Tree-based approaches for censored survival data and model selection

Abdissa Negassa, PhD
Albert Einstein College of Medicine
Of Yeshiva University

OUTLINE OF PRESENTATION

1. Introduction

2. Prognostic Classification

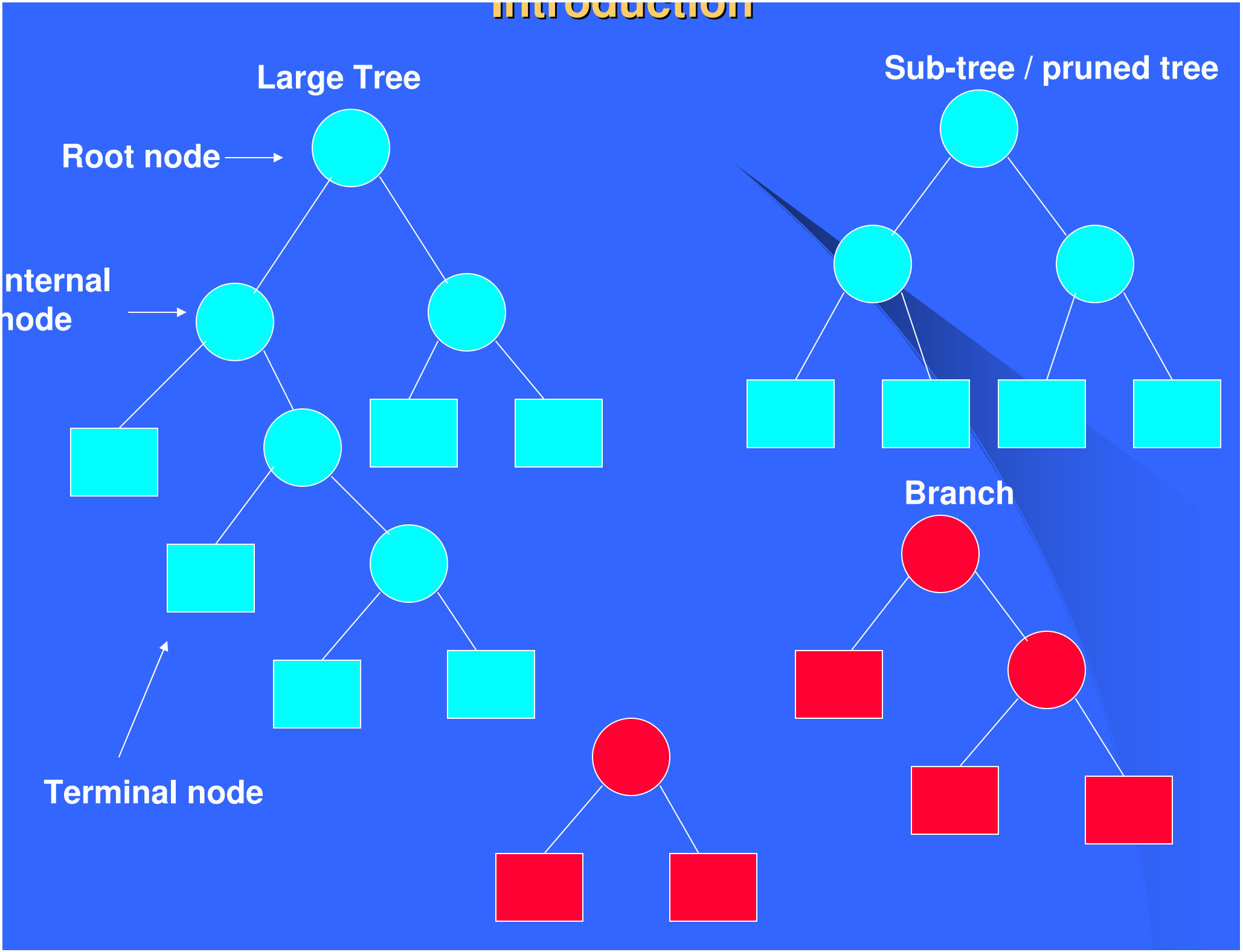
3. Previous work

4. Methods

5. Results
   - i) Simulation
   - ii) Real data set

6. Subgroup Analysis

Introduction



PROGNOSTIC CLASSIFICATION

PROBLEM:

GIVEN :

DATA = { t_i, δ_i, Z_i}

Where t is a time random variable (time to the event of interest), δ a censoring indicator, Z a vector of covariates.

FIND:

A classification of individuals with classes *homogeneous* and *distinct* with respect to survival experience, described by a tree

Gordon and Olshen (1985)
- Wasserstein distance

Davis and Anderson (1988)
- Likelihood Ratio Statistic (LRS)

Ciampi *et al* (1987, 1992, 1995)
- Log-rank statistic, Partial LRS

Segal (1988)
- Log-rank statistic

Ahn and Loh (1994)
- Patterns of Cox-residuals; two-sample *t* test

LeBlanc and Crowley (1992, 1993)
- Full likelihood, Log-rank statistic

RECPAM TREE CONSTRUCTION STEPS

Step 1. Build a binary tree

a) *Split function* : partial Likelihood Ratio Statistic (LRS)

b) examine every allowable split on the basis of simple statement on z

Step 2. Determine the *right size* tree

a) prune the large tree: construct a sequence of nested rooted subtrees based on *Information Weight*

