

Clustering patients into subgroups differing in optimal treatment alternative: QUINT

Elise Dusseldorp

Singapore, March 25, 2014



Universiteit
Leiden
The Netherlands

TNO

KATHOLIEKE UNIVERSITEIT
LEUVEN

Starting point: Two treatment alternatives



Treatment A



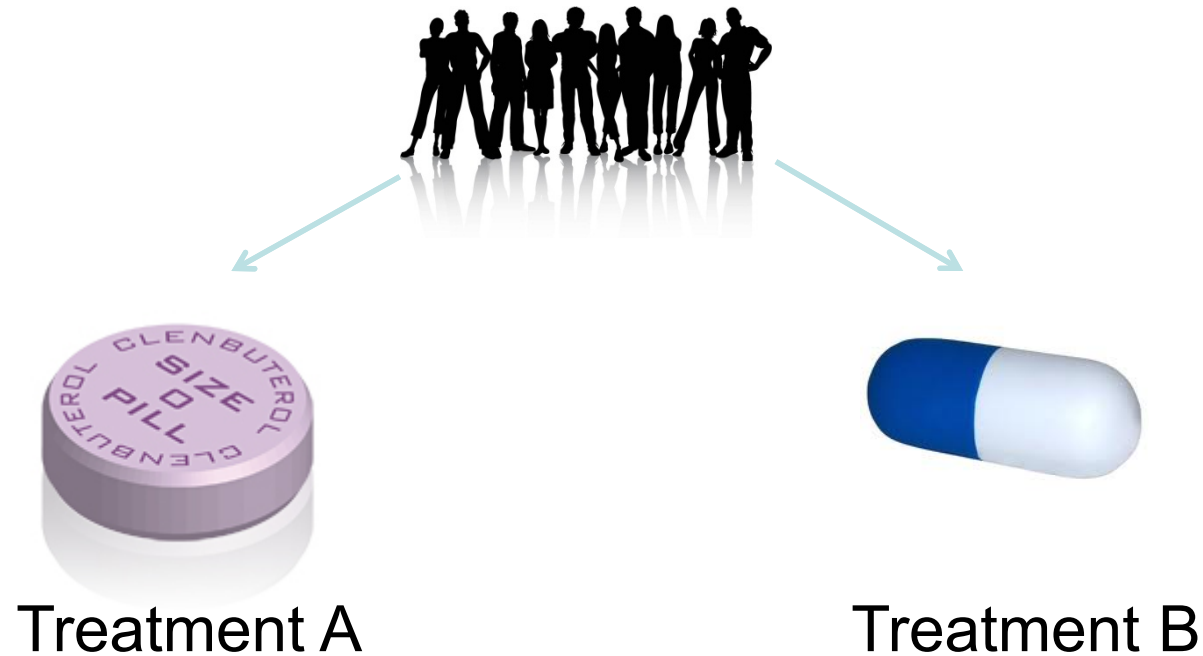
Treatment B

Professional: Which treatment for which patient?

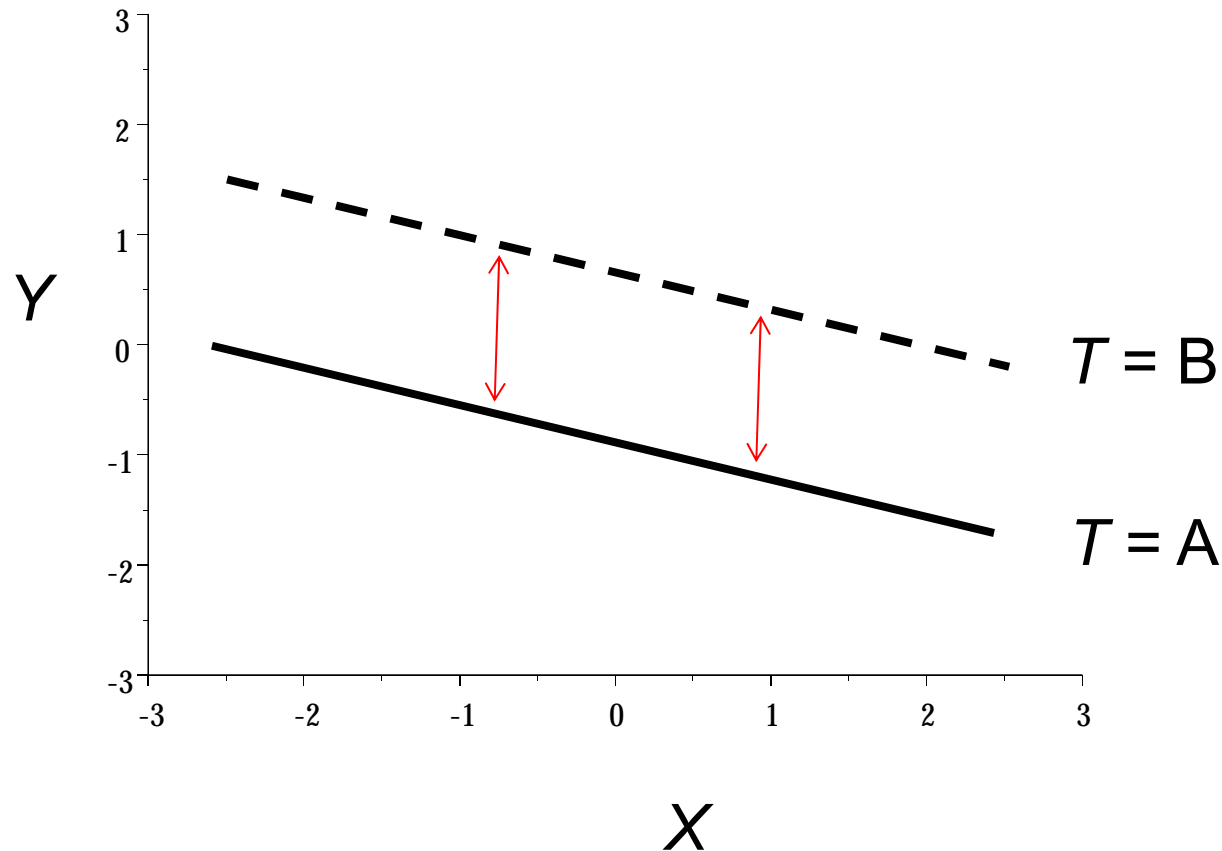
Patient: Which treatment is best for me?

Policy maker: Overall effectiveness \uparrow

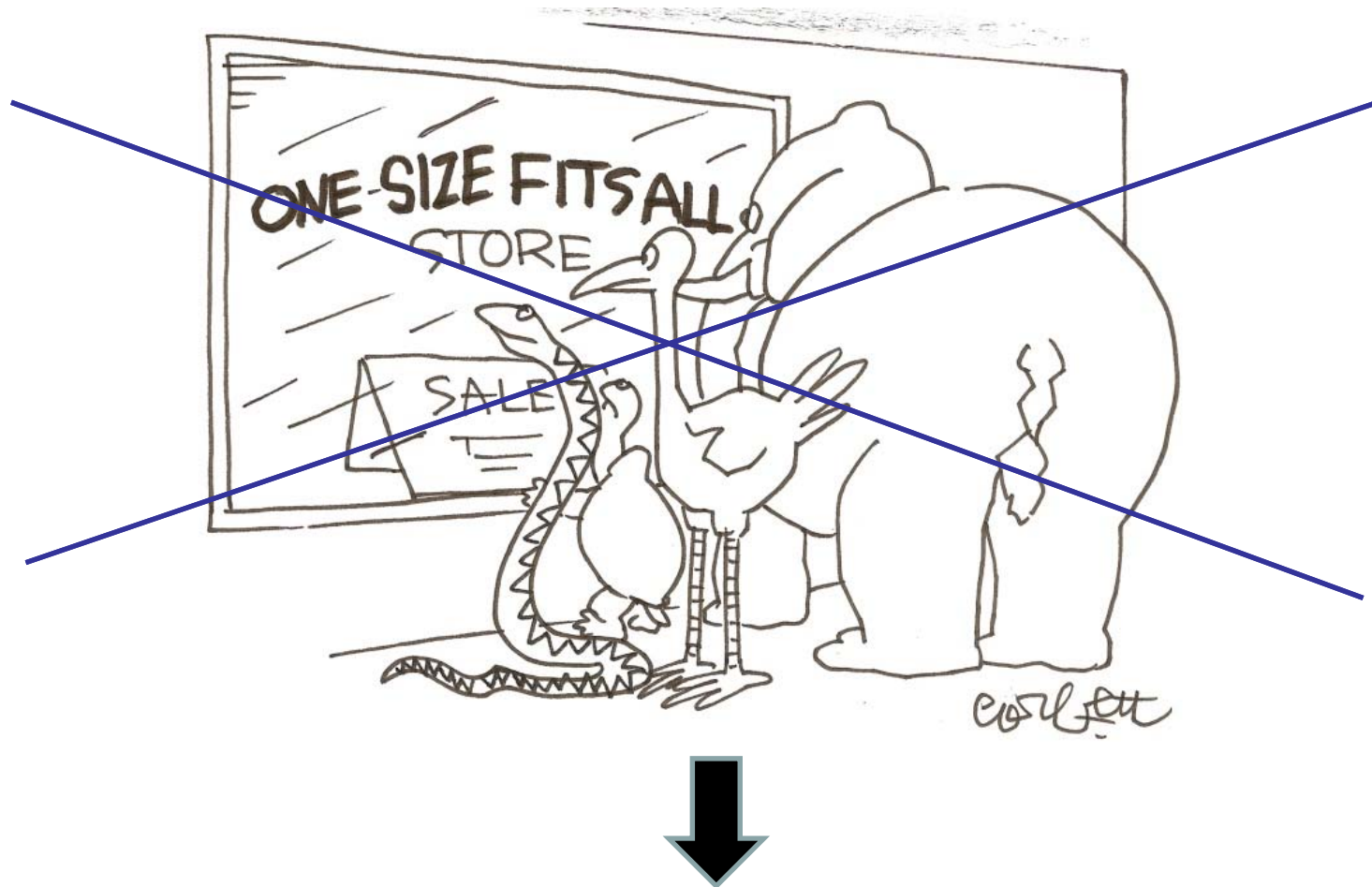
Randomized controlled trial



We measure the treatment outcome (Y) of the persons.



