Novartis R&D

Predicting subgroup treatment effects for a new study

Giulia Capestro and **Silvia Zaoli**, Data Scientists
AMLD 2022, Lausanne
March 28, 2022

NOVARTIS | Reimagining Medicine

What is subgroup identification and why we need it

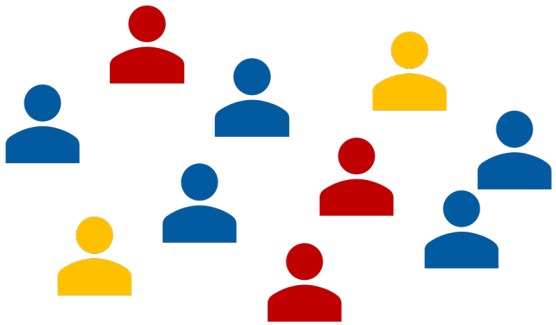
This situation is relatively common:

- **50%** of drugs fail first submission, **13.2%** of them due to inadequate efficacy wrt Standard of Care
- **16%** among drugs that are never approved

Sacks et al, *Scientific and Regulatory Reasons for Delay and Denial of FDA Approval of Initial Applications for New Drugs, 2000-2012*, JAMA

What is subgroup identification and why we need it

Patients are heterogeneous, and so is their response to a drug



Baseline covariates or biomarkers

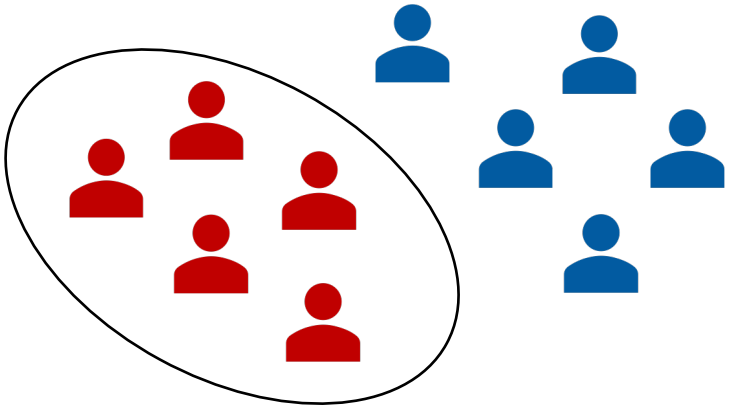
- Demographic (age, sex, …)
- Disease sub-types, severity scores…
- Genetic biomarkers…

Covariates can be:

- **Prognostic** : impact the outcome regardless of treatment received
- **Predictive** : inform about the effect of treatment

What is subgroup identification and why we need it

If we find out which baseline covariates are predictive, we can identify a *subgroup with increased treatment effect*



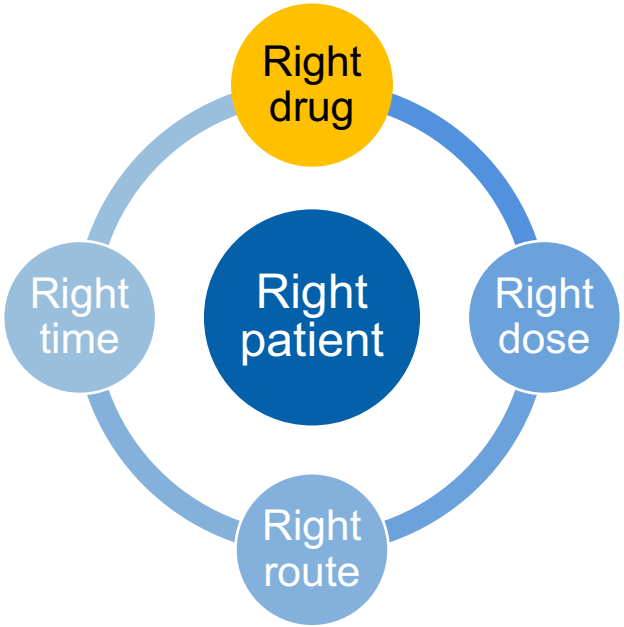
Ex: patients younger than 60 that have symptom X

Subgroup identification

≠

Clustering

Subgroup identification as precision medicine



What is subgroup identification and why we need it

Finding a subgroup which will replicate in a new trial is an extremely hard problem:

- Trial was not designed with this purpose
- Treatment effect in subgroups noisy due to small sample size
- Example: Aspirin harmful for people born under Libra and Gemini!

