The Effect Of Nasal Mucus Viscosity On Annual Nosebleed Rates

Data Science Group Technical Presentation
Bayer AG

Dr. Marcel Preising

August 16th, 2018

Scenario Introduction

- Bayer is in-licensing a new treatment (superdupripine) for unmet clinical need of recurrent serious nosebleeds.

- Bayer is considering whether or not to fund a large-scale Phase III clinical trial.

$\Longrightarrow$ **Objectives:**

- Detect whether a treatment effect of superdupripine on counted number of nosebleeds exists.

- Effectiveness of superdupripine depending on further predictors (viscosity of nasal mucus, amount of paper tissues, country)?

- How to predict nosebleeds?

- How to implement Phase III simulation?

Data Set Description

- Number of individuals: 444
  - → Patients in treatment group: 223
  - → Patients in placebo group: 221

- Number of variables: 8

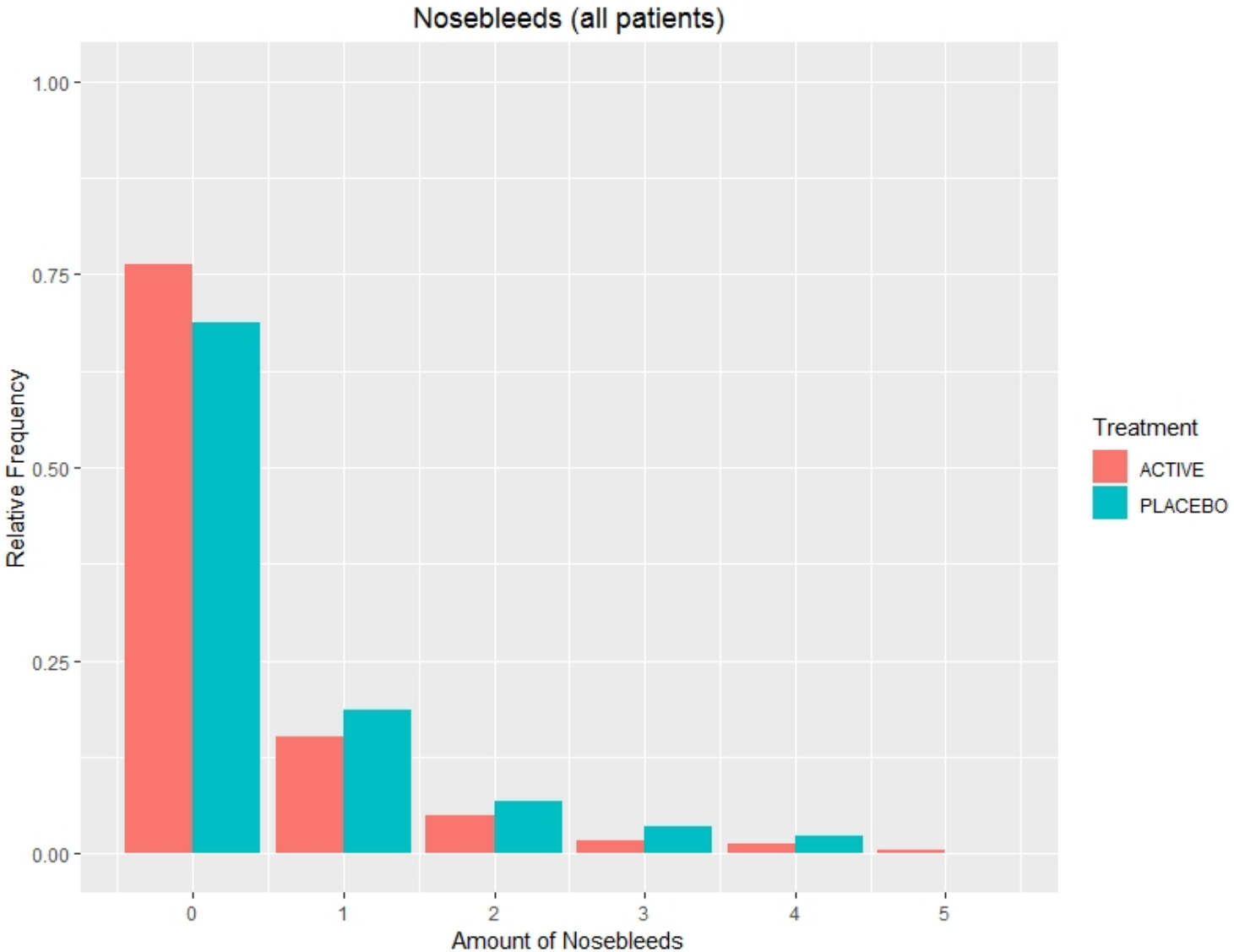
Variable name	Description
arm	Assigned treatment arm (Active/Placebo)
country	Country of the subject (A to J)
eye.colour	Eye colour (Black/Blue/Brown)
mucus.viscosity	Nasal mucus viscosity
tissue.use	Amounts of tissues subject uses (High/Medium)
previous.year	# of nosebleeds requiring hospitalization last year
nosebleeds	# of nosebleeds observed on study
duration	Time that the subject was on the study (Days)