b) choose the *"honest tree"*

Step 3. *Amalgamate* successively the leaves of

INFORMATION MEASURES WITHIN RECPAM

1) *Information Content (IC)* at a node

LRS comparing the models

$$h_1(t; Q(z)) = \exp\{\gamma Q(z)\} h_0(t)$$

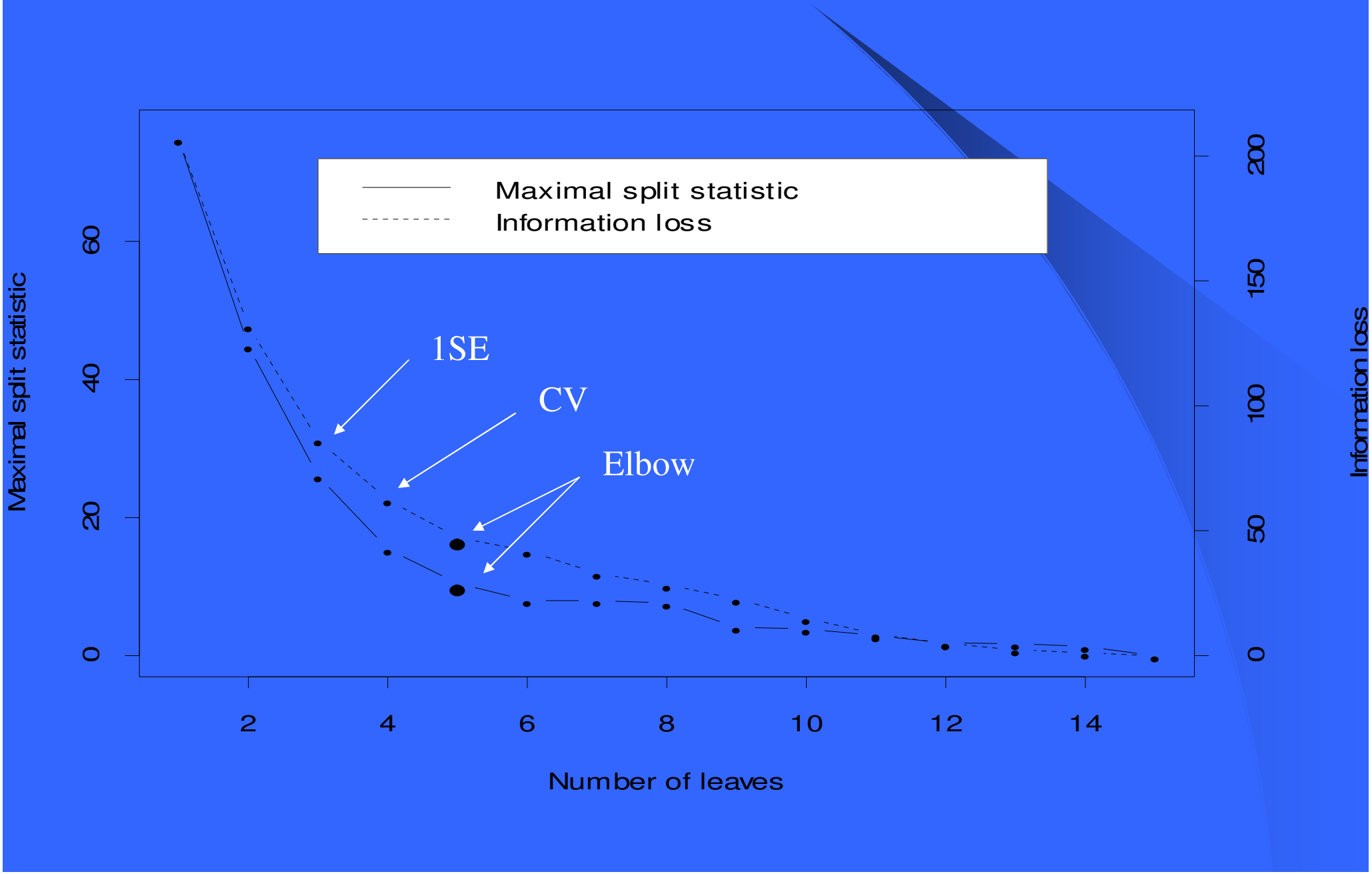
and

$$h_1(t; Q(z)) = h_0(t)$$

2) *Information Weight (IW)* of an internal node g

Information Loss (IL) of a subtree T - Tg with respect to the large tree T:

Illustration of the elbow rule using maximal split statistic and information loss



ASSESSMENT OF PERFORMANCE

1. Percent recovery of the correct structure

2. Optimism (following the outline of Efron 1983)

$$op = I\hat{C}(T_{\max}) - IC(T_{true}) \; with \; E_F(op) = \omega_{\max}$$

$$\hat{op} = I\hat{C}(T_{\max}) - I\hat{C}(T_{sel}) \; with \; E_F(\hat{op}_{sel}) = \hat{\omega}_{sel}$$

$$bias = \hat{\omega}_{sel} - \omega_{\max}$$

$$MSE = E_F(I\hat{C}(T_{sel})) - IC(T_{true}))^2$$

$$REL = \frac{MSE - MSE^{ic}}{MSE^{zero} - MSE^{ic}}$$

Where: $MSE^{zero} = MSE$ of $I\hat{C}(T_{\max})$

$MSE^{ic} = MSE$ of the "ideal constant" estimator, $I\hat{C}(T_{\max}) - \omega_{ma}$

a selection criterion with the smallest REL is expected

SIMULATION

- 4 scenarios (0% & 50% censoring and presence / absence of underlying structure)

- 150 replications from each scenario with n=300

- survival and censoring times are generated from the exponential distribution according to :