Crucial to find reliable methods, but the problem remains challenging!

Overview of existing learning methods

Tree-based

- VirtualTwin
- SIDES
- GUIDE
- MOB

…

- Decision trees to split patients based on their characteristics
- Criteria to choose best split depends on method
- Model returns combination of thresholds which define the subgroup

Regression-based

- Lasso & Ridge, GLMnet
- Boosting
- 'FindIT' (SVT+Lasso)
- STIMA (hybrid)

…

- Identify predictive variables thanks to the regression on the outcome
- Does not directly return subgroup definition

The challenge aims

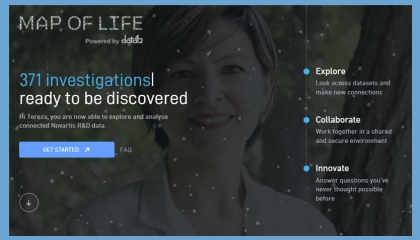
Realistic setting

- Assess teams in a real situation
- Subgroup tested on unseen data, to verify if it replicates

Innovative way of learning

Opportunity to test and compare various methods

Novartis data42 platform



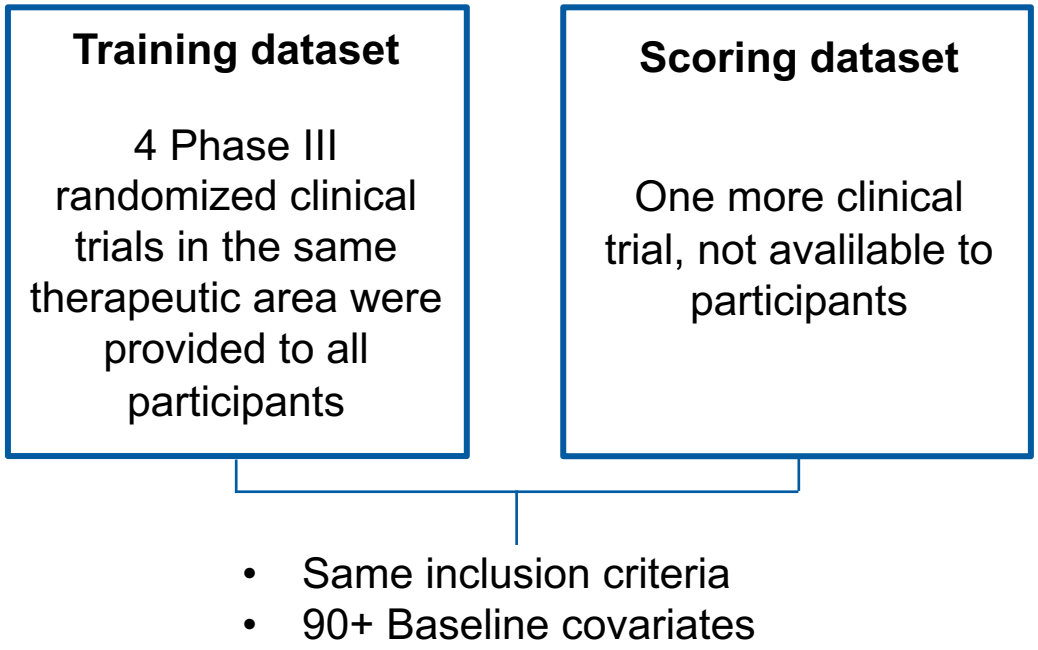
- Large amount of clinical and RWE data
- Several data analysis tools
- Fosters collaboration and reproducibility

NOVARTIS | Reimagining Medicine

The challenge data

Challenge reproduces realistic situation:

- Identified subgroup is tested on new, unseen data
- Allows to test reproducibility



The challenge data

Endpoint

Binary response index

Combination of measures of improvement e.g.:

- patient global assessment
- physician global assessment
- result of health questionnaire
- results of lab tests

Treatment effect

TE = p(treatment) – p(control)

The challenge task

Participants had to submit:

1. A definition of the subgroup for which they predict increased treatment effect in the new trial (ex. “AGE<60”)

2. Their prediction of the treatment effect to be observed for the subgroup in the new trial:

   $$\delta_{pred} = \text{p(treatment)} - \text{p(control)}$$

   and its uncertainty σ_{pred}

3. A description of their methodology, with clinical/biological justification of the subgroup

How we scored the submissions

Does the subgroup have increase treatment effect in the new trial?

