Model Partitioning with an Application to Subgroup Identification

Heidi Seibold, Achim Zeileis and Torsten Hothorn

Epidemiology, Biostatistics and Prevention Institute

University of Zurich

University of Zurich[UZH]

June 16, 2015

Subgroup analyses

Identifying groups of patients for whom the treatment has a different effect than for others.

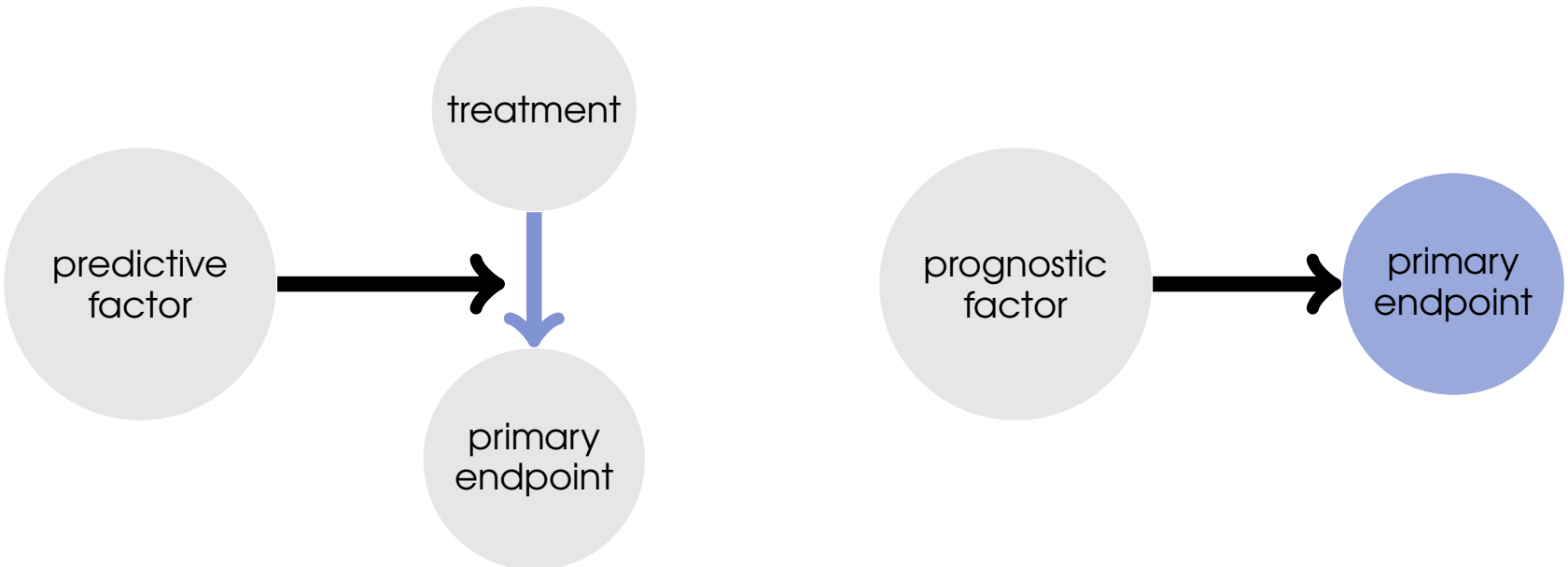
Effect is:

- Stronger
- Lower
- Contrary

Subgroup analyses

Goal: Find predictive factors

$\hat{=}$ covariate $\times$ treatment interactions



Trees are great for detecting interactions. Combine trees with models (knowledge about treatment effect).
$\Rightarrow$ Model-based recursive partitioning.

Data

PRO-ACT database (1)