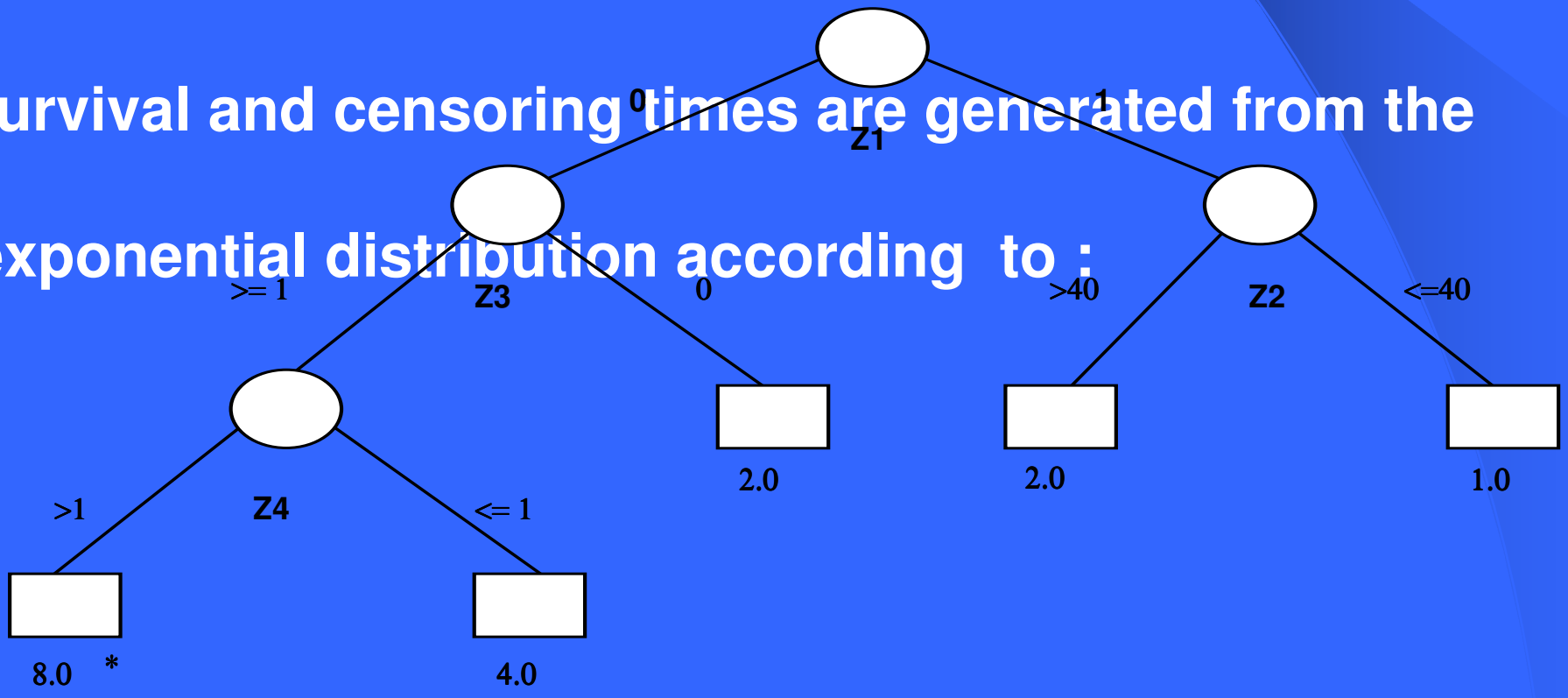
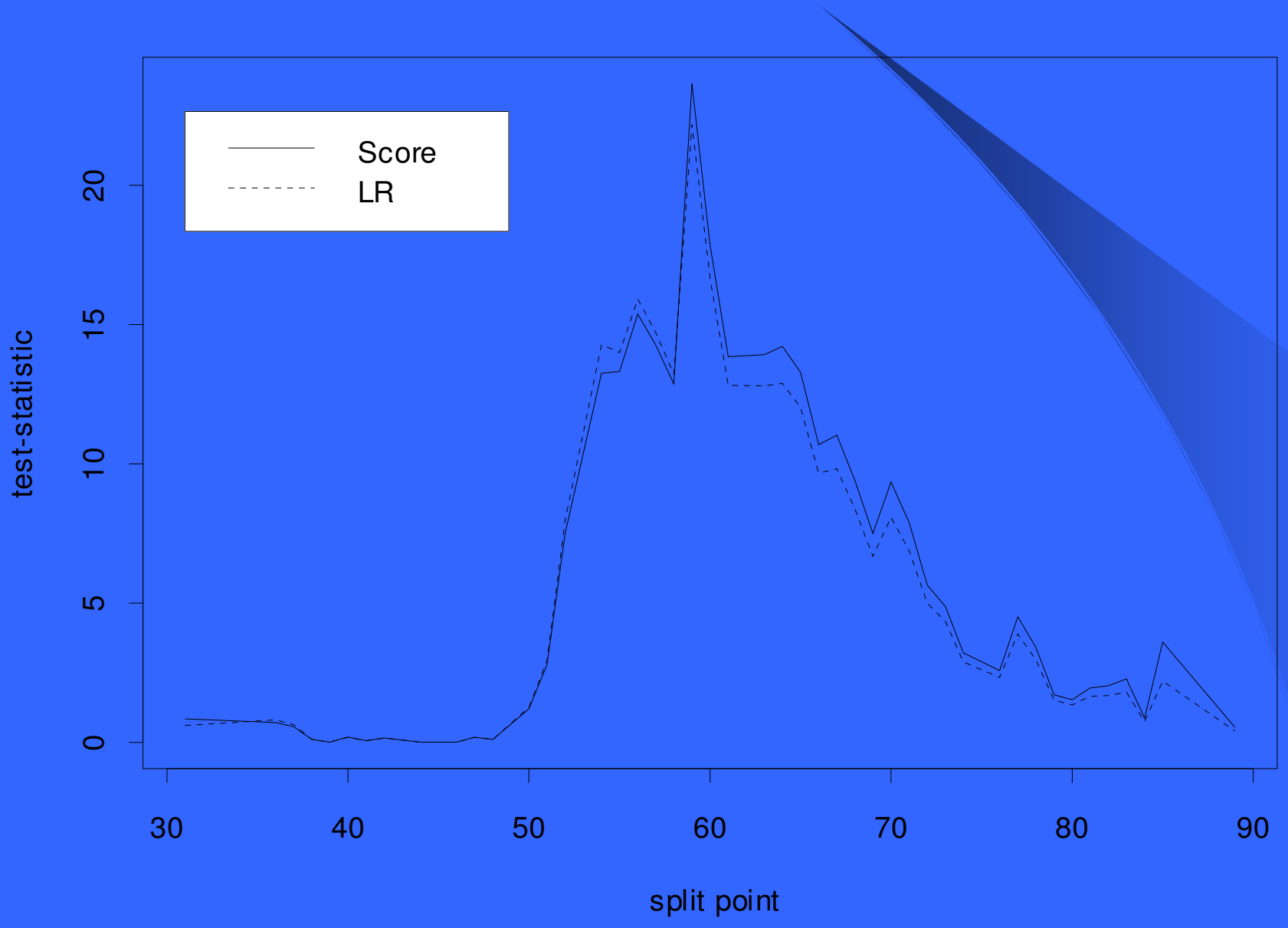