- Probability p_i of patient *i* to respond modeled as

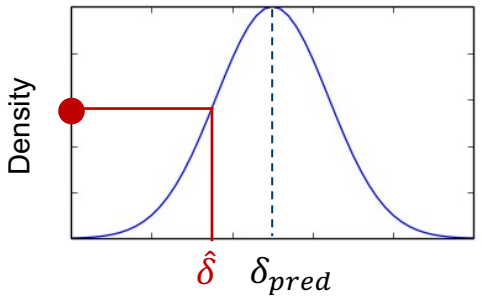
$$logit\,(p_i) = \beta_0 + \beta_{trt} t_i + \beta_s s_i + \beta_{interaction} t_i s_i + x_i' \beta$$

- t_i = treatment (0 – control, 1 – treatment)
- s_i = subgroup (1 – subgroup, 0 – complement)
- x_i covariates as in the primary analysis model for the new trial

- Score obtained as: $\dfrac{\hat{\beta}_{interaction}}{s.e.(\hat{\beta}_{interaction})}$

Is the prediction of treatment effect accurate?

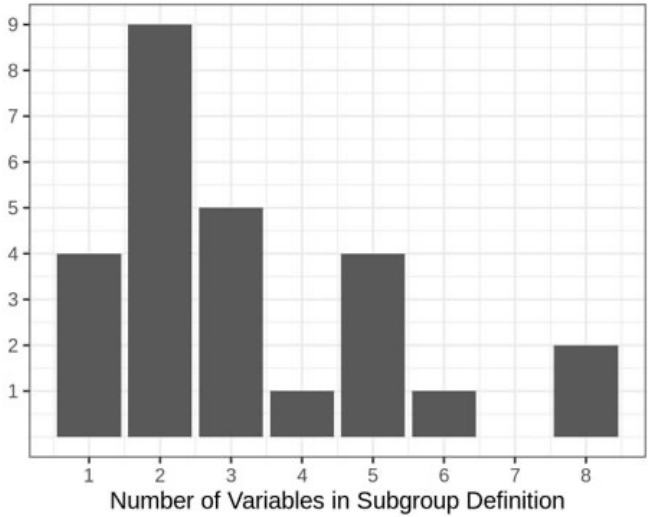
- $\hat{\delta}$ = treatment effect observed in the subgroup in the new trial
- The score is the log-likelihood of $\hat{\delta}$ according to N(δ_{pred}, σ_{pred})



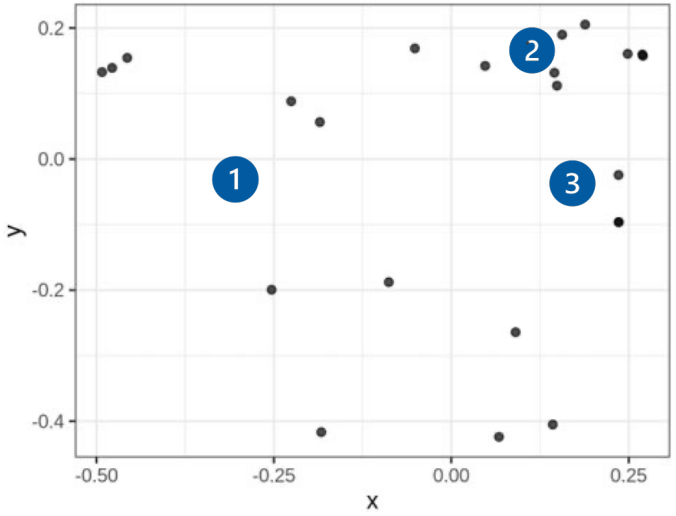
Treatment difference in subgroup

Challenge results

Typically few covariates used in subgroup definition
A few variables were used very often (e.g. age)