Implicit assumption: Mean difference between A and B is the same for all persons (i.e., only treatment main effect)



Personalized Medicine:

An individual will be given the treatment that is most effective given his or her characteristics.

New question: What works better for whom?

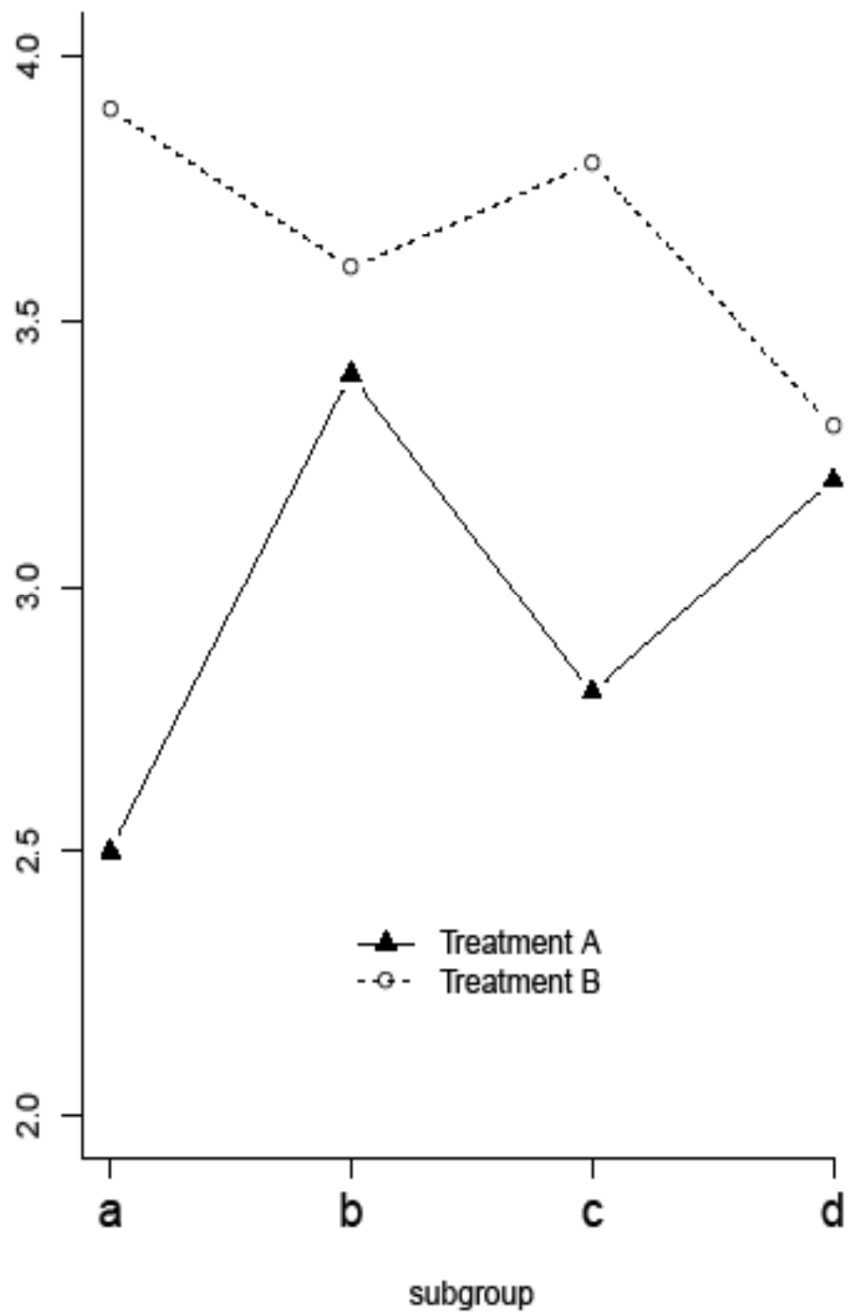
Treatment-subgroup interaction:

The effect of the treatment variable T on the outcome Y depends on the value of one or more **pretreatment** characteristics X_j ($j = 1, \dots, J$)

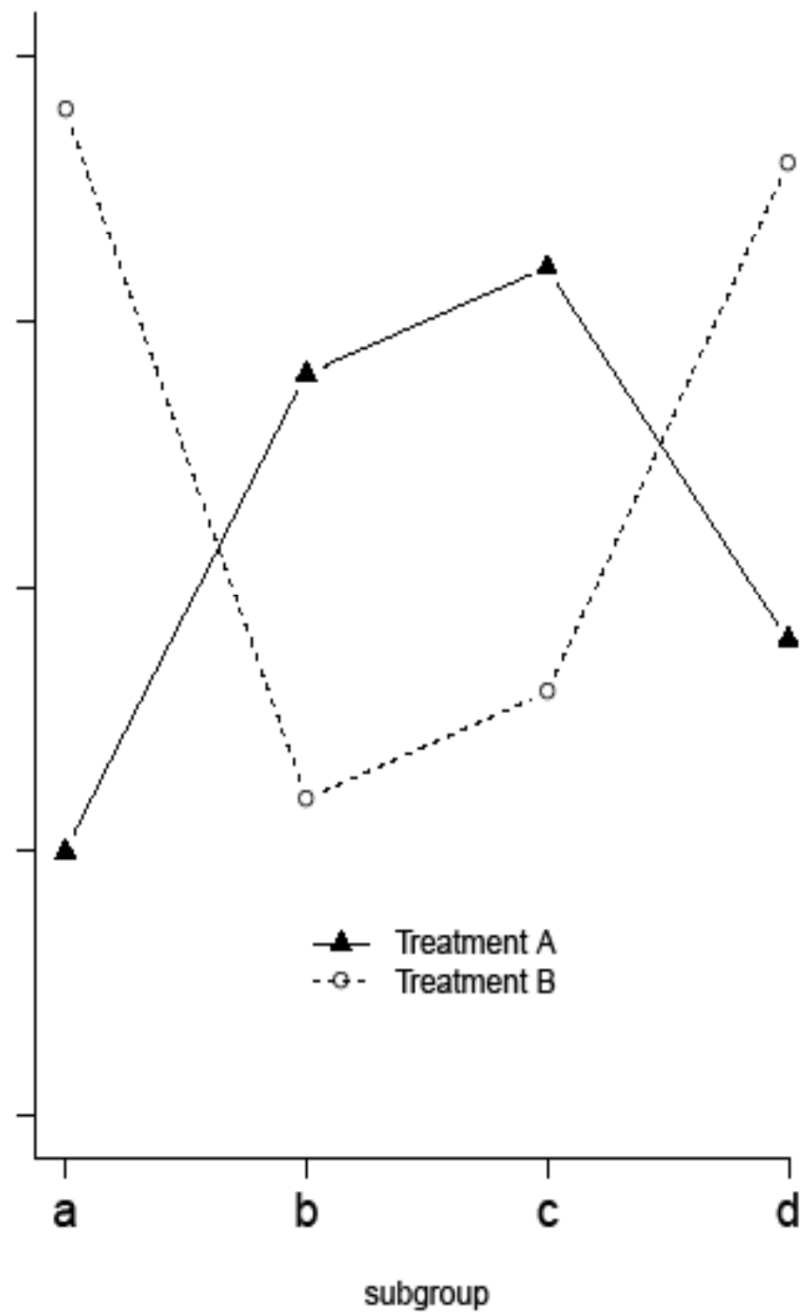
Pretreatment characteristics are also called:

effect modifiers or
moderators

Quantitative



Qualitative



Tree-based approaches to treatment-subgroup interactions

- 1) Regression trunk, STIMA (Dusseldorp et al. 2004, 2010)
- 2) Interaction Trees (Su et al. 2008, 2009)
- 3) Model-based recursive partitioning (Zeileis et al., 2008)
- 4) Virtual Twins (Foster et al., 2011)
- 5) SIDES (Lipkovich et al., 2011)
- 6)

A comparison of five recursive partitioning methods to find person subgroups involved in meaningful treatment–subgroup interactions

L. L. Doove · E. Dusseldorp ·
K. Van Deun · I. Van Mechelen

Received: 6 March 2013 / Revised: 5 December 2013 / Accepted: 9 December 2013
© Springer-Verlag Berlin Heidelberg 2013

Abstract In case multiple treatment alternatives are available for some medical problem, the detection of treatment–subgroup interactions (i.e., relative treatment effectiveness varying over subgroups of persons) is of key importance for personalized medicine and the development of optimal treatment assignment strategies. Randomized Clinical Trials (RCT) often go without clear a priori hypotheses on the subgroups involved in treatment–subgroup interactions, and with a large number of pre-treatment characteristics in the data. In such situations, relevant subgroups (defined in terms of pre-treatment characteristics) are to be induced during the actual data analysis. This

Advantages:

- Handle many moderators
- Model complex nonlinear relationships
- Simple interpretation and graphical representation
- Procedures to avoid spurious interaction effect

Disadvantage:

- No control over the type of interactions involved in the tree (quantitative or qualitative). *Qualitative* interactions have serious implications to treatment assignment.

Qualitative interaction trees: a tool to identify qualitative treatment–subgroup interactions

Elise Dusseldorp^{a,b,*†} and Iven Van Mechelen^b

When two alternative treatments (A and B) are available, some subgroup of patients may display a better outcome with treatment A than with B, whereas for another subgroup, the reverse may be true. If this is the case, a qualitative (i.e., disordinal) treatment–subgroup interaction is present. Such interactions imply that some subgroups of patients should be treated differently and are therefore most relevant for personalized medicine. In case of data from randomized clinical trials with many patient characteristics that could interact with treatment in a complex way, a suitable statistical approach to detect qualitative treatment–subgroup interactions is not yet available. As a way out, in the present paper, we propose a new method for this purpose, called QUalitative Interaction Trees (QUINT). QUINT results in a binary tree that subdivides the patients into terminal nodes on the basis of patient characteristics; these nodes are further assigned to one of three classes: a first for which A is better than B, a second for which B is better than A, and an optional third for which type of treatment makes no difference. Results of QUINT on simulated data showed satisfactory performance, with regard to optimization and recovery. Results of an application to real data suggested that, compared with other approaches, QUINT provided a more pronounced picture of the qualitative interactions that are present in the data. Copyright © 2013 John Wiley & Sons, Ltd.