Treatment effect of superdupripine I

- Treatment effect detectable due to superdupripine on nosebleeds?

- Compare whether distribution of Active group differs from that of Placebo group.
  - → Graphical hints.
  - → Statistical tests.

Treatment effect of superdupripine II



Treatment effect of superdupripine III

- T-Test: Significant difference in means (90%-level) of Active and Placebo Group concerning amount of nosebleeds:

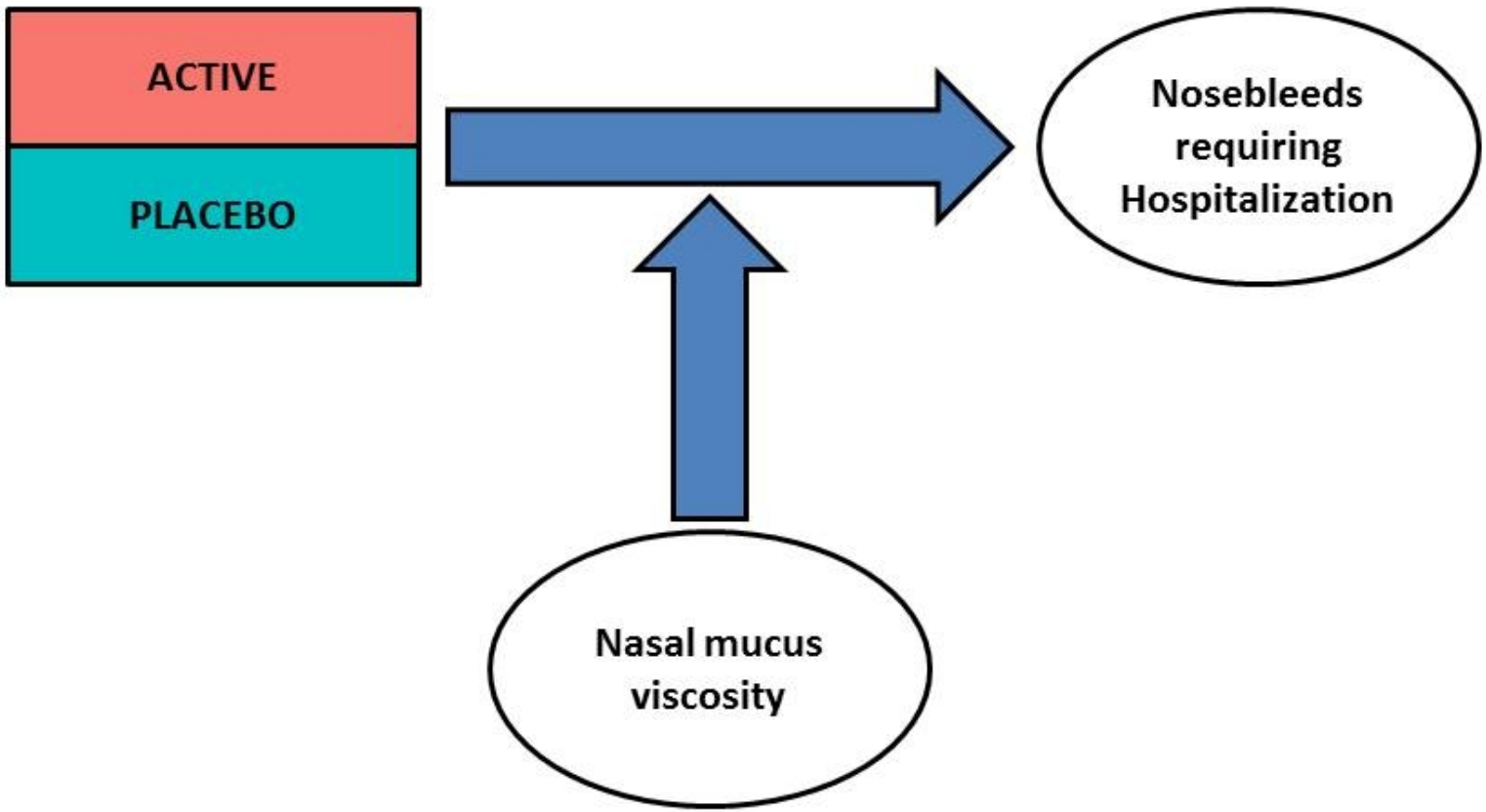
$$\begin{aligned} \text{Mean amount of nosebleeds (ACTIVE):} &= 0.38 \\ \text{Mean amount of nosebleeds (PLACEBO):} &= 0.52 \\ \text{Difference} &= -0.14 \end{aligned}$$