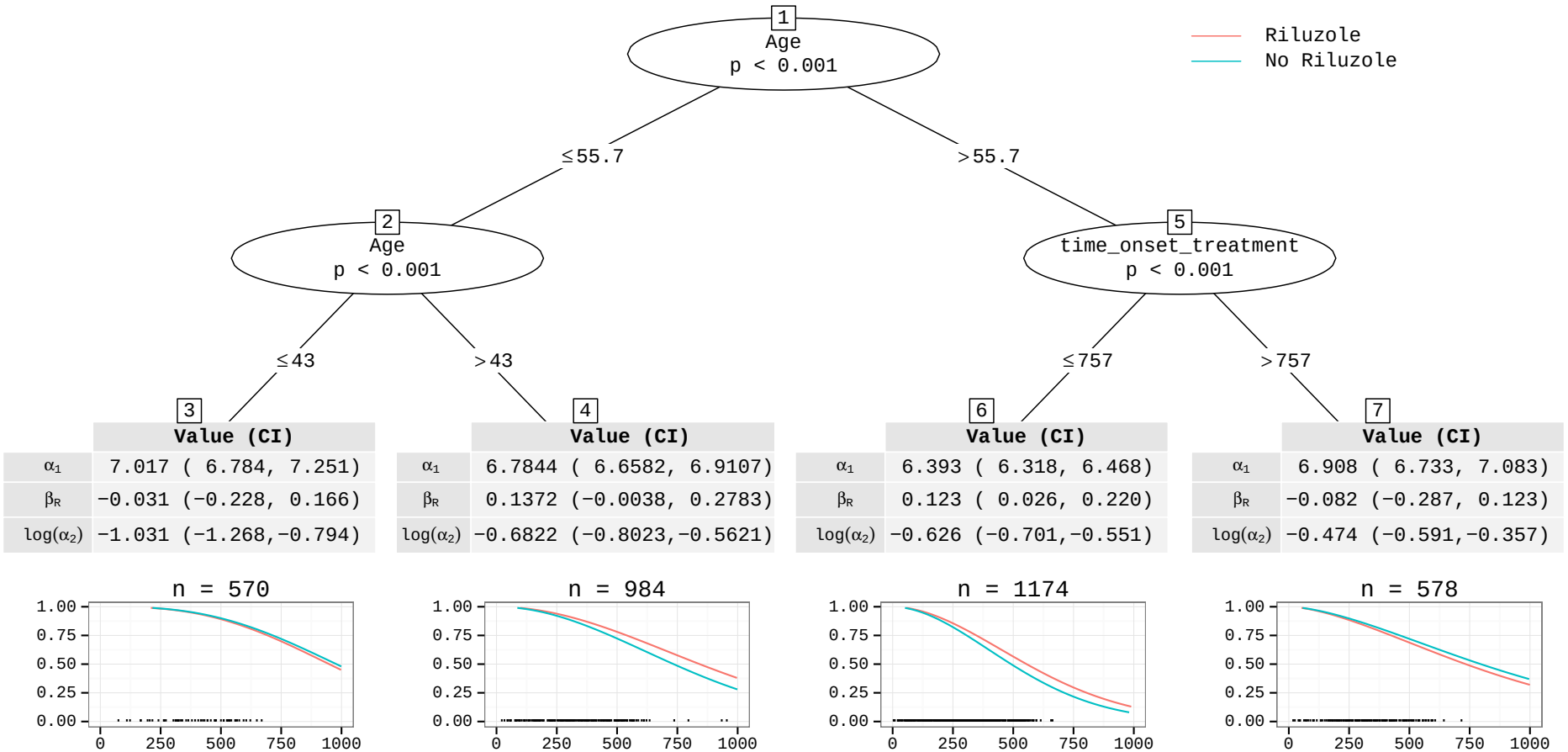
- Amyotrophic lateral sclerosis (ALS) patients
- Data of several clinical trials
- Treatment of interest: Riluzole
- Primary endpoints of interest: Survival time

Riluzole modestly prolongs life expectancy

But: Are there any groups of patients for whom it is better or worse?

$\Rightarrow$ Subgroup analyses

Subgroups with differential Riluzole effect on the survival



	Value (CI)
α_1	7.017 (6.784, 7.251)
β_R	−0.031 (−0.228, 0.166)
$\log(\alpha_2)$	−1.031 (−1.268,−0.794)

	Value (CI)
α_1	6.7844 (6.6582, 6.9107)
β_R	0.1372 (−0.0038, 0.2783)
$\log(\alpha_2)$	−0.6822 (−0.8023,−0.5621)

	Value (CI)
α_1	6.393 (6.318, 6.468)
β_R	0.123 (0.026, 0.220)
$\log(\alpha_2)$	−0.626 (−0.701,−0.551)

	Value (CI)
α_1	6.908 (6.733, 7.083)
β_R	−0.082 (−0.287, 0.123)
$\log(\alpha_2)$	−0.474 (−0.591,−0.357)

MOB Basics

MOB: Model-based recursive partitioning (2)

Start with model $\mathcal{M}((Y, \mathbf{X}), \boldsymbol{\vartheta})$ with

$$\boldsymbol{\vartheta} = \begin{pmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta} \\ \boldsymbol{\gamma} \\ \boldsymbol{\nu} \end{pmatrix} \begin{matrix} \text{intercept(s)} \\ \text{treatment effect} \\ \text{other parameter(s) of interest} \\ \text{nuisance parameter(s),} \end{matrix}$$

which fits data $(Y, \mathbf{X})$.