Keywords: qualitative interaction; moderator; subgroup analysis; binary tree; partitioning; treatment efficacy

1. Introduction

1. Explanation of the method with examples
 - Partitioning criterion
 - Tree algorithm
2. Evaluation of performance
3. Demonstration of *R*-package quint

Breast Cancer Recovery Project

Scheier MF, Helgeson VS, et al. (JCO, 2007)

Patients:

Young women with early-stage breast cancer who had undergone lumpectomy combined with radiation and/or chemotherapy

Three-arm randomized controlled trial:

A) Nutrition information: how to adopt a low-fat diet ($n = 78$; $T = A$)

B) Education: provision of coping skills ($n = 70$; $T = B$)

[C) Control condition]

Outcome (Y):

Improvement in depression from pretest to posttest

Possible moderators (X_j):

Nationality, Marital status, Age, Weight-change, Treatment extensiveness, Comorbidity, Dispositional optimism, Unmitigated communion, Negative social environment

Result

?

VIP - project: Vitality in Practice

Coffeng, Hendriksen et al. (BMC Public Health, 2012)

Patients:

Employees of a Dutch financial service provide

Two-arm randomized controlled trial:

A) Group motivational interviewing (GMI): motivational sessions about how to increase your daily physical activity level and relaxation at work ($n = 149$)

B) Control: no motivational sessions ($n = 163$)

Outcome (Y):

Improvement in “need for recovery” from pretest to posttest

Possible moderators (X_j with $J = 25$):

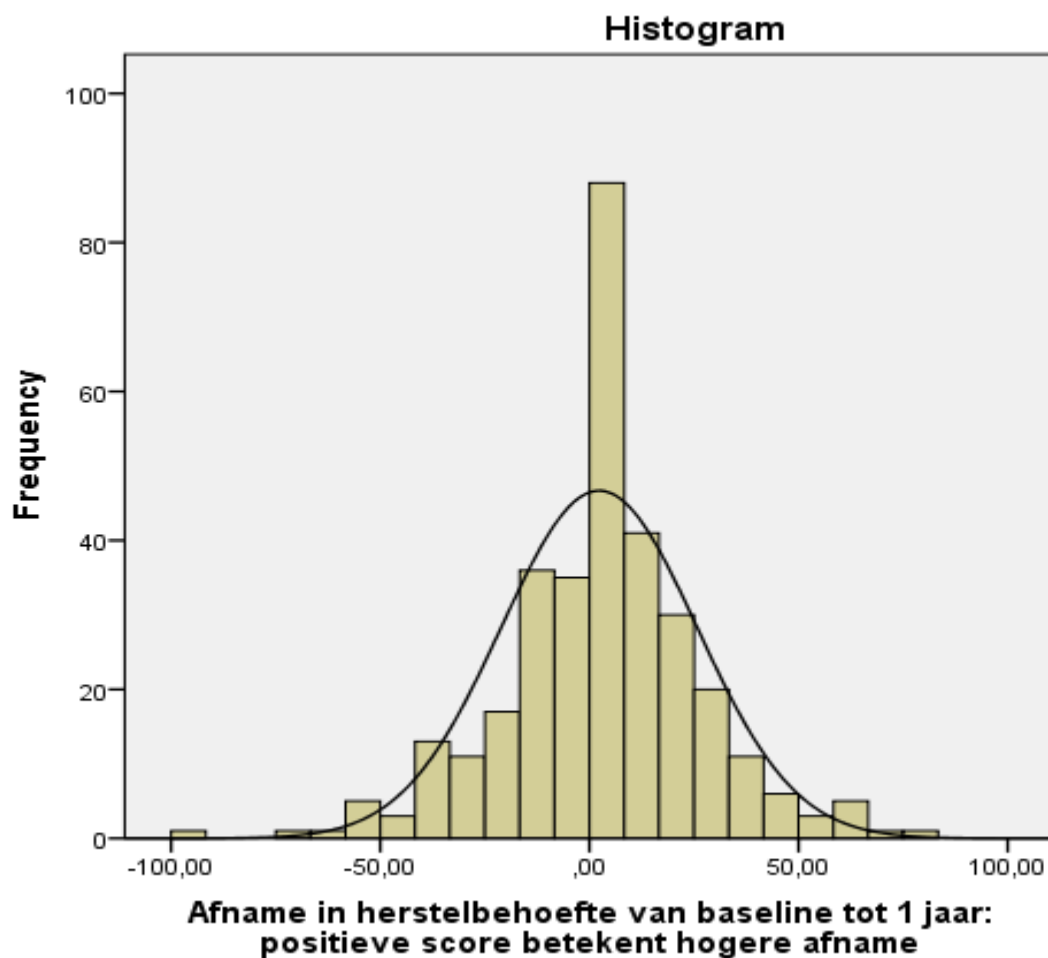
Background: Age, Sex, Level of education, Cohabiting, Mother country;

Health: BMI, General health, Mental health, (...);

Work-related: Team commitment, Organizational commitment, Supervisor support, Job demands, Working overtime, (...);

Primary Outcome variable Y

Improvement in need for recovery: Pretest - posttest



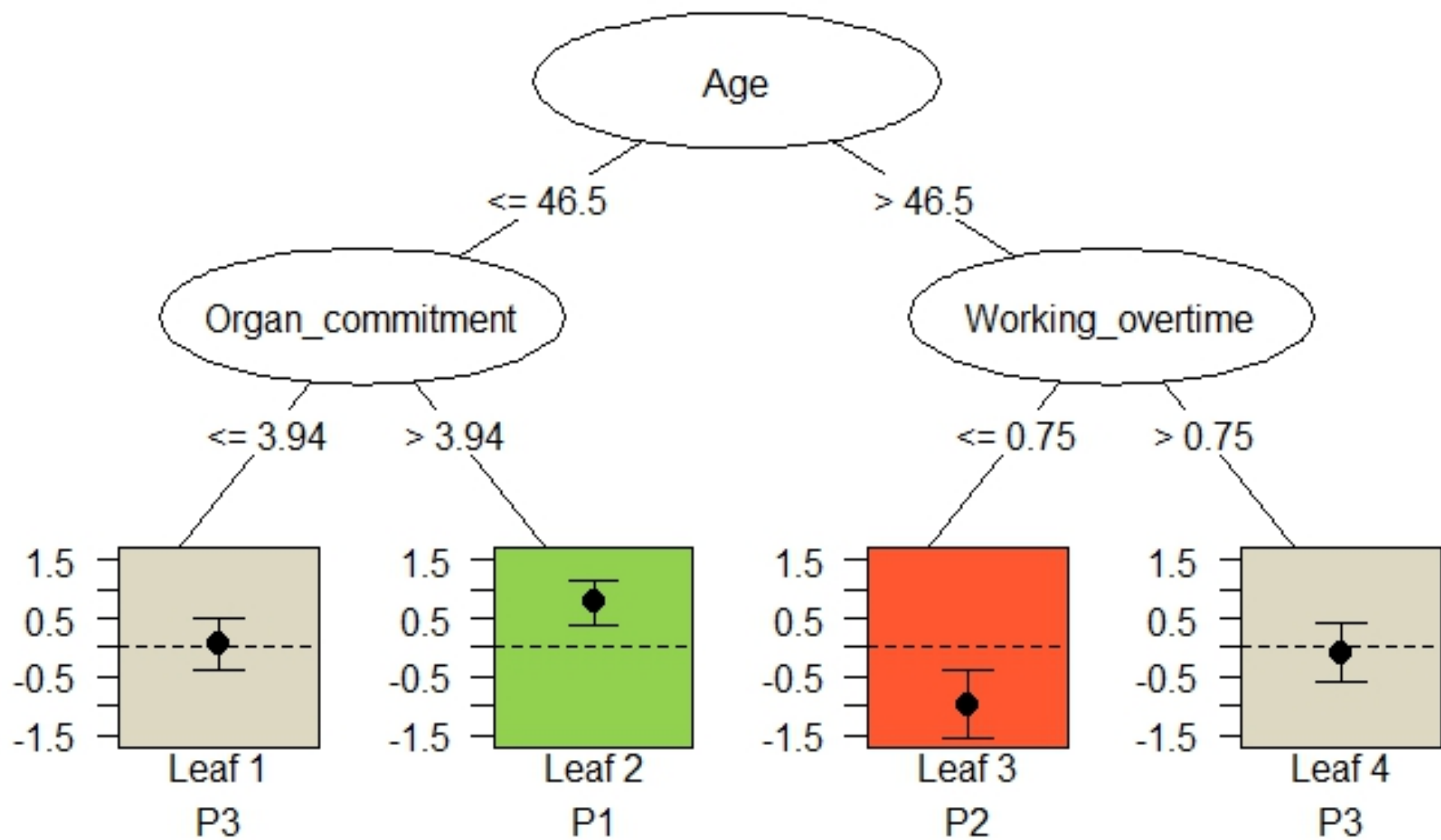
Range: -100 to 82
Mean (SD): 2.3 (23.4)
N = 329

