics 19.5 (2020): 541-560. https://onlinelibrary.wiley.com/doi/pdfdirect/10.1002/pst.2012

•Wonderful Wednesdays December 2020. https://vis-sig.github.io/blog/posts/2020-12-03-wonderful-wednesdays-december-2020/

•Brigdan, Matthew, et al. "Novel interactive data visualization: Exploration of the ESCAPE trial (endovascular treatment for small core and anterior circulation proximal occlusion with emphasis on minimizing CT to recanalization times) data." Stroke 49.1 (2018): 193-196. http://escapevisualization.herokuapp.com/novel_vis

•Lipkovich, Ilya, Alex Dmitrienko, and Ralph B D'Agostino Sr. "Tutorial in biostatistics: data-driven subgroup identification and analysis in clinical trials." Statistics in medicine 36.1 (2017): 136-196.

•Mathias Harrer and Pim Cuijpers and Toshi Furukawa and David Daniel Ebert (2019). dmetar: Companion R Package For The Guide 'Doing Meta-Analysis in R', R package version 0.0.9000, http://dmetar.protectlab.org/.

•Olkin, Ingram, Issa J Dahabreh, and Thomas A Trikalinos. 2012. GOSH a Graphical Display of Study Heterogeneity. Research Synthesis Methods 3 (3). Wiley Online Library: 214
23. https://onlinelibrary.wiley.com/doi/abs/10.1002/jrsm.1053

References / Resources (2)

•Pajouheshnia Romin, Groenwold Rolf H H, Peelen Linda M, Reitsma Johannes B, Moons Karel G M. When and how to use data from randomised trials to develop or validate prognostic models BMJ 2019; 365:l2154 https://www.bmj.com/content/365/bmj.l2154.full

•Chalkou, K., Steyerberg, E., Egger, M., Manca, A., Pellegrini, F., & Salanti, G. (2020). A two-stage prediction model for heterogeneous effects of many treatment options: application to drugs for Multiple Sclerosis. arXiv preprint arXiv:2004.13464. https://arxiv.org/abs/2004.13464

•Hua, H., Burke, D. L., Crowther, M. J., Ensor, J., Tudur Smith, C., & Riley, R. D. (2017). One stage individual participant data meta-analysis models: estimation of treatment covariate interactions must avoid ecological bias by separating out within trial and across trial information. Statistics in medicine, 36(5), 772-789. https://onlinelibrary.wiley.com/doi/full/10.1002/sim.7171

•Tamara Munzner. Visualization Analysis and Design. A K Peters Visualization Series, CRC Press, 2014 https://www.cs.ubc.ca/~tmm/vadbook/

•Muysers, Christoph, et al. "A systematic approach for post hoc subgroup analyses with applications in clinical case studies." Therapeutic innovation & regulatory science 54.3 (2020): 507-518. https://link.springer.com/content/pdf/10.1007/s43441-019-00082-6.pdf