Model

Parametric model: **Weibull model**

$$\mathbb{P}(Y \leq y | X = x) = F\left(\frac{\log(y) - \alpha_1 - \beta x_R}{\alpha_2}\right)$$

y (right-censored) survival time,
F cumulative distribution function of Gompertz distribution

$\boldsymbol{\alpha} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}$ defines the shape of the baseline hazard

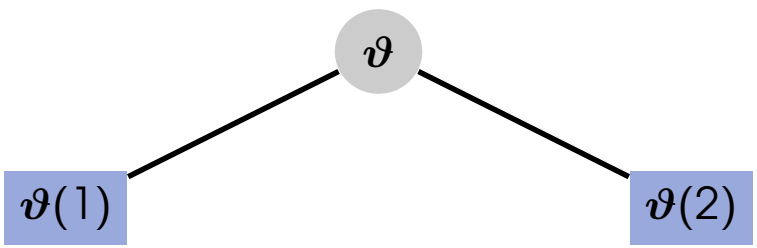
$\Rightarrow$ use as "intercept"

MOB Basics

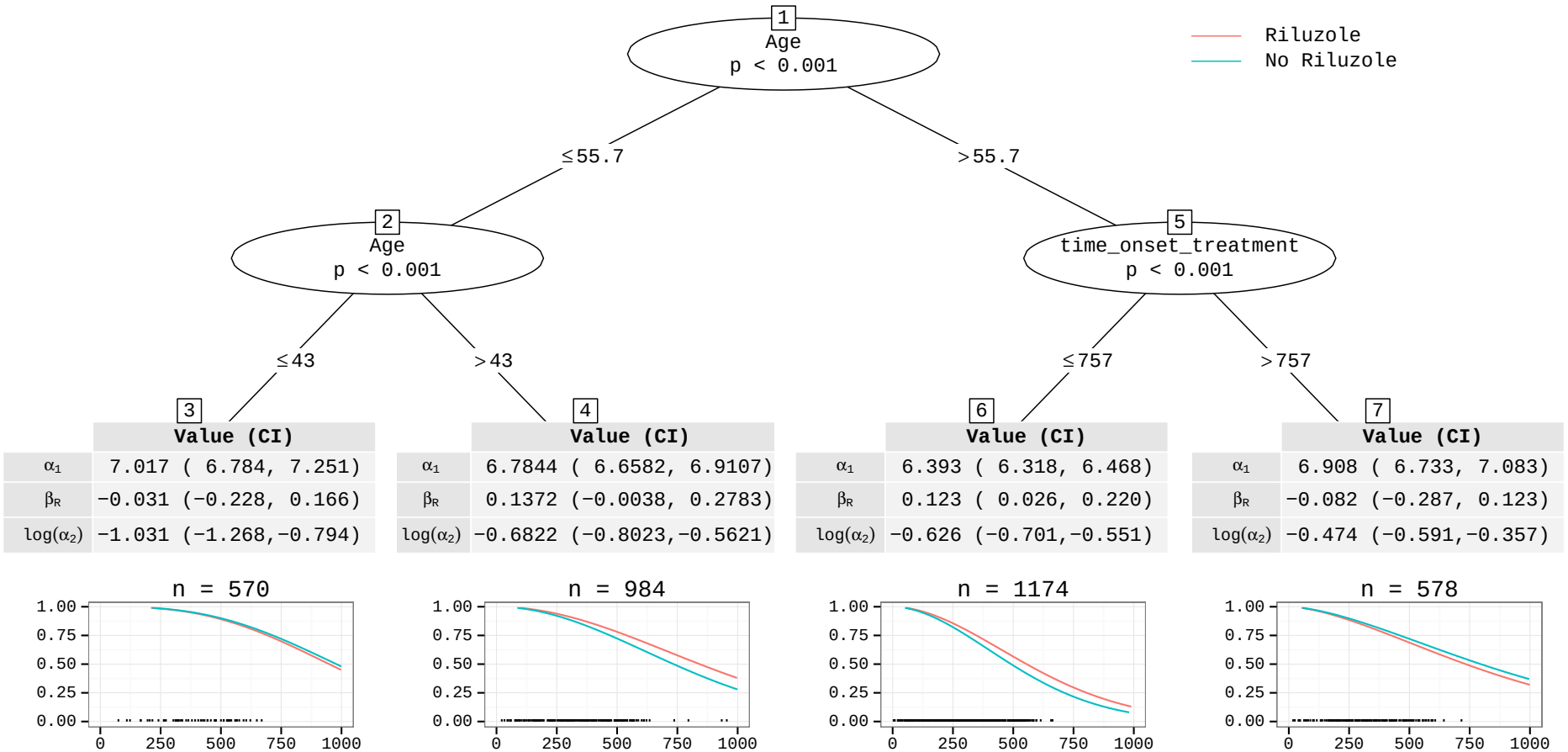
Maybe the treatment effect is not the same for all patients, but depends on their characteristics.