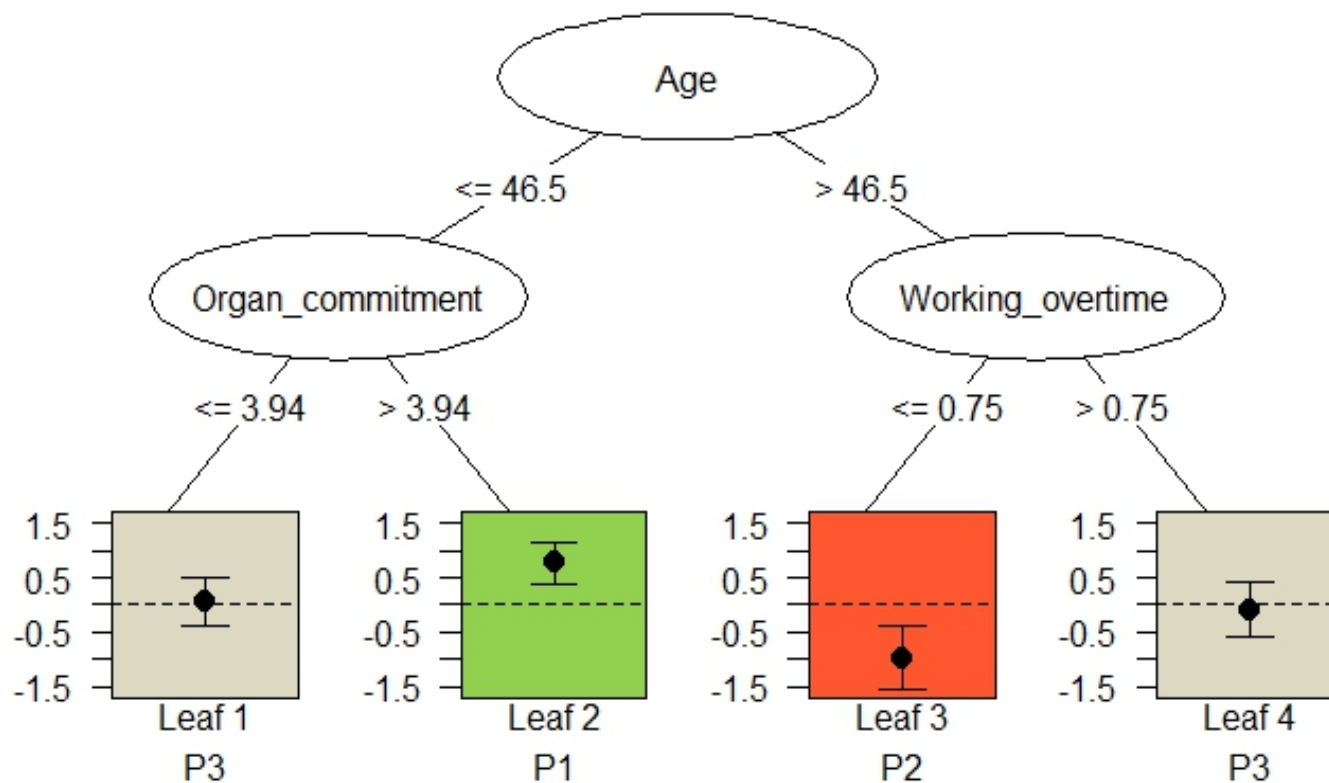
Descriptives of outcome Y per group

GMI		Control		Difference in means		Effect size
$\bar{Y}_{T=A}$ (s)	$\bar{Y}_{T=B}$ (s)	$\bar{Y}_{T=A} - \bar{Y}_{T=B}$ (s_{pooled})			d	
3.82 (25.32)	1.17 (21.98)	2.65 (23.63)			0.11	

Quint

- **Goal:** partitioning of patients on the basis of pretreatment characteristics into three subgroups that are involved in an optimal qualitative treatment-subgroup interaction
- **Three subgroups of patients:**
 - P_1 : for whom Treatment A is better than Treatment B
 - P_2 : for whom Treatment B is better than Treatment A
 - P_3 : for whom it does not make any difference





Leaf information:

	#(T=A)	meanY T=A	SD T=A	#(T=B)	meanY T=B	SD T=B	<i>d</i>	se
Leaf 1	43	2.11	22.98	38	1.08	18.54	0.05	0.23
Leaf 2	47	13.95	20.22	69	-4.06	25.35	0.77	0.20
Leaf 3	25	-11.71	26.00	27	9.83	17.95	-0.97	0.30
Leaf 4	34	3.41	28.43	29	5.63	17.90	-0.09	0.26

Quint

- Goal: partitioning of patients on the basis of pretreatment characteristics into three subgroups that are involved in an optimal qualitative treatment-subgroup interaction
- Three subgroups of patients:
 - P_1 : for whom Treatment A is better than B
 - P_2 : for whom Treatment B is better than A
 - P_3 : for whom it does not make any difference
- Partitioning criterion (C):
 - two components that are to be jointly maximized:
 - (1) Difference in treatment outcome component: in P_1 and P_2
 - (2) Cardinality component: Sample size in P_1 and P_2

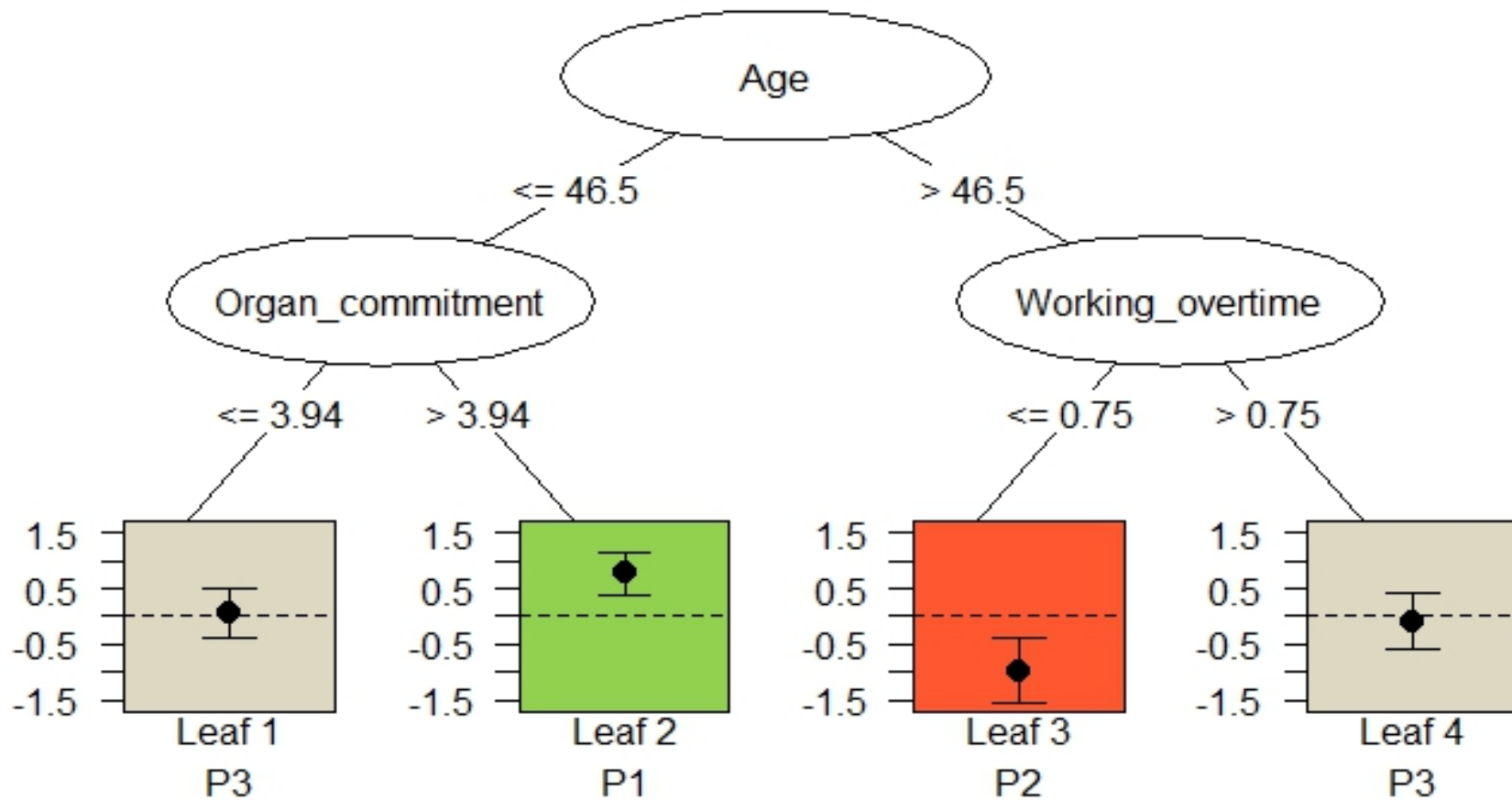
Difference in treatment outcome component

Define treatment difference in a leaf node R_ℓ ($\ell = 1, \dots, L$) as:

$$\alpha_\ell \left(\bar{Y}_{T=A,\ell} - \bar{Y}_{T=B,\ell} \right), \text{ with}$$

$$\alpha_\ell = 1 \text{ or } \alpha_\ell = 1 / s_\ell$$

Difference in means or Effect size



Let $f(\ell)$ assign each node R_ℓ to the partition classes;
 in this example: $f(1) = 3$; $f(2) = 1$; $f(3) = 2$; $f(4) = 3$;

Difference in treatment outcome in P_1 = mean difference in treatment outcome of all nodes R_ℓ assigned to P_1 :

$$D_1 = \frac{\sum_{\ell=1}^L I(f(\ell) = 1) \#R_\ell \alpha_\ell \left(\bar{Y}_{T=A,\ell} - \bar{Y}_{T=B,\ell} \right)}{\sum_{\ell=1}^L I(f(\ell) = 1) \#R_\ell}$$

→ similarly for difference in treatment outcome in P_2 :

$$D_2 = \frac{\sum_{\ell=1}^L I(f(\ell) = 2) \#R_\ell \alpha_\ell \left(\bar{Y}_{T=B,\ell} - \bar{Y}_{T=A,\ell} \right)}{\sum_{\ell=1}^L I(f(\ell) = 2) \#R_\ell}$$

Cardinality component

Cardinality of P_1 : $\sum_{\ell=1}^L I(f(\ell) = 1) \# R_\ell$

Cardinality of P_2 : $\sum_{\ell=1}^L I(f(\ell) = 2) \# R_\ell$

$$\left(\sum_{\ell=1}^L I(f(\ell) = 1) \# R_\ell \right) \times \left(\sum_{\ell=1}^L I(f(\ell) = 2) \# R_\ell \right)$$