- Poisson regression: Coefficient confirms reduction of amount of nosebleeds due to treatment: -0.36 (significant on 95%-level).

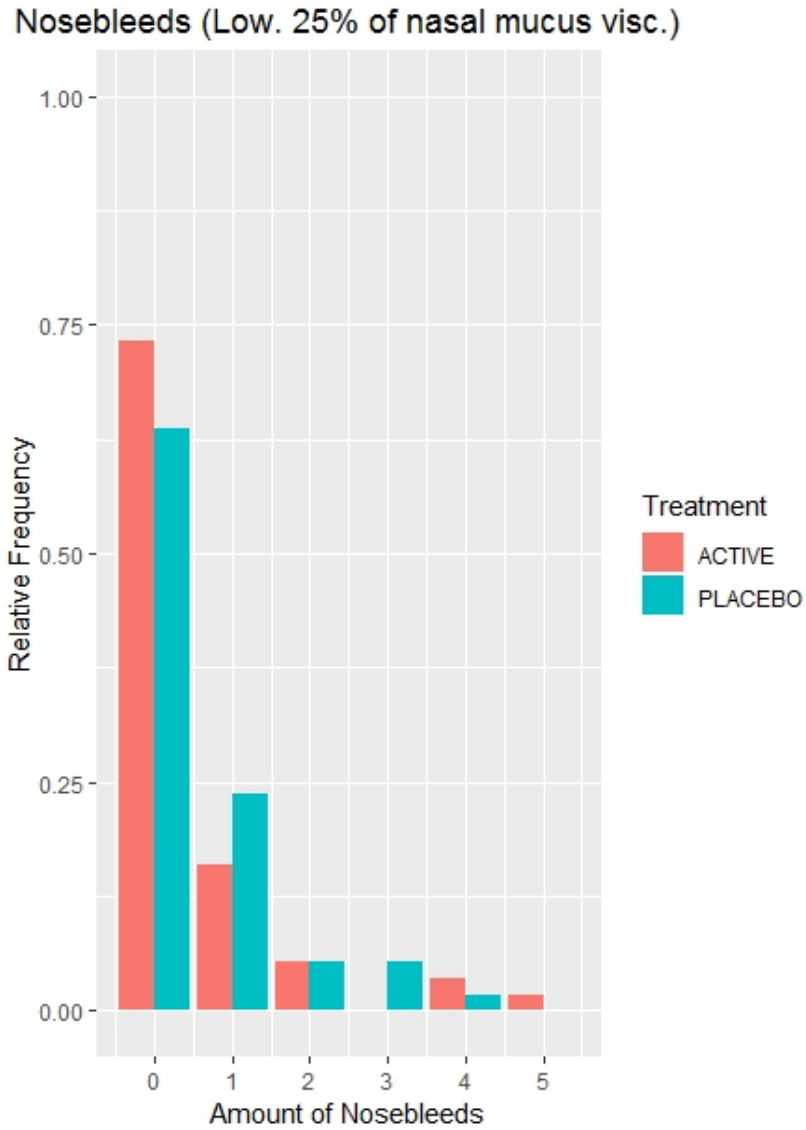
$\Longrightarrow$ **Treatment leads to less expected amounts of nosebleeds!**