$\Rightarrow$ Find partitions $\{\mathcal{B}_b\}$ $(b = 1, ..., B)$ based on patient characteristics $\mathbf{Z} = (\mathbf{Z}_1, ..., \mathbf{Z}_J) \in \mathcal{Z}$

$\Rightarrow$ Fit separate models $\mathcal{M}((Y, \mathbf{X}), \vartheta(b))$ in partitions.



Subgroups with differential Riluzole effect on the survival



	Value (CI)
α_1	7.017 (6.784, 7.251)
β_R	−0.031 (−0.228, 0.166)
$\log(\alpha_2)$	−1.031 (−1.268,−0.794)

	Value (CI)
α_1	6.7844 (6.6582, 6.9107)
β_R	0.1372 (−0.0038, 0.2783)
$\log(\alpha_2)$	−0.6822 (−0.8023,−0.5621)

	Value (CI)
α_1	6.393 (6.318, 6.468)
β_R	0.123 (0.026, 0.220)
$\log(\alpha_2)$	−0.626 (−0.701,−0.551)

	Value (CI)
α_1	6.908 (6.733, 7.083)
β_R	−0.082 (−0.287, 0.123)
$\log(\alpha_2)$	−0.474 (−0.591,−0.357)

How to find the Partitions?

$\Rightarrow$ Test of independence between the **partial score functions** and each **patient characteristic**:

$$
\begin{aligned}
H_0^{\alpha,j}: \quad & \psi_\alpha((Y,\mathbf{X}),\hat{\boldsymbol{\vartheta}}) \quad \perp \quad \mathbf{Z}_j \\
H_0^{\beta,j}: \quad & \psi_\beta((Y,\mathbf{X}),\hat{\boldsymbol{\vartheta}}) \quad \perp \quad \mathbf{Z}_j, \quad j = 1,\ldots,J
\end{aligned}
$$

$$
\boldsymbol{\psi} = \begin{pmatrix}
\psi_\alpha((y,\mathbf{x})_1,\boldsymbol{\vartheta}) & \psi_\beta((y,\mathbf{x})_1,\boldsymbol{\vartheta}) & \cdots \\
\vdots & \vdots & \\
\psi_\alpha((y,\mathbf{x})_N,\boldsymbol{\vartheta}) & \psi_\beta((y,\mathbf{x})_N,\boldsymbol{\vartheta}) & \cdots
\end{pmatrix}
$$