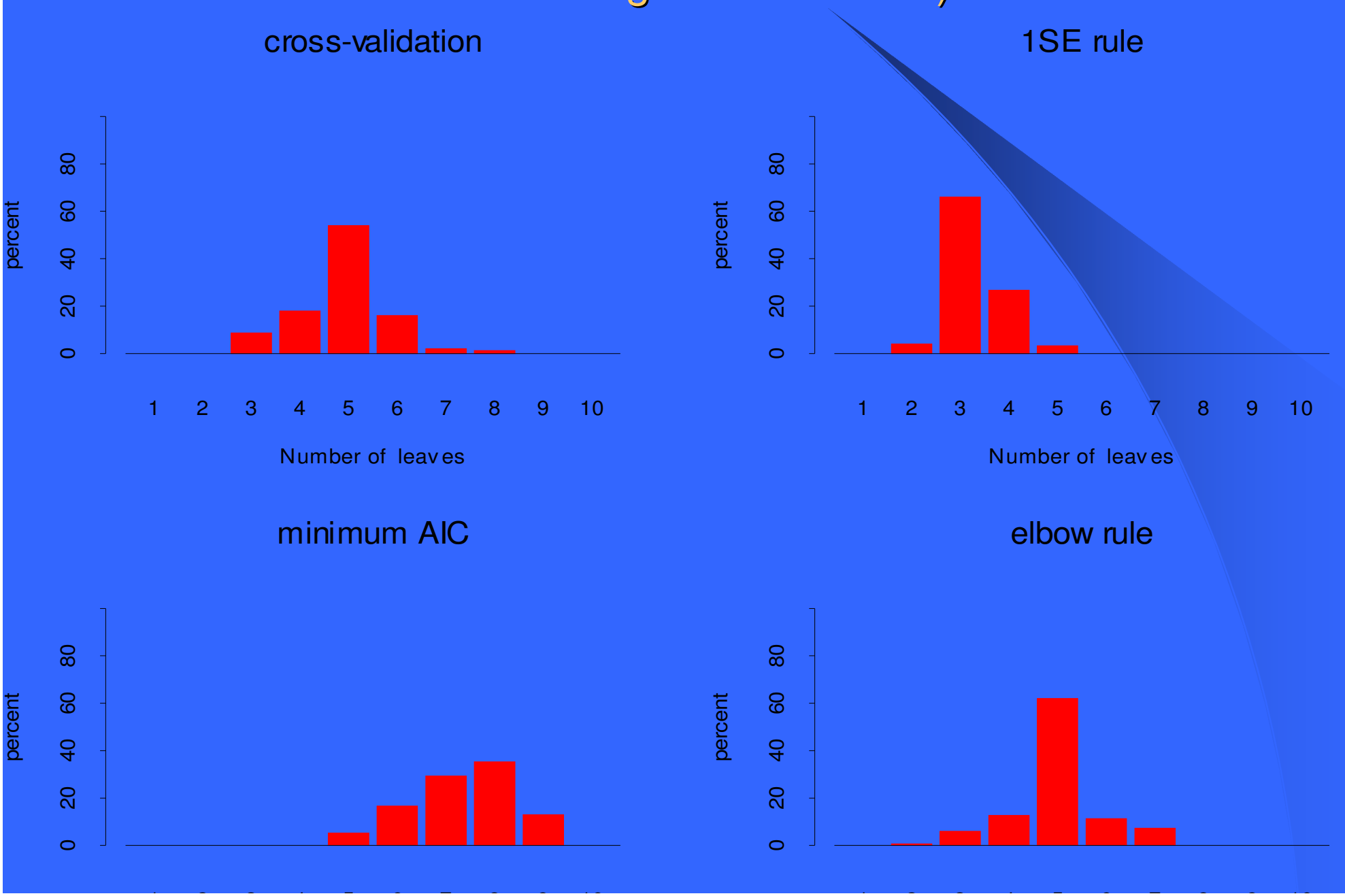
Figure 2. Split function



selection: Prognostic Classification

(The true structure has five terminal leaves and 0%

censoring in the data set)