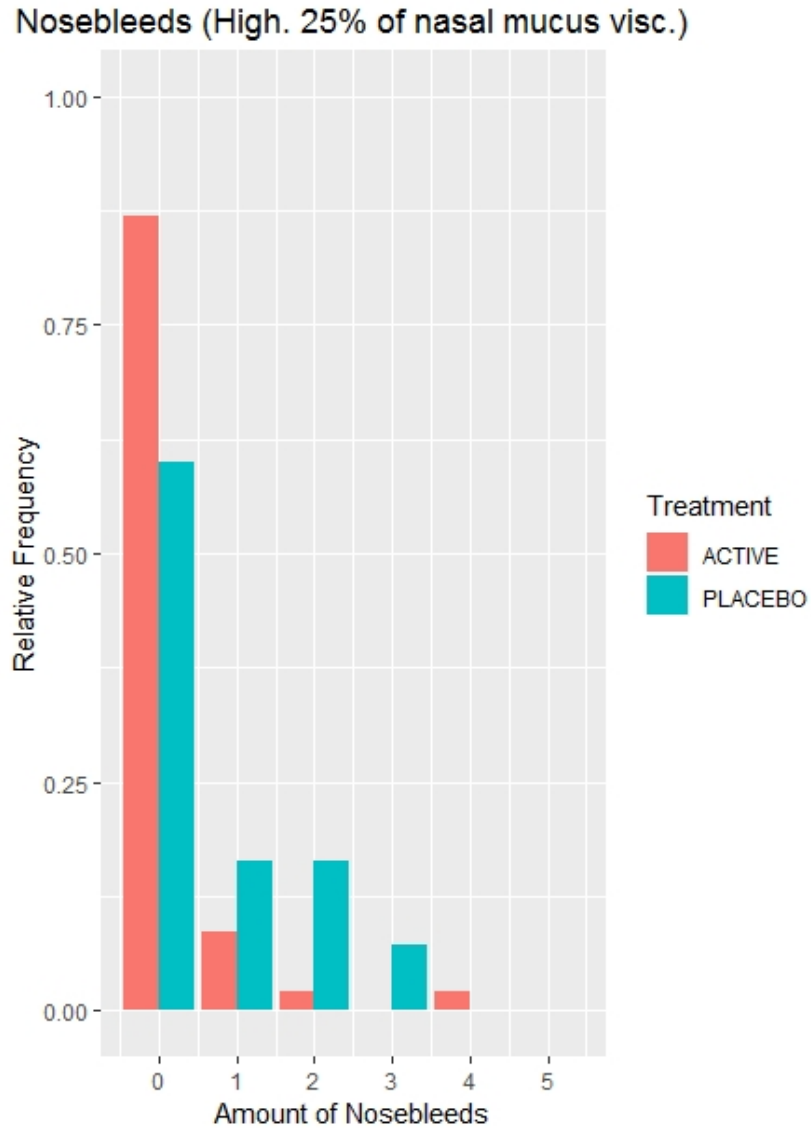
Effect of nasal mucus viscosity I

- Nasal mucus viscosity as predictor of the effect of superdupripine?



Effect of nasal mucus viscosity II





Effect of nasal mucus viscosity III

→ T-tests for mean differences of nosebleeds:

Mucus visc.	All patients	Low. 25%	High. 25%	High. 10%
Active	0.38	0.50	0.22	0.09
Placebo	0.52	0.58	0.71	0.81
Difference	-0.14*	-0.08	-0.49**	-0.73**

Significance level: *:p<0.10, **:p<0.05.