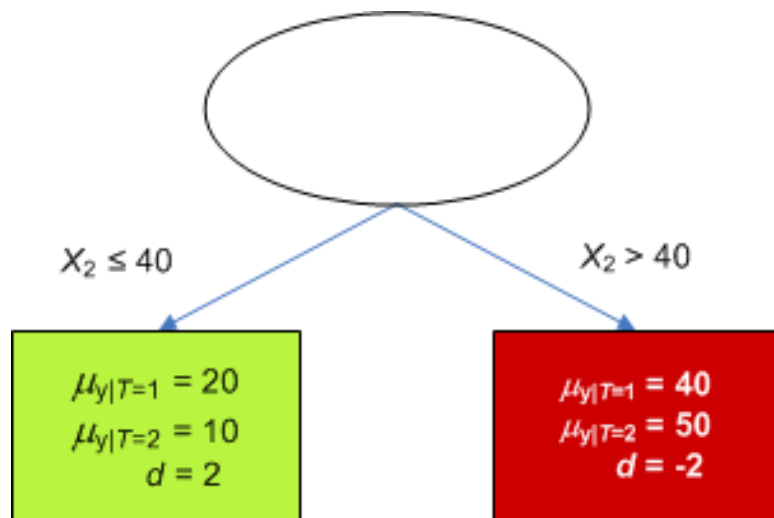
Partitioning criterion of QUINT:

Put the components on comparable measurement scales, by giving them suitable weights:

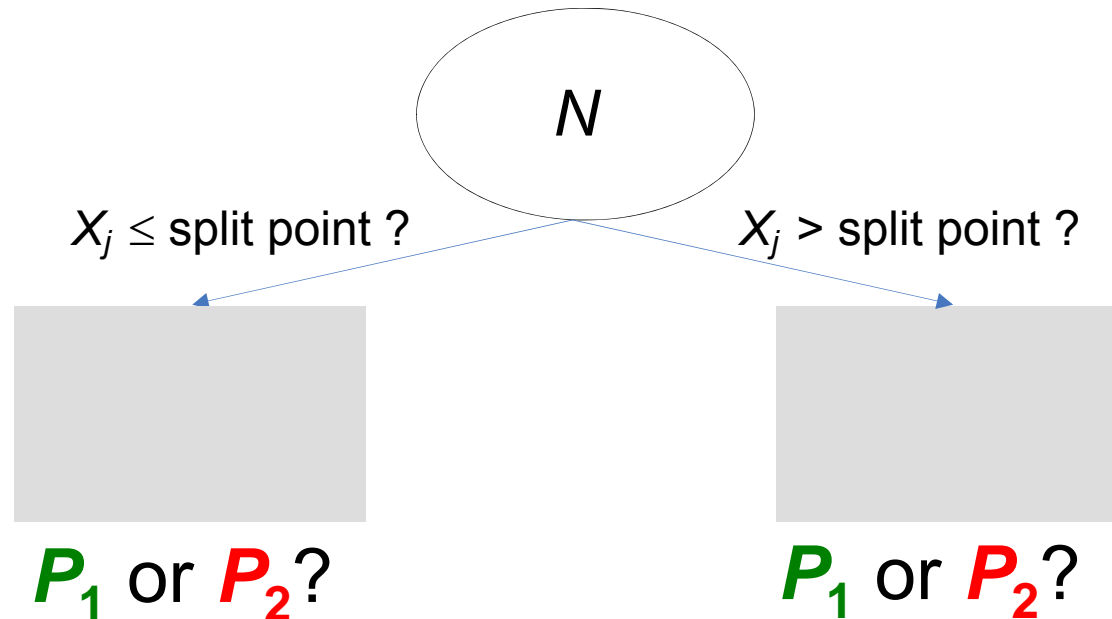
$$C = w_1 \left[\log(1 + D_1) + \log(1 + D_2) \right] + w_2 \left[\log \left(\sum_{\ell=1}^L I(f(\ell) = 1) \# R_\ell \right) + \log \left(\sum_{\ell=1}^L I(f(\ell) = 2) \# R_\ell \right) \right]$$

Boundary conditions of QUINT

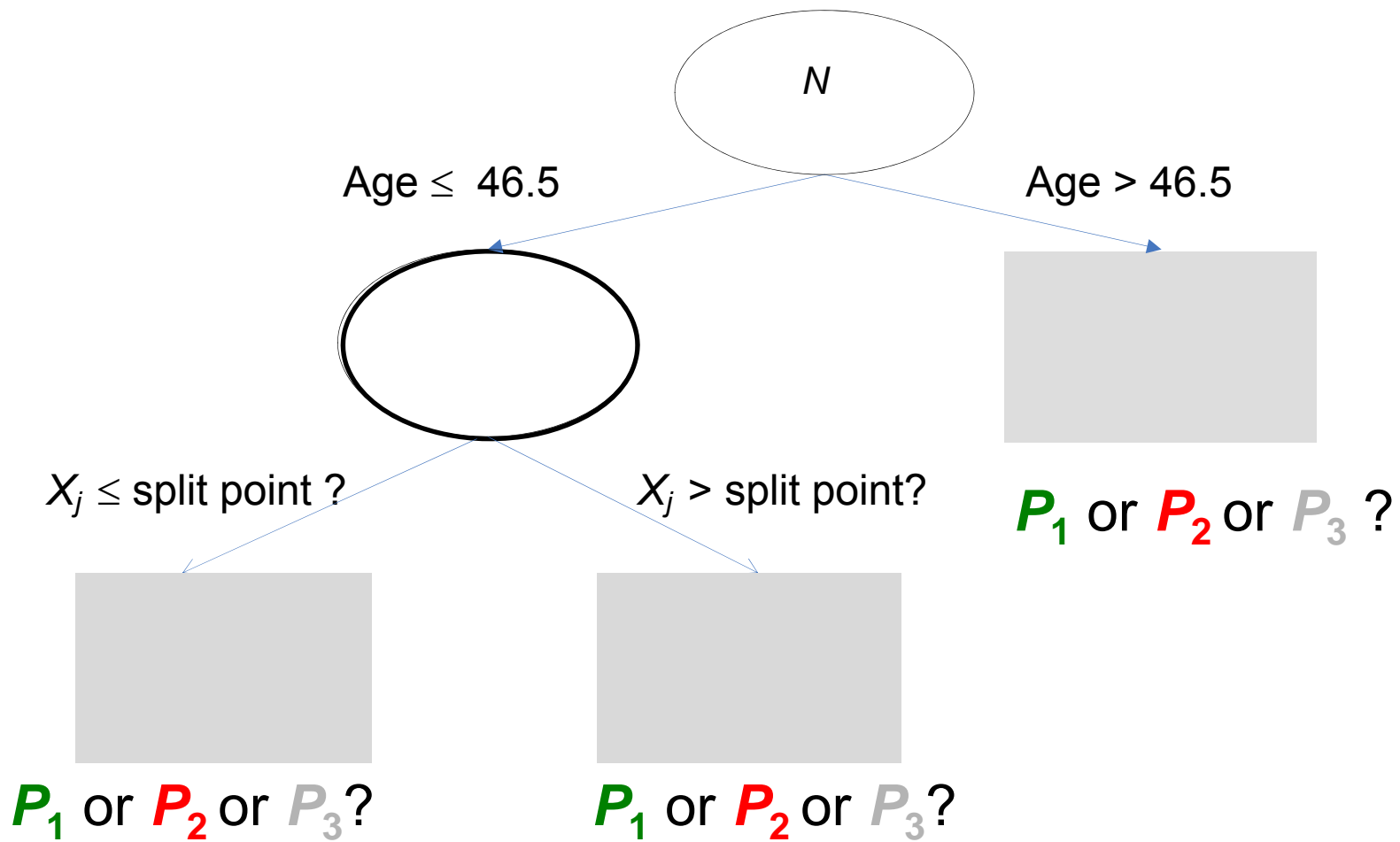
- 1) Nonempty partition class condition: P_1 and P_2 may not be empty
- 2) After first split, the qualitative interaction condition:
in each of the two leaves $|d| \geq d_{\min}$