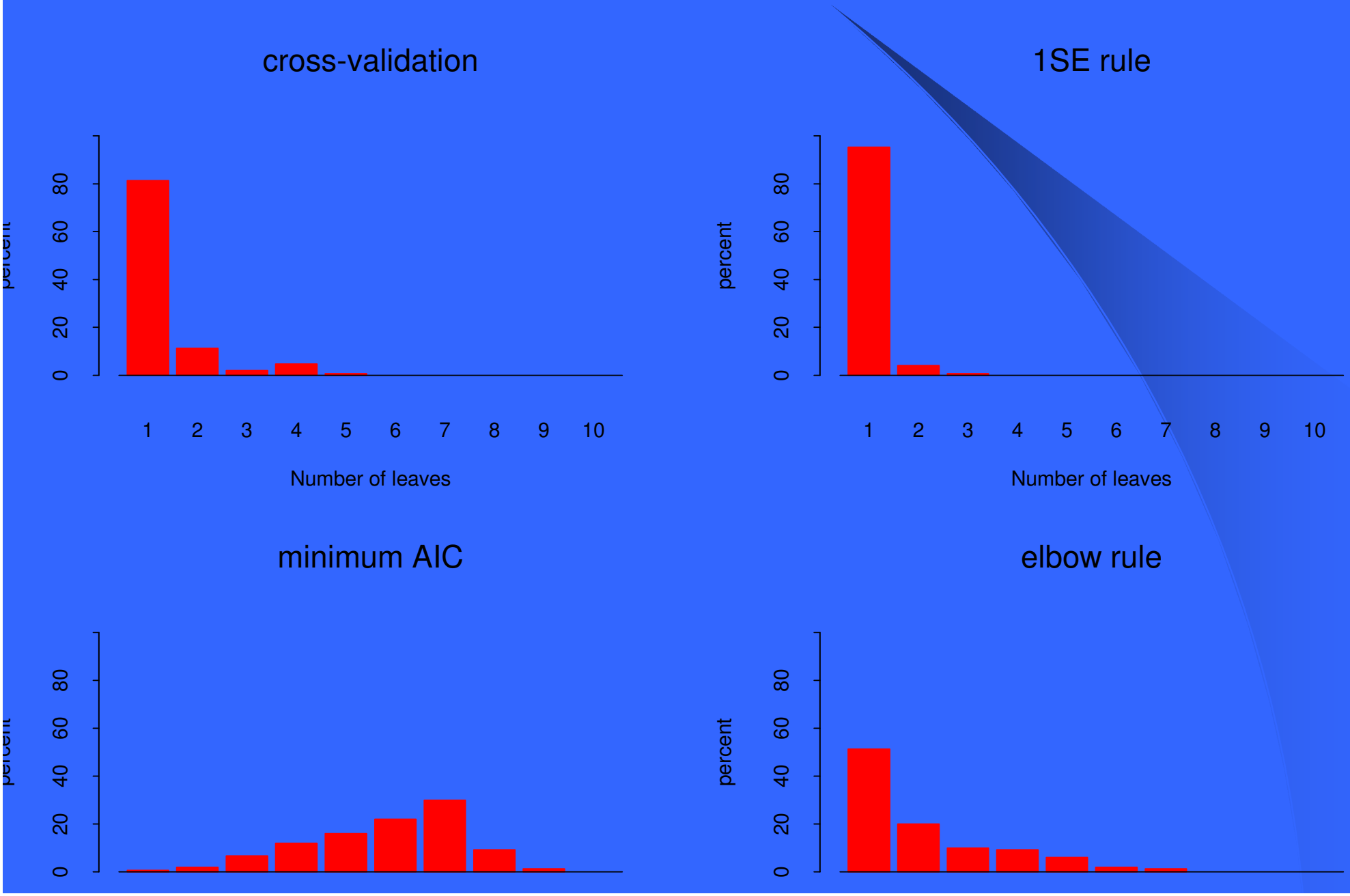


Bias, standard deviation and relative inefficiency by method of model selection: Prognostic Classification with structure

Method	Bias	$SD(o\hat{p}_{sel})$	REL
0% censoring			
CV	0.73	10.77	0.38
Elbow	-0.28	11.33	0.30
Min. AIC	-13.23	3.85	0.57
1SE rule	23.55	12.76	1.89
50% censoring			
CV	-5.23	9.31	0.38
Elbow	-7.48	6.64	0.21
Min. AIC	-18.54	3.30	0.74
1SE rule	11.14	11.23	0.45

Figure 4. Number of terminal leaves by method of tree selection: Prognostic Classification

(No structure and 0% censoring in the data set)



Bias, standard deviation and relative inefficiency by method of model selection: Prognostic Classification without structure

Method	Bias	$SD(o\hat{p}_{sel})$	REL
0% censoring			
CV	-1.67	6.88	-0.04
Elbow	-5.49	6.14	0.08
Min. AIC	-18.06	3.67	0.67
1SE rule	-0.35	6.35	-0.07
50% censoring			
CV	-2.66	6.05	0.01
Elbow	-4.99	5.12	0.08
Min. AIC	-15.52	3.43	0.65
1SE rule	-0.31	6.08	-0.08

Bias, standard deviation and relative efficiency by method of model selection: Prognostic Classification with structure

Method	Bias	$SD(o\hat{p}_{sel})$	REL
	0% censoring		
Elbow	-0.28	11.30	0.30
Two-stage with CV	-0.28	11.30	0.30
Two-stage with 1SE	-0.28	11.30	0.30
	50% censoring		
Elbow	-7.47	6.64	0.21
Two-stage with CV	-7.47	6.64	0.21
Two-stage	-6.68	8.84	0.30

Table 4
Bias, standard deviation and relative inefficiency by method of model selection: Prognostic Classification without structure

Method	Bias	$SD(o\hat{p}_{sel})$	REL
	0% censoring		
Elbow	-5.49	6.14	0.08
Two-stage with CV	-1.56	7.02	-0.03
Two-stage with 1SE	-0.65	6.73	-0.05
	50% censoring		
Elbow	-4.99	5.12	0.08
Two-stage with CV	-2.42	5.94	0.01
Two-stage	-0.24	6.07	-0.08