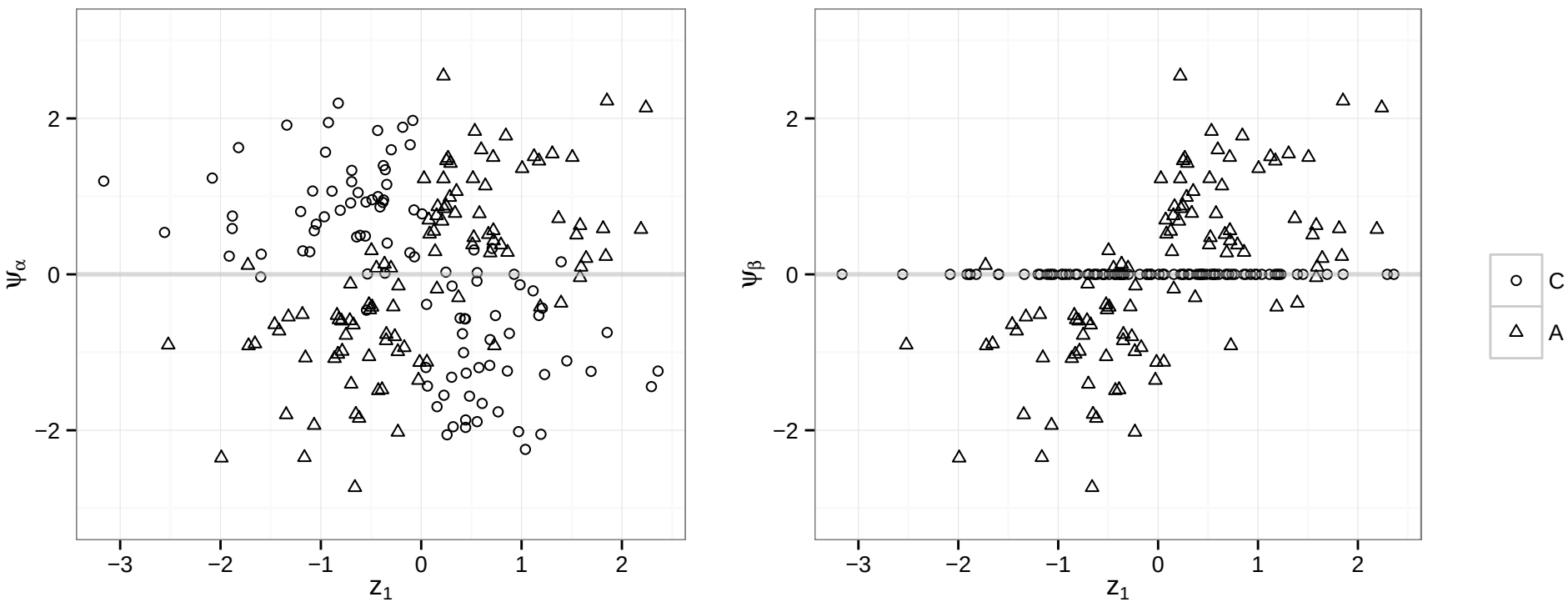
- Partition if global p-value smaller than significance level
- Use as split variable the one with the smallest p-value

Why does this work?