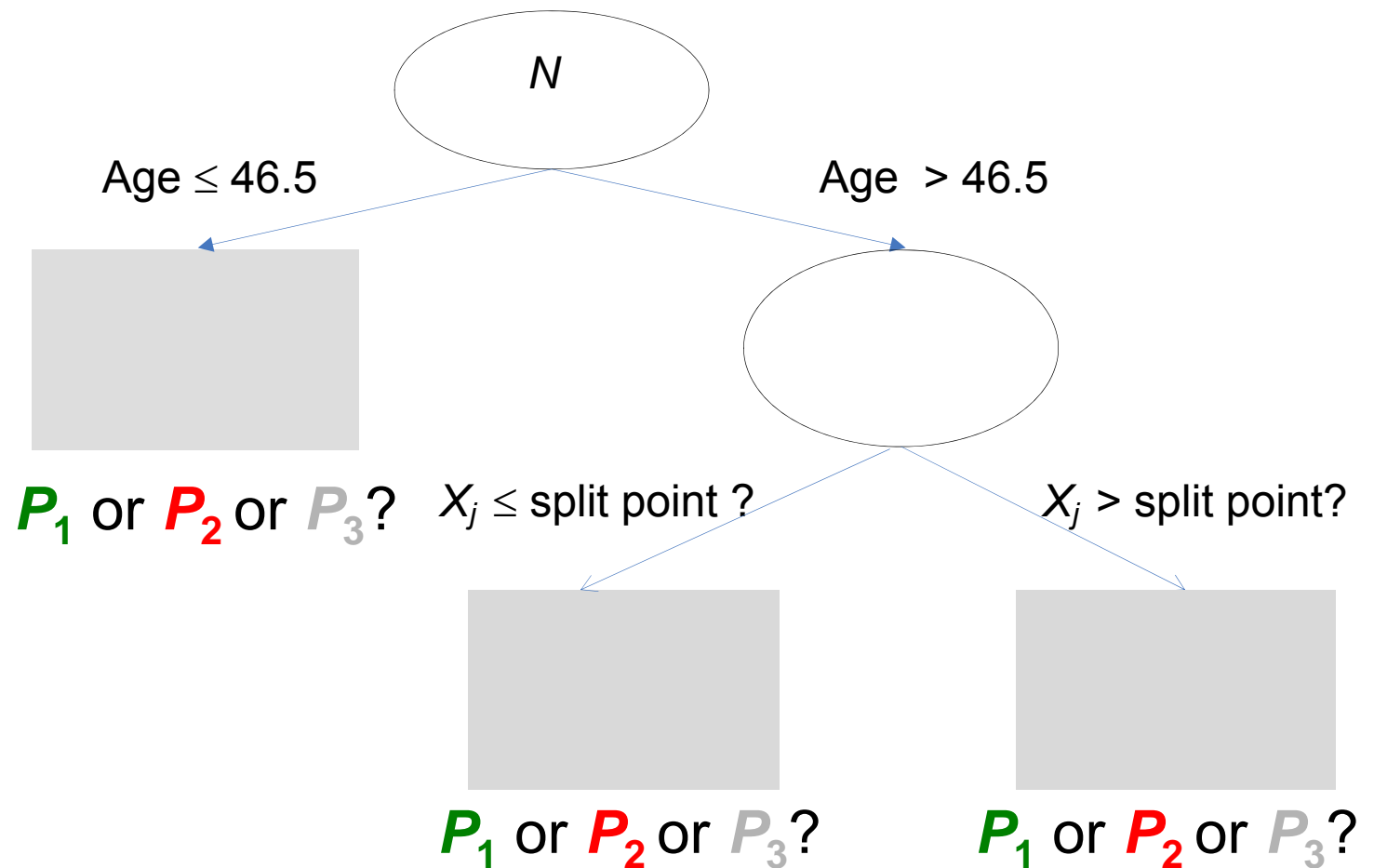


Tree growing algorithm of QUINT



Step 1: Determine the optimal combination (X_j , split point, assignment):
⇒ Select the one that induces the highest value of C





Step 2: Across all parent nodes: Select the one with the optimal combination (X_j , split point, assignment) that maximizes C

When to stop growing?

Tree growing process stops if further splitting with increase in C is impossible

Additional stopping criterion

Minimal size per treatment condition in a leaf :

a_1 = minimum sample size in $T = A$, and

a_2 = minimum sample size in $T = B$

Final step: Pruning

Bias-corrected bootstrap (Efron, JASA, 1983; LeBlanc & Crowley, JASA, 1993)

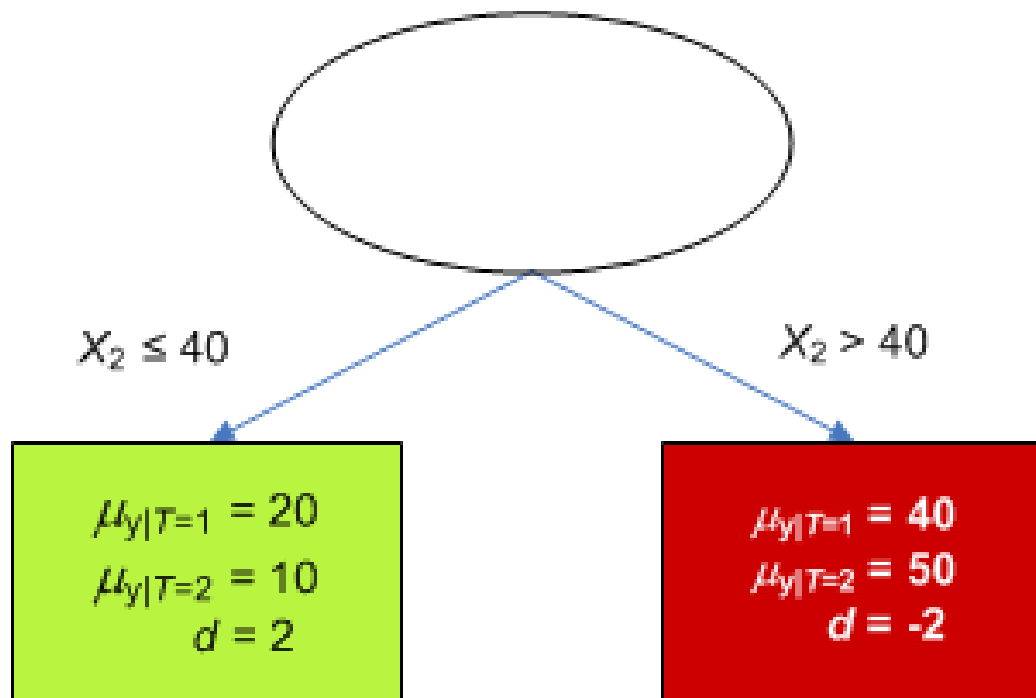
- Grow a tree of size L on the original sample and compute the partitioning criterion: C_L^{app}
- Generate B bootstrap samples
- For each bootstrap sample b ($b= 1, \dots, B$), grow a tree of size L and compute the partitioning criterion: $C_{b,L}^{boot}$
Freeze the tree and calculate C for original sample: $C_{b,L}^{orig}$
Calculate the overoptimism: $O_{b,L} = C_{b,L}^{boot} - C_{b,L}^{orig}$
- Across all b 's, compute overall optimism: $\bar{O}_L = 1/B \sum_{b=1}^B O_{b,L}$
- Calculate the bias-corrected value of C as:

$$C_L^{bc} = C_L^{app} - \bar{O}_L$$

1. Explanation of the method with an example
 - Partitioning criterion
 - Tree algorithm
2. Evaluation of performance
3. Demonstration of *R*-package quint

Simulation study

True model:



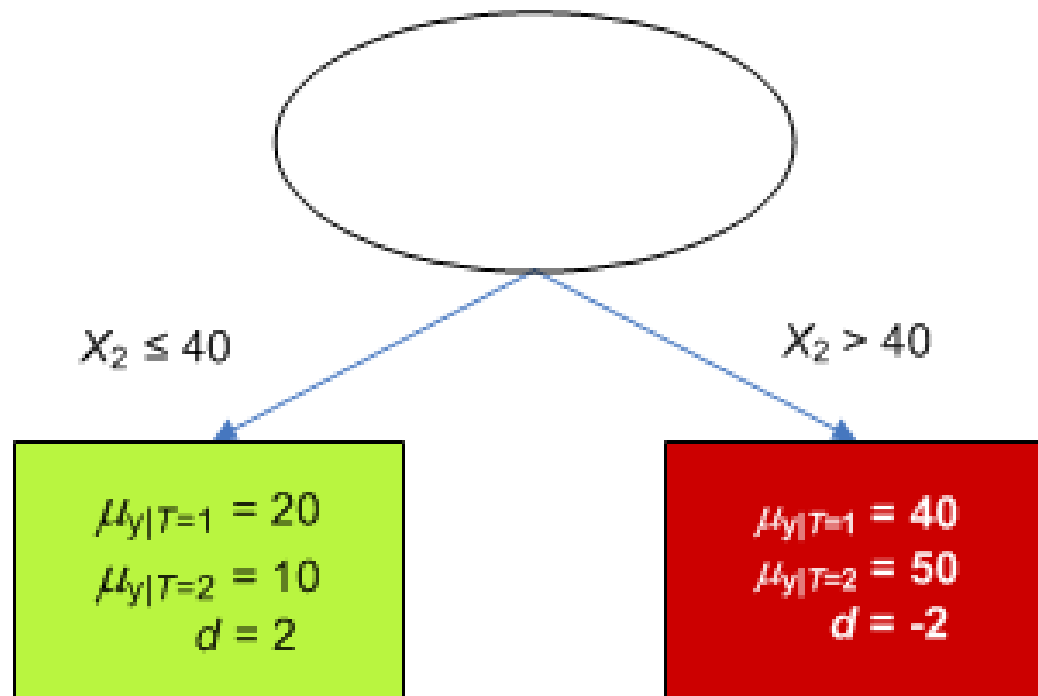
Simulation study: design factors

- Number of moderators (J): 5, 10, 20
- Intercorrelation of moderators (ρ): 0, 0.20
- Sample size (N): 200, 300, 400, 500, 1000
- True effect size $|d|$: 0.50, 1.00, 2.00

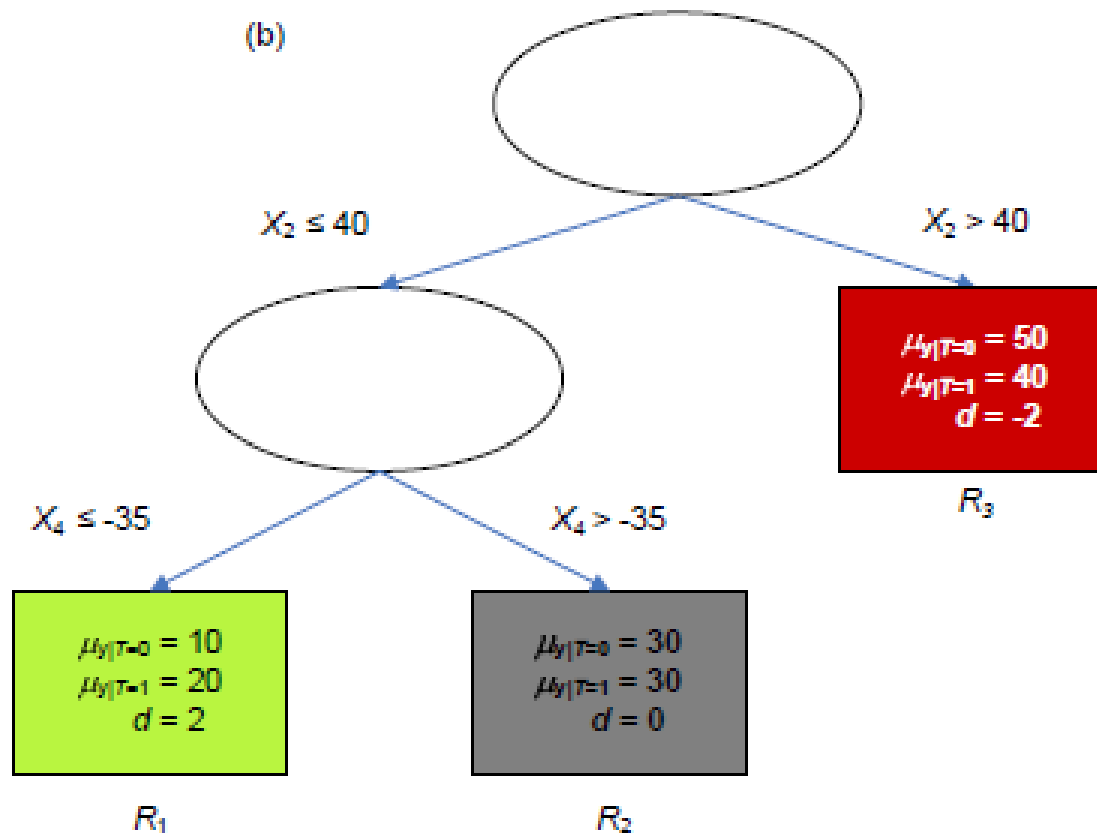
100 data sets per cell: $3*2*5*3 \rightarrow 9,000$ analyses per model