Acute Lymphoblastic Leukemia (ALL) Data Set

N=2725

Median follow-up time was 2017 days

66.1% censoring

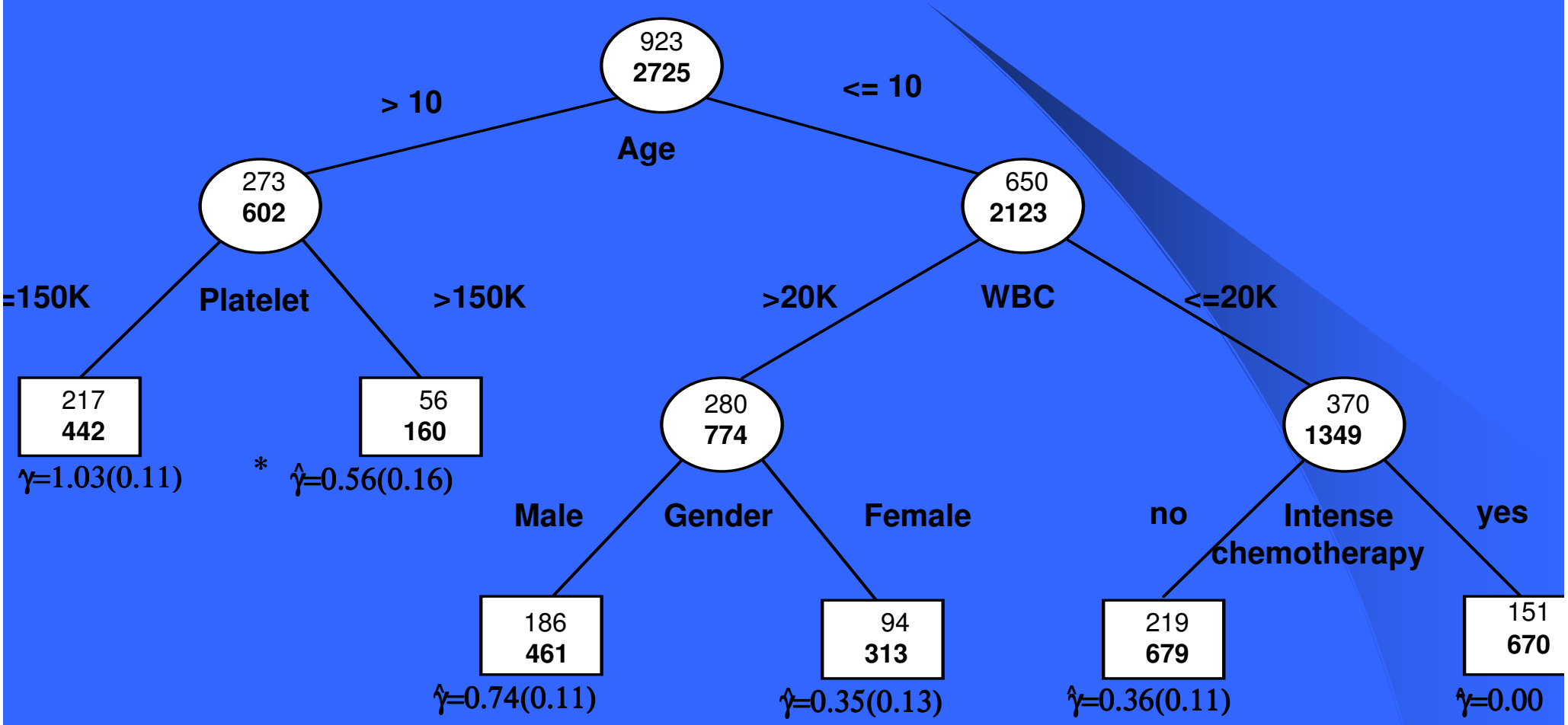
Eighteen covariates

Event Free Survival = number of days from study entry to the first *major* study event or time to last follow-up.

- Failure to achieve remission in the initial treatment phase, i.e., induction therapy phase
- Death during induction without achieving remission
- Relapse after achieving remission
- Death during remission

Figure 5

Prognostic Classification for ALL



* Cox regression coefficient (SE)