→ Poisson regression estimating effects of superdupripine on nosebleed:

Mucus visc.	All patients	Low. 25%	High. 25%	High. 10%
Coefficient	-0.31**	-0.06	-1.15**	-2.14**

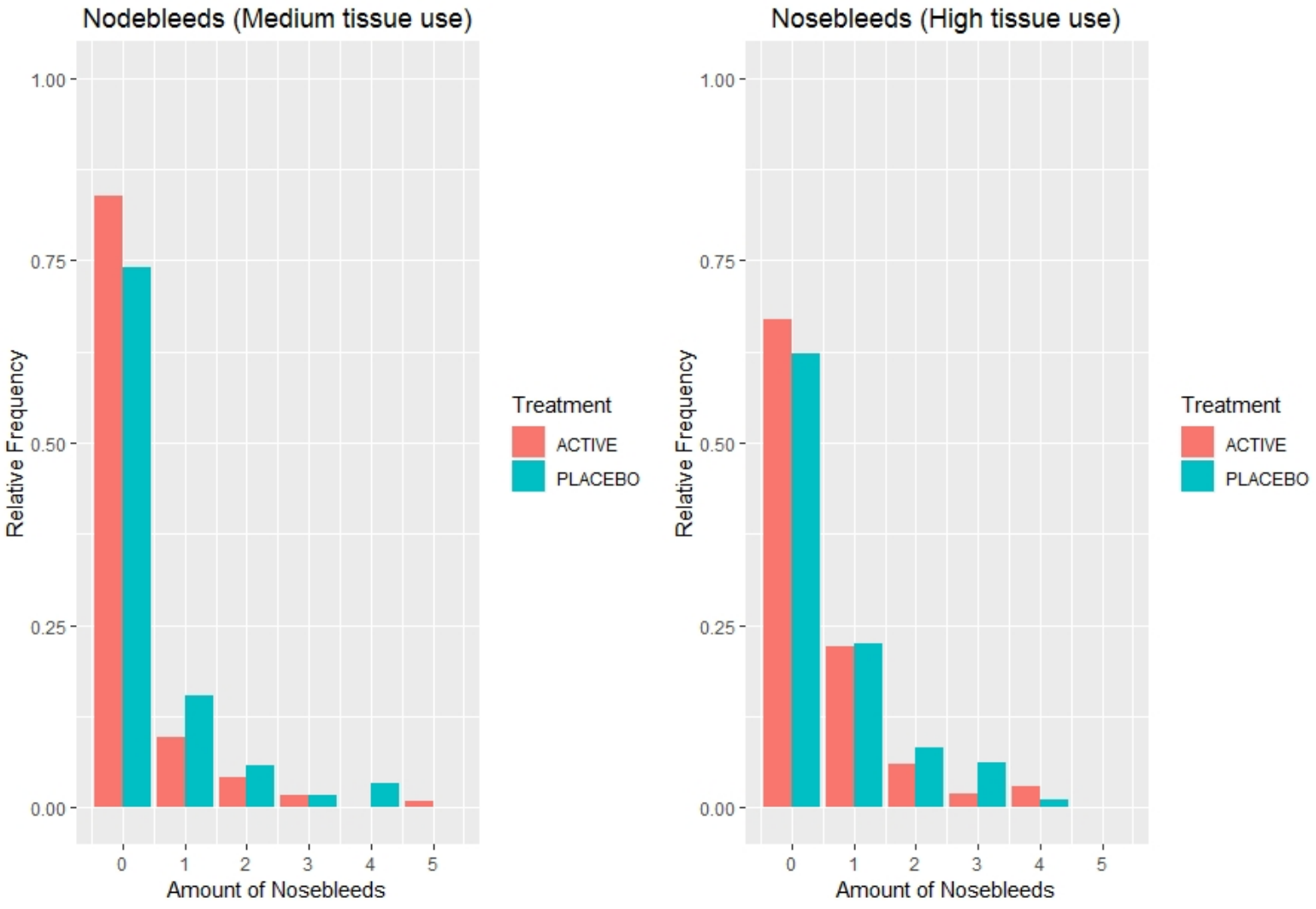
Significance level: **:p<0.05.