True models were varied in complexity (A through E)

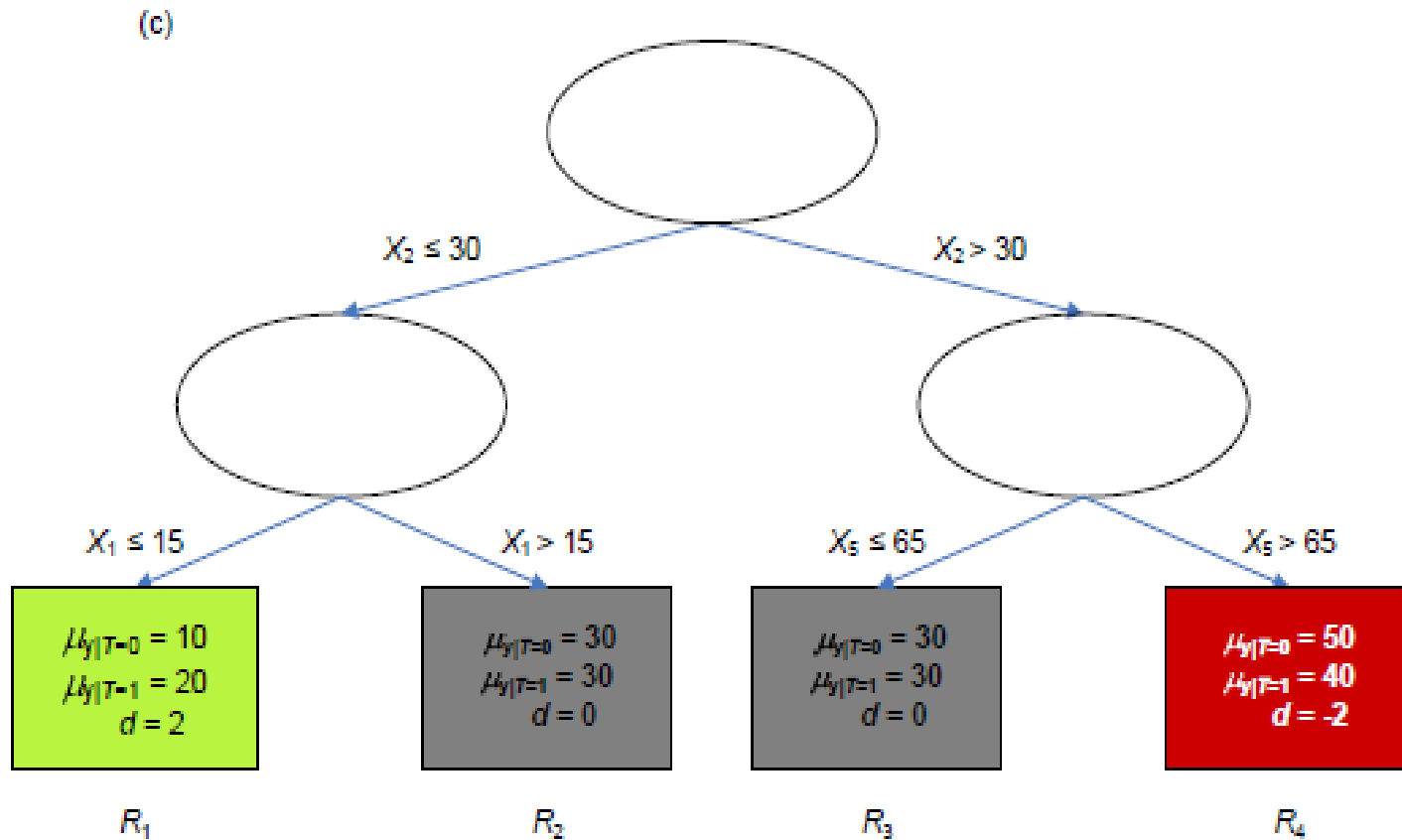
True model A:



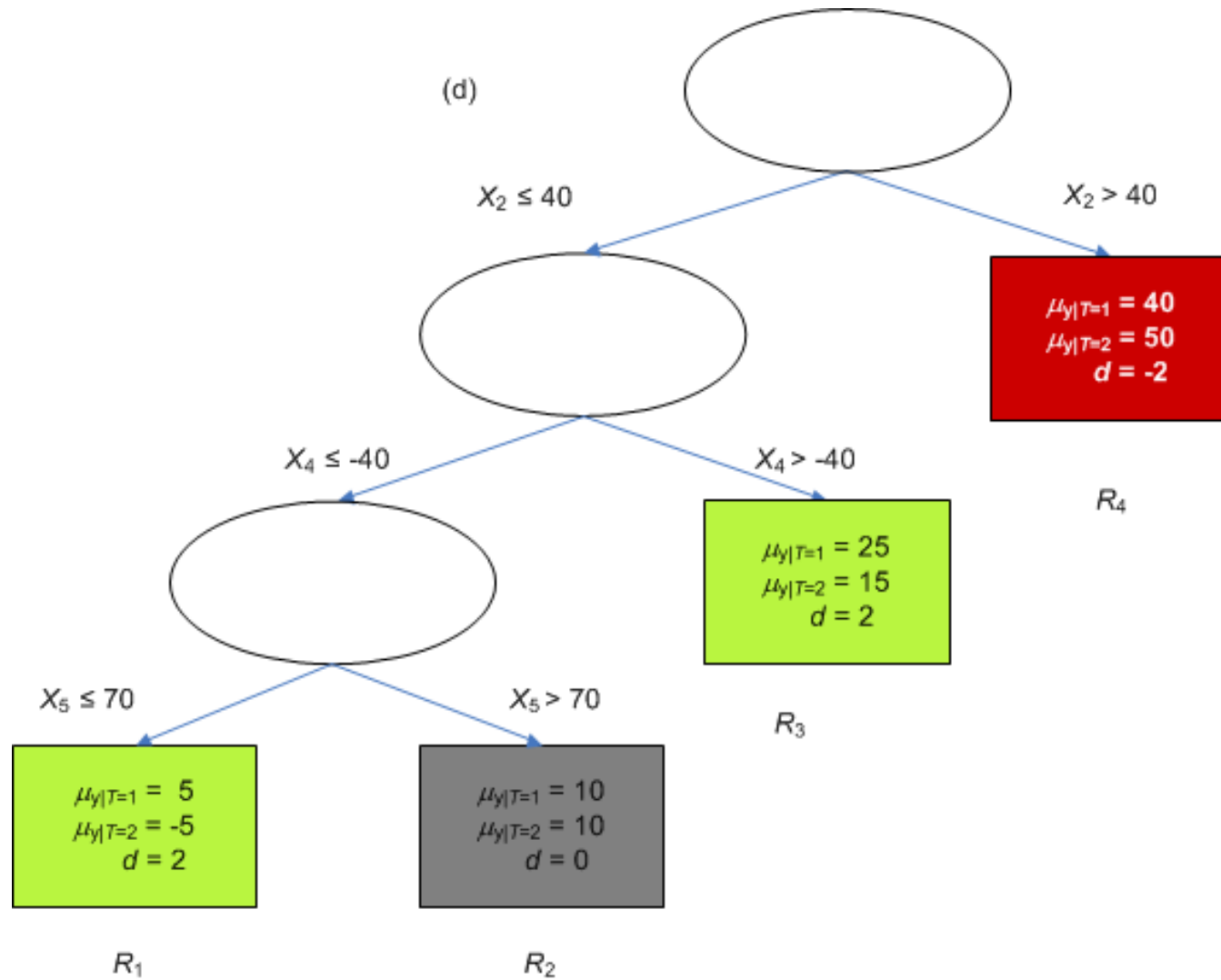
True model B:



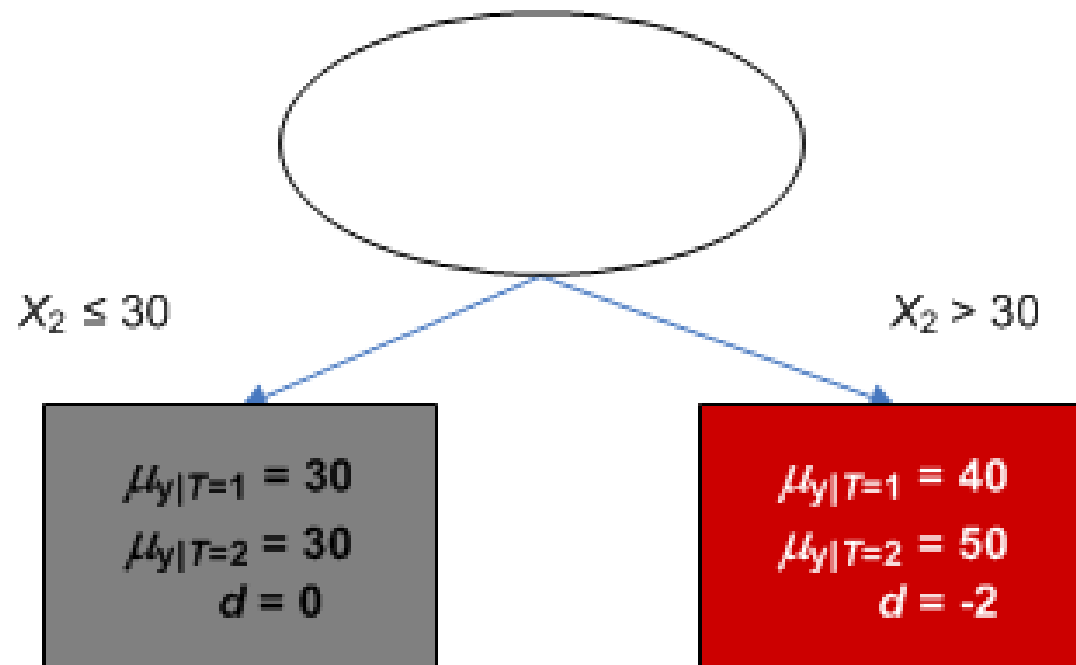
True model C:



True model D:



True model E:



Results simulation study

- **Optimization performance:** $\hat{C} \geq C_{true}$
97% non-suspect solutions
- **Recovery performance:**
 - (a) Type I and Type II errors
Evaluated for different values of d_{min} : 0.20 to 0.40
Good balance if $d_{min} = 0.30$
Caution when $N \leq 300$
 - (b) Recovery of tree complexity
Good for all models if true value of d was large (2.00)
If $d < 2.00$ and $N \leq 300$: recovery was unsatisfactory for more complex models (C and D)

Results (continued)

(c) Recovery of splitting variables and split points

- depends on sample size, true effect size, and on number of moderators
- satisfactory for $N \geq 400$ and $|d| \geq 1$

(d) Recovery of assignments to partition classes

- Cohen's K varied between .93 to .56 for models A to D

Conclusions:

- Largely acceptable for all models when $N \geq 400$ and true effect sizes ($|d|$) ≥ 1 .
- Influence of number of moderators and intercorrelation is very small

Demonstration of *R*-package quint

```
> library(quint)
```

```
Loading required package: partykit
```

```
Loading required package: grid
```

```
Loading required package: Formula
```

```
Loading required package: rpart
```

```
> vipdats[1:3,] [these data are not included in the package]
```

	ID	Gender	Country	Working_overtime	BMI	Age	GMI	...
1	100	1	1	0	27.131	43	1	
2	108	0	1	0	19.840	48	2	
3	110	0	1	0	23.030	46	1	

```
> form1<- NFRch ~ GMI | NFR_T0 + Gender + Country+  
Working_overtime + BMI + Age + Organ_commitment + ....
```

Start quint analysis: Grow a large tree

```
> set.seed(1)
> quint1vip <- quint(form1, data = vipdats)
```

The sample size in the analysis is 312

split 1

#leaves is 2

Bootstrap sample 1

Bootstrap sample 2

Bootstrap sample 3

...

Bootstrap sample 25

split 2

#leaves is 3

Bootstrap sample 1

...

splitting process stopped after number of leaves equals 4 because new value of C was not higher than current value of C .

Inspect the results

```
> summary(quint1vip)
```

```
Fit information:
```

```
              Criterion
-----
split #leaves  apparent  biascorrected  se
  1      2      3.49      3.29      0.03
  2      3      3.54      3.30      0.02
  3      4      3.59      3.33      0.02
```

```
Split information:...
```

```
Leaf information: ...
```

Example of other results (BCRP-example of quint paper)

```
> summary(quint1bcrp)
```

```
Fit information:
```

		Criterion		

split	#leaves	apparent	biascorrected	se
1	2	2.51	2.25	0.03
2	3	2.66	2.23	0.05
3	4	2.78	2.27	0.05
4	5	2.86	2.32	0.05
5	6	2.89	2.30	0.05
6	7	2.90	2.24	0.10

```
Split information:...
```

```
Leaf information: ...
```

Prune the tree and plot

```
> quint1vippr <-prune(quint1vip)
```

```
The sample size in the analysis is 312
```

```
split 1
```

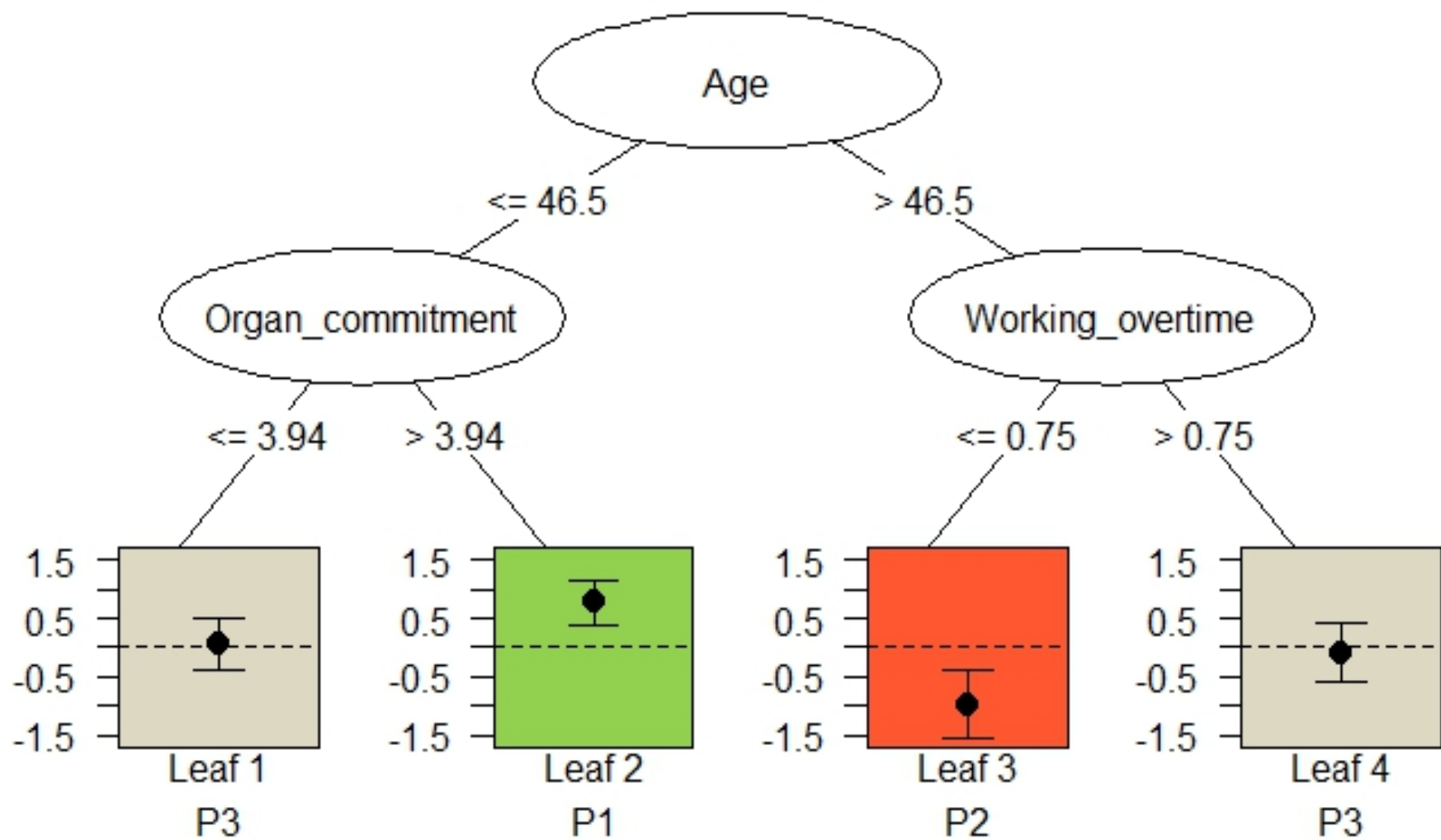
```
#leaves is 2
```

```
...
```

```
split 3
```

```
#leaves is 4
```

```
> plot(quint1vippr)
```



Inspect the leaf information of the pruned tree:

```
> summary (quint1vippr)
```

```
Fit information:...
```

```
Split information: ...
```

```
Leaf information:
```

	#(T=1)	meanY T=1	SD T=1	#(T=2)	meanY T=2	SD T=2	d	se
Leaf 1	43	2.11	22.98	38	1.08	18.54	0.05	0.23
Leaf 2	47	13.95	20.22	69	-4.06	25.35	0.77	0.20
Leaf 3	25	-11.71	26.00	27	9.83	17.95	-0.97	0.30
Leaf 4	34	3.41	28.43	29	5.63	17.90	-0.09	0.26

Change default options:

```
> cnew <- quint.control (B = 200, crit = "dm", a1 = 25, a2 = 25)
```

```
> set.seed(4)
```

```
> quint2vip<-quint(form1, data= vipdats, control=cnew)
```

Conclusions

- Quint partitions the total group of patients into subgroups that differ in treatment efficacy
- The criterion optimizes qualitative treatment subgroup interactions
- The splits leading to the leaves of the tree can be used to set up optimal treatment regimes

Future

- Multiple treatment groups (> 2)
- Splits on categorical variables
- Validation procedure of estimated effect sizes
- Globally optimized qualitative interaction trees

Acknowledgement

- Iven van Mechelen (KU Leuven)
- Lisa Doove (KU Leuven)
- Margriet Formanoy (TNO)
- Erwin Tak (TNO)
- Ingrid Hendriksen (TNO)
- Jacqueline Meulman (Leiden University)

More info: elise.dusseldorp@math.leidenuniv.nl