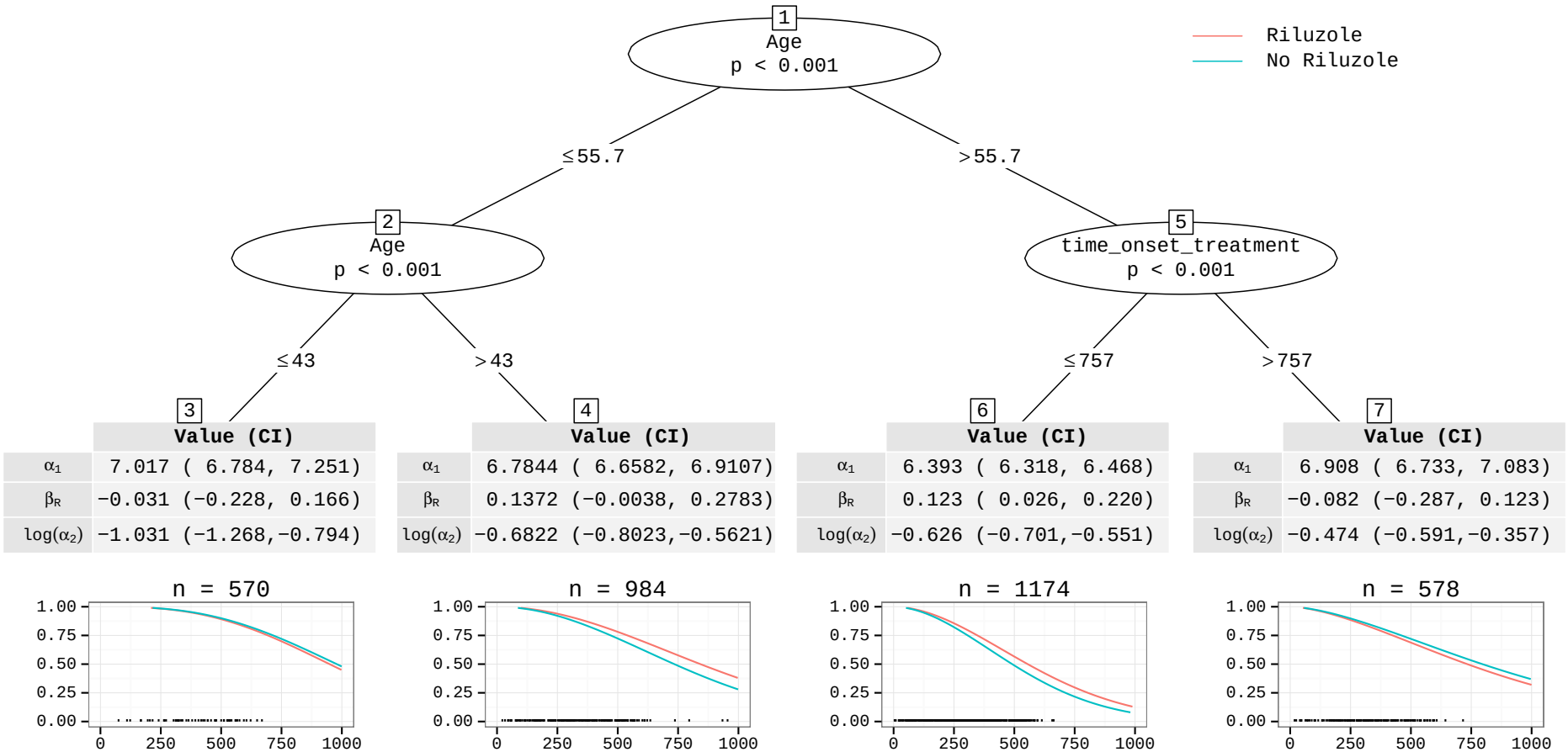
Example (3):

Data generating process:
$Y|\mathbf{X} = \mathbf{x}, \mathbf{Z} = \mathbf{z} \sim \mathcal{N}(1.9 + 0.2 \cdot x_A + 1.8 \cdot \mathbb{1}(z_1 < 0) + 3.6 \cdot \mathbb{1}(z_1 > 0) \cdot x_A,\ 0.7)$

Linear model:
$Y|\mathbf{X} = \mathbf{x} \sim \mathcal{N}(\alpha + \beta \cdot x_A,\ \sigma^2)$



Subgroups with differential Riluzole effect on the survival



	Value (CI)
α_1	7.017 (6.784, 7.251)
β_R	−0.031 (−0.228, 0.166)
$\log(\alpha_2)$	−1.031 (−1.268,−0.794)

	Value (CI)
α_1	6.7844 (6.6582, 6.9107)
β_R	0.1372 (−0.0038, 0.2783)
$\log(\alpha_2)$	−0.6822 (−0.8023,−0.5621)

	Value (CI)
α_1	6.393 (6.318, 6.468)
β_R	0.123 (0.026, 0.220)
$\log(\alpha_2)$	−0.626 (−0.701,−0.551)

	Value (CI)
α_1	6.908 (6.733, 7.083)
β_R	−0.082 (−0.287, 0.123)
$\log(\alpha_2)$	−0.474 (−0.591,−0.357)

Summary

- MOB partitions a large class of models suitable for the treatment effect on the primary endpoint of interest.
- Score functions capture instabilities and thus help to identify predictive and prognostic variables.
- Easy interpretation and good to communicate to clinicians.
- Coming soon: model-based forest for personalised medicine.

Literature

Massachusetts General Hospital, Neurological Clinical Research Institute
Pooled Resource Open-Access ALS Clinical Trials Database
`https://nctu.partners.org/ProACT/`, 2013.

A. Zeileis, T. Hothorn, K. Hornik
Model-Based Recursive Partitioning
Journal of Computational and Graphical Statistics, 2008.

W. Loh, X. He, M. Man
A regression tree approach to identifying subgroups with differential treatment effects
Statistics in Medicine, 2015.

Contact: heidi.seibold@uzh.ch