Figure 6

Prognostic Classification for ALL

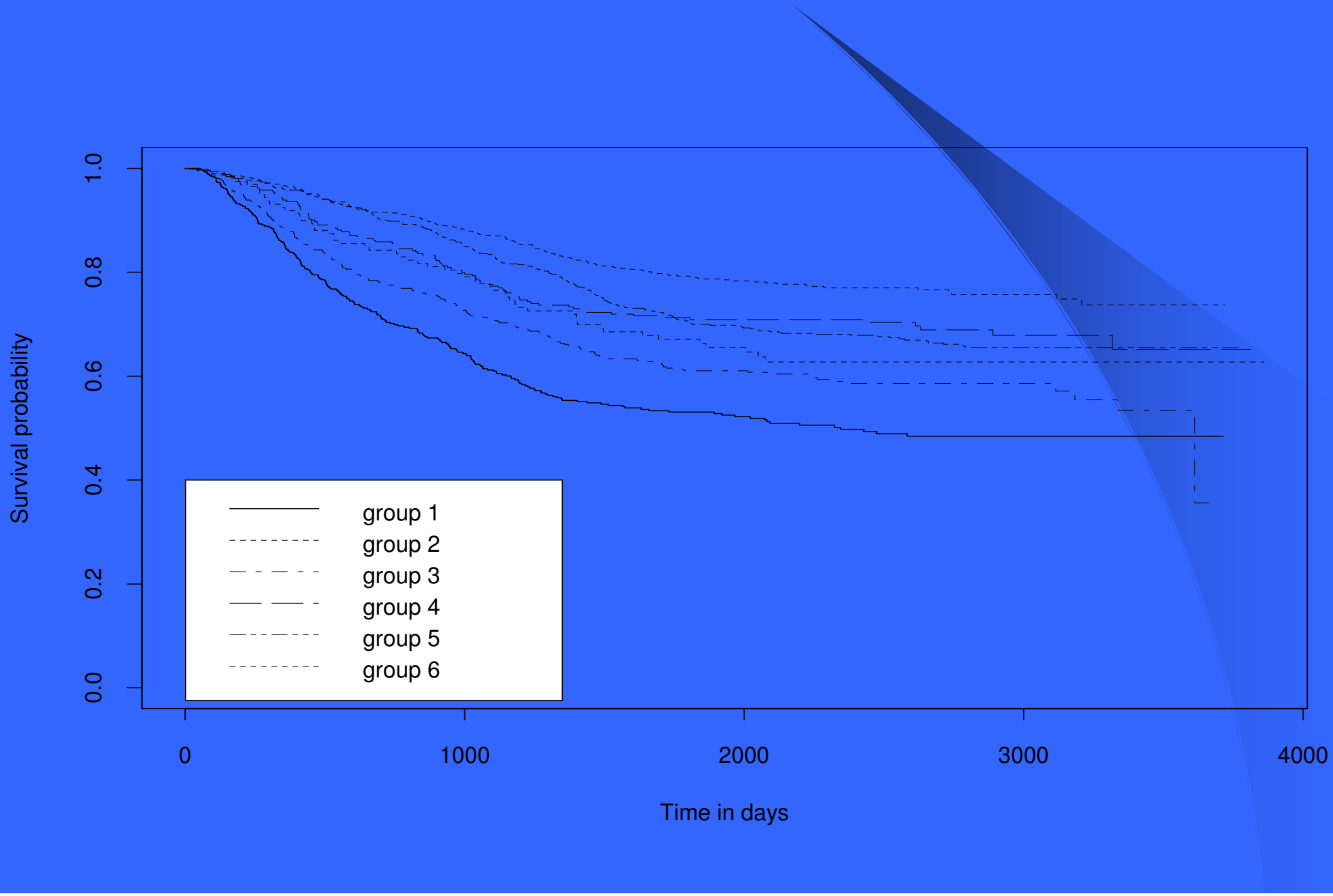
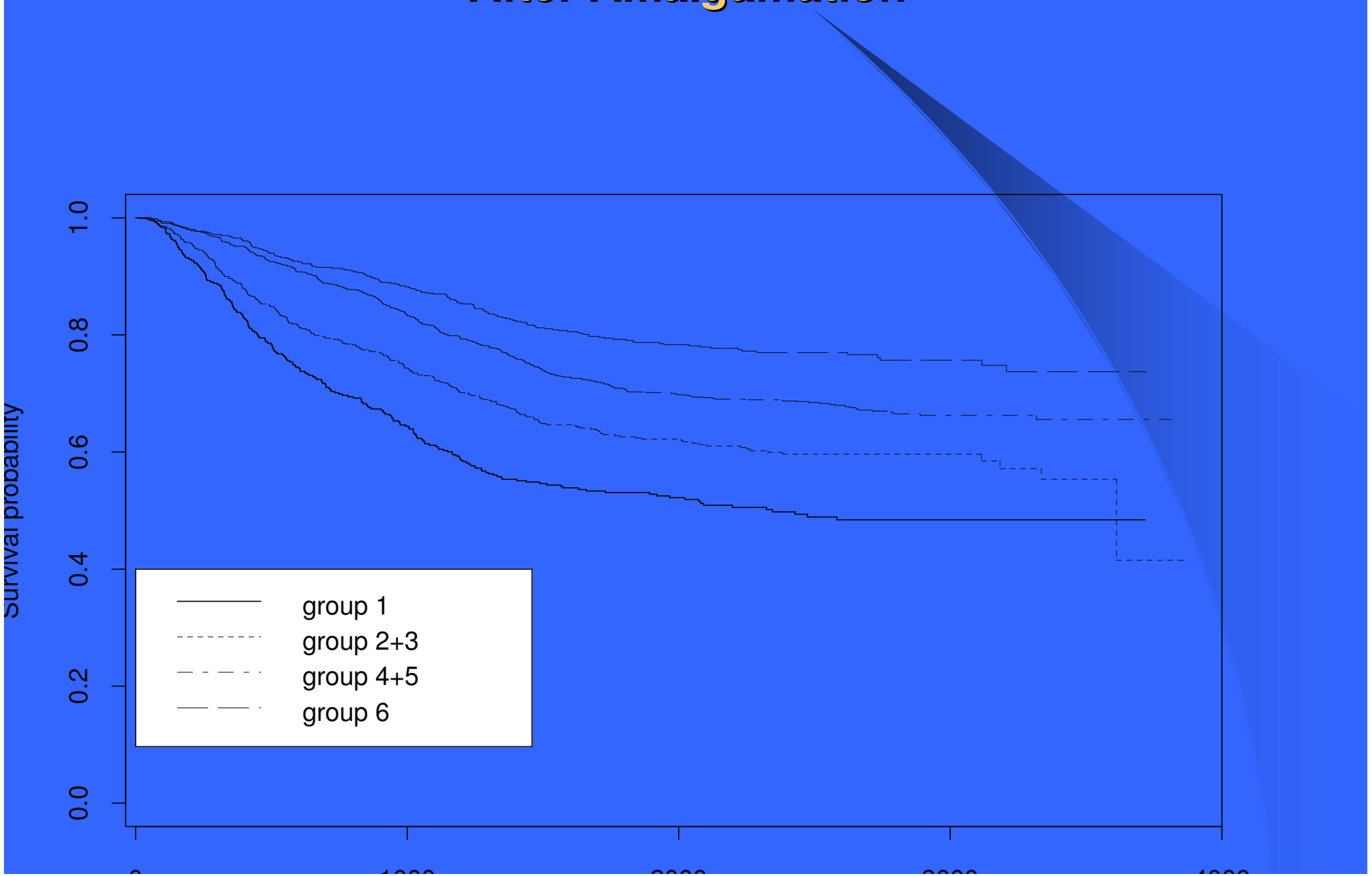


Figure 7
Prognostic Classification for ALL Data Set

After Amalgamation



Subgroup Analysis

Subgroup analysis refers to analysis that is aimed at uncovering possible variation in treatment effect in different patient subgroups such as male/female, young/old, with distinct molecular profile etc.