⟹ **The higher the nasal mucus viscosity, the higher the superdupripine effect!**

Effect of paper tissues use I

- "Most medically serious cases are those where the patient buys a large amount of paper tissues."
- Differing distributions for effects by superdupripine on nosebleeds for patients with medium and high paper tissues demand?

Effect of paper tissues use II



Effect of paper tissues use III

→ T-tests for mean differences of nosebleeds:

Tissue use	Medium	High
Active	0.27	0.52
Placebo	0.45	0.61
Difference	-0.18*	-0.09

Significance level: *:p<0.10.

→ Verify again by Poisson regression coefficients estimating the effects of superdupripine on nosebleed:

Tissue Use	Medium	High
Coefficient	-0.51**	-0.16

Dependent variable: nosebleeds.

Significance level: **:p<0.05.

⟹ **High tissue use does not support higher treatment effect by superdupripine.**

Local medical practices as predictor of treatment I

- Nosebleeds differing across countries?
- Treatment effects different in particular countries?

Local medical practices as predictor of treatment II

- Countries differ in average nosebleeds significantly.

$\Longrightarrow$ Analysis of variance confirms differences of mean nosebleeds over countries.

$\Longrightarrow$ Poisson regression reveals significant country-specific effects on nosebleeds:

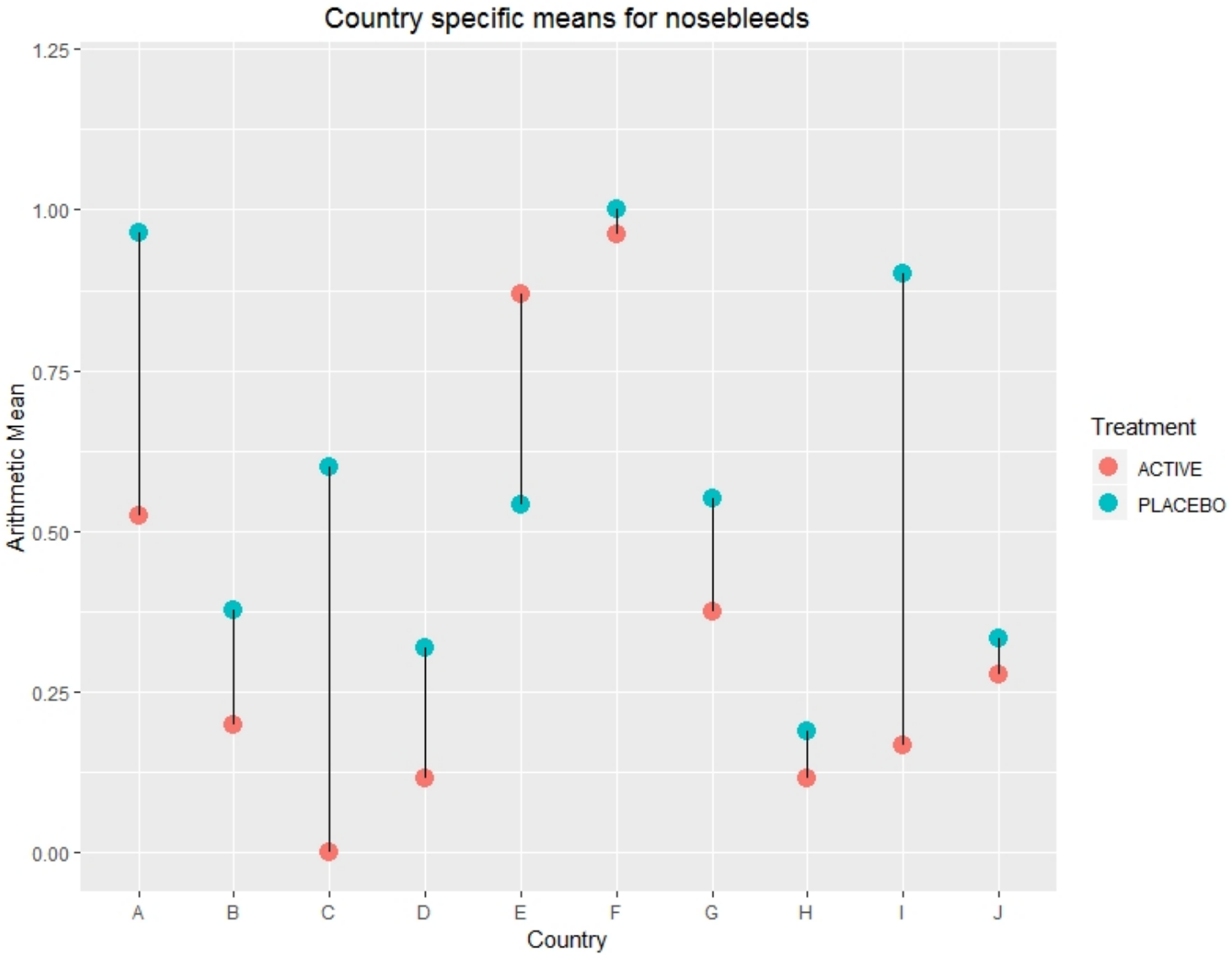
Country	B	D	G	H	J
Coefficient	-0.98	-1.32	-0.50	-1.63	-0.95