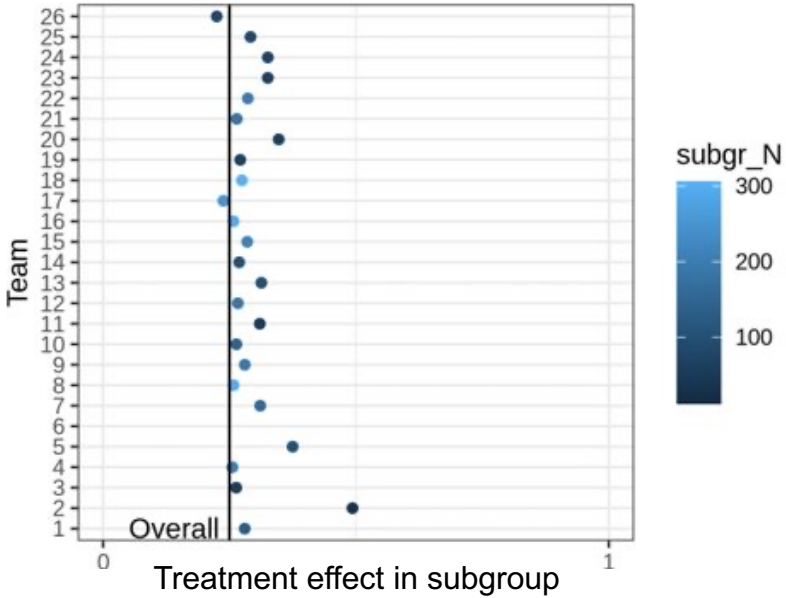
Proposed subgroups are quite different, also the top 3



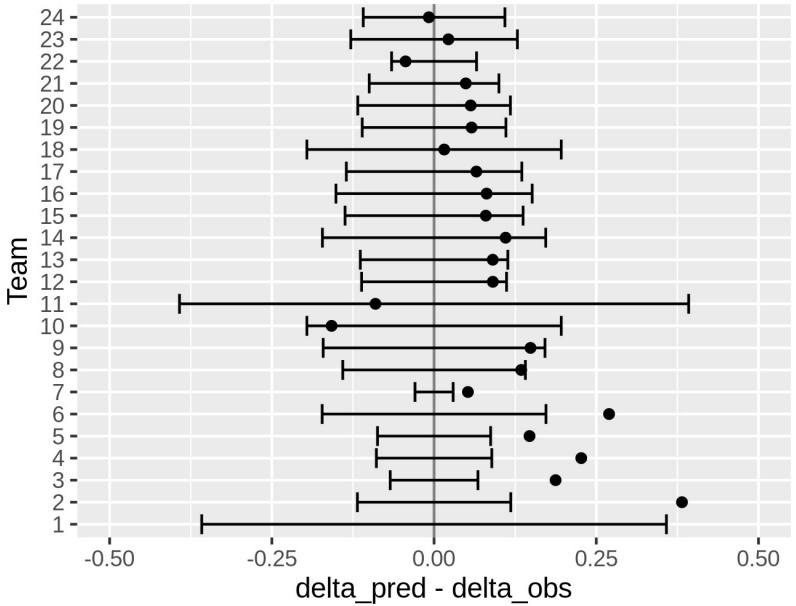
NOVARTIS | Reimagining Medicine

Challenge results

All but two teams identified a subgroup with tr. eff. higher than overall in the new trial



Systematic overestimation of treatment effect (regression to the mean)



NOVARTIS | Reimagining Medicine

Methods used by the participants

Focusing on top teams:

Penalized-regression
to identify
predictive covariates

- LASSO regression
- Dose-response model with shrinkage priors

Find threshold

Try different combinations of thresholds, evaluate best by bootstrap

Subgroup definition

Tree-based

MOB

Subgroup definition

Learnings and insights from the participants

HARD PROBLEM

- No obvious solution
- Various methods could identify different subgroups with increased tr. eff.

REPRODUCIBILITY

- Cross-validation with hold-out sets
- Clinical insights

DATA NOT SUFFICIENT?

- Even with the best method, need to complement data with external information

Many thanks to

- The other organizers: Bjorn Bornkamp, Carsten Müller, Conor Moloney, Mark Baillie, Michela Azzarito, Ruvie Martin, Jana Starkova
- The participants of the challenge
- data42
- All of you for listening!