The question to be answered by this type of analysis is -- for whom does treatment work best?

It is reasonable to perform subgroup analysis in clinical trials *only* after the main comparison is shown to be significant (Bulpitt, 1988).

Possibility of clinically significant effects *within* subgroups leading to *overall null* effect (Gail and Simon, 1985)

- Such a scenario is very *unlikely* in clinical trials (Yusuf *et al.* 1991)

Exploratory analysis

Veteran Administration Lung Cancer Data Set

N=137

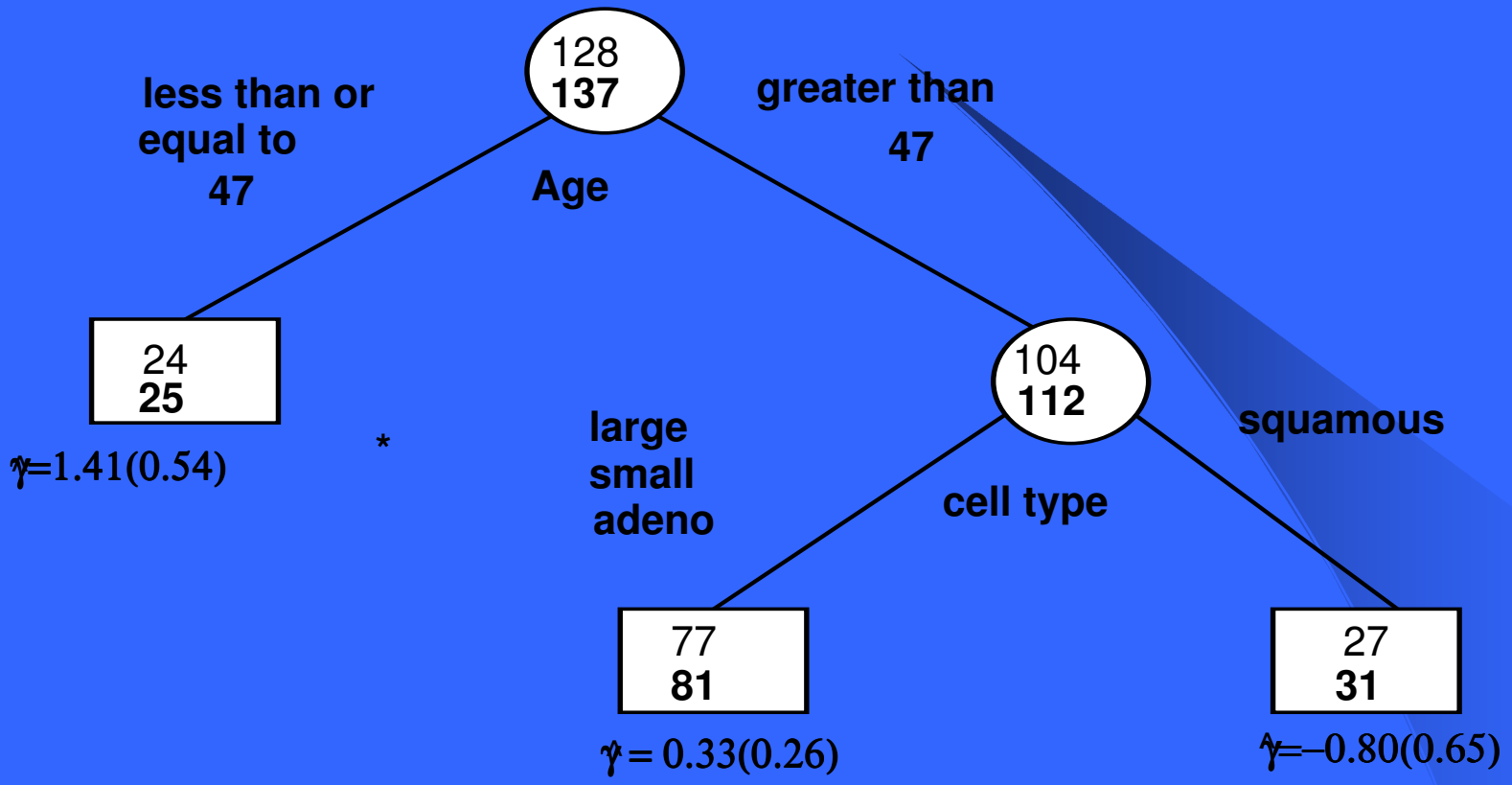
Six covariates: Performance status, disease duration, age, prior therapy, cell type, treatment (standard vs test)

Median follow-up time was 80 days

6.6% censoring

Employing Cox proportional hazards model, treatment didn't reach statistical significance after adjusting for the other covariates

Subgroup Analysis for Lung Cancer



* Cox regression coefficient for treatment effect(SE)
Deviance = 536.864

Treatment effect after adjustment for prognostic factors that did not appear in the subgroup tree-structure

Leaf	$\hat{\beta}\ (SE)$	HR	95% CI
1	0.97(0.62)	2.64	(0.79, 8.83)
2	0.41(0.26)	1.51	(0.91, 2.52)
3	-0.84(0.73)	0.43	(0.10, 1.79)