Dependent variable: nosebleeds. Predictor: country.

Reference category: Country A.

Only significant regression coefficients indicated (95%-level).

Local medical practices as predictor of treatment III



Prediction of nosebleed rates I

- Possible selection of statistical methods and models:

$\Longrightarrow$ Parametric statistical models, e.g.

- Poisson regression model.
- Zero-inflated poisson regression: *zeroinfl()* provided in R package *pscl*.

$\Longrightarrow$ Non-parametric statistical models, e.g.

- Classification and regression trees (e.g. Packages *rpart, tree*).
- Random forests (e.g. Package *randomForest*).

Prediction of nosebleed rates II

Prediction Steps:

1. Regress nosebleed on predictors:
   → *zeroinfl(nosebleeds ~ arm + country +...,data=Data)*

2. Use prediction function with argument *type="prob"* to predict probability of nosebleeds amount for new data:
   → *predict(newdata, type="prob")*

Nosebleeds	0	1	2	3	4	5
Patient						
1	0.65	0.15	0.11	0.06	0.02	0.00
2	0.46	0.32	0.16	0.05	0.01	0.00
3	0.59	0.18	0.13	0.06	0.02	0.0

The simulation of phase three I

- Gain distributional properties and dependencies within and across variables for patients with focus on desired characteristics:
  - → Patients with high nasal mucus viscosity.
  - → High demands for paper tissues.
  - → Countries with beneficial medical practices.

The simulation of phase three II

- Simulate new covariate data.
- Draw treatment variable randomly (simulate coin flip).
- Use regression coefficients from Phase II and *predict()*.
  $\Longrightarrow$ Vary uncertainty by stochastic components.
  $\Longrightarrow$ Vary predictor's influences, internal variability, and correlations.
- Simulate high amounts of synthetic data sets.
  $\Longrightarrow$ Analyze estimator properties!
  (Bias, confidence intervals, efficiency, prediction intervals)

The simulation of phase three II

- Vary sampling properties: e.g. detect properties of analysis models for small sample sizes.
- Include e.g. random coefficients to account for individual-specific heterogeneity in predictor's effects.
- Use panel models:
  - $\Longrightarrow$ Time-specific coin flip to decide whether patients is part of treatment or placebo group.
  - $\Longrightarrow$ Account for latent heterogeneity within individuals (random or fixed effect).

$\Longrightarrow$ **Work together with project team to decide which modifications the treatment effects needs to withstand to justify a Phase III.**

Conclusion and outlook I

- Data reveal treatment effect by superdupripine!
- Treatment effect particularly strong for patients with high degree of nasal mucus viscosity.
- Higher Nosebleed rates for patients buying high amounts of paper tissues.
  $\Longrightarrow$ However, effect of superdupripine NOT stronger for patients with high paper tissue demand.
- Effects increased for certain countries.
  $\Longrightarrow$ Project team knows about certain local medical practices?

Conclusion and outlook II

- Conduct Phase III Monte Carlo simulation to supervise for estimation robustness and behavior of treatment effect under varying stochastic influencing factors.

$\Longrightarrow$ Simulation study helps to decide whether to fund and how to optimize Phase III!

Thank you for your attention!