

metacart 3.0

Classification and regression trees for Exploratory Moderator Analysis in Meta-Analysis

Juan Claramunt Gonzalez, Xinru Li, Elise Dusseldorp, Xiaogang Su,
and Jacqueline Meulman



Universiteit
Leiden
The Netherlands



July 2026

- Meta-analysis: summarize results reported in multiple studies on the same topic and analyze which predictors have an effect.
- Common approach: Meta-regression.
- Statistical problem: Linear model is sometimes not suitable or interaction effects are ignored.
- Solution: Use a tree algorithm! MetaCART.

Problems:

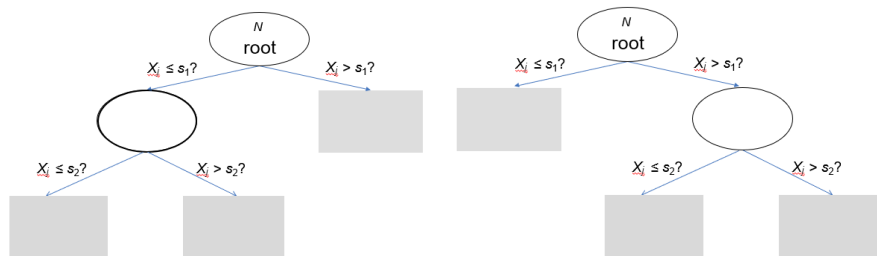
- 1 The split may be based on a local optimum

Solutions:

- 1 Splitting using look ahead

Tree construction, convergence and stability

```
REtree_lookahead <- REMrt(form_Levine, vi = vi, c.pruning = 0,  
  data = SpermCon, lookahead = TRUE)
```



- Find the best scenario for each split by considering multiple splits across different layers.
- Maximize Q_B^* , the between-subgroups heterogeneity.

Problems:

- 1 The split may be based on a local optimum
- 2 The algorithm might favor continuous study characteristics to categorical ones

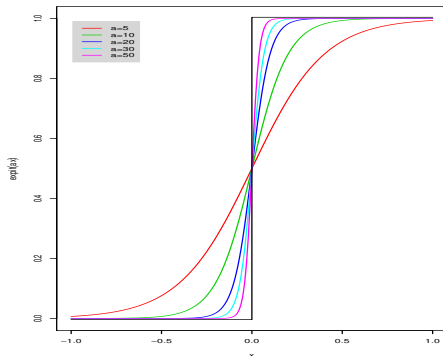
Solutions:

- 1 Splitting using look ahead
- 2 Use smooth sigmoid surrogate strategy

Tree construction, convergence and stability

```
REtree_sss <-REmrt(form_Levine, vi = vi, c.pruning = 0,  
  data = SpermCon, sss = TRUE)
```

Transform step functions into spline functions that resemble the step functions.



Problems:

- 1 The split may be based on a local optimum
- 2 The algorithm might favor continuous study characteristics to categorical ones
- 3 Subgroups are identified and tested using the same data

Solutions:

- 1 Splitting using look ahead
- 2 Use smooth sigmoid surrogate strategy
- 3 Permutation test of Q_B^*

Permutation procedure

Q_B^* obtained via permutation is used to determine the statistical significance of the between-subgroups heterogeneity.

```
set.seed(111)
REtree_perm <- REMrt(form_Levine, vi = vi, c.pruning = 0, data = SpermCon,
                     lookahead = TRUE, sss = TRUE, perm = TRUE)
summary(REtree_perm)
```

Test for Between-Subgroups Heterogeneity under RE assumption:

qb = 61.371 (df = 2), Permutation p-value 0.001996;

The estimate without bootstrap for the residual heterogeneity tau2 = 227.914

Random Effects Meta-analysis Results without Bootstrap:

	K	est	se	zval	pval	ci.lb	ci.ub	
2	69	76.530	2.021	37.875	0.000	72.570	80.490	***
4	87	90.627	1.900	47.701	0.000	86.904	94.351	***
5	88	70.764	1.739	40.700	0.000	67.357	74.172	***

Problems:

- 1 The split may be based on a local optimum
- 2 The algorithm might favor continuous study characteristics to categorical ones
- 3 Subgroups are identified and tested using the same data
- 4 CIs of summary effect sizes in subgroups are too optimistic

Solutions:

- 1 Splitting using look ahead
- 2 Use smooth sigmoid surrogate strategy
- 3 Permutation test of Q_B^*
- 4 Bootstrap-based confidence intervals

Bias correction bootstrap procedure

New bootstrap procedure to reduce bias in the estimate of the confidence intervals in the subgroups.

```
set.seed(111)
Results_boot <- BootCI(REplustree, nboot = 50)

summary(Results_boot)
```

Random Effects Meta-analysis Results with Bootstrap:

	K	est	se	ci.lb	ci.ub
2	69	77.378	2.635	72.213	82.542
4	87	91.154	2.435	86.381	95.926
5	88	71.029	2.290	66.540	75.519

Compare with previous result:

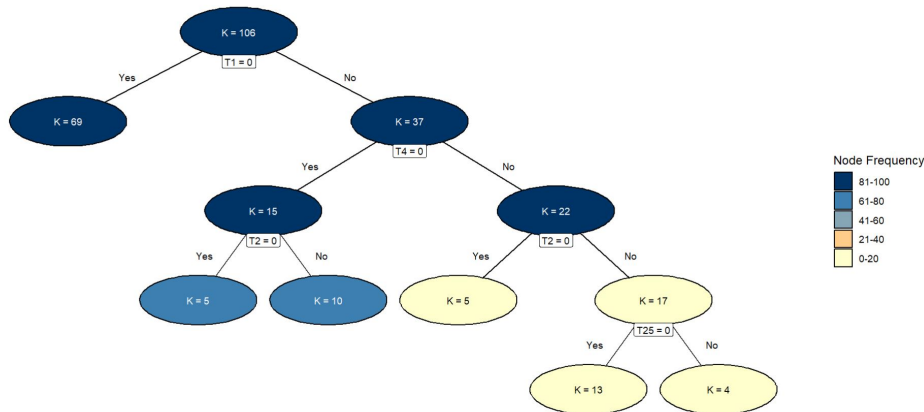
Random Effects Meta-analysis Results without Bootstrap:

	K	est	se	zval	pval	ci.lb	ci.ub	
2	69	76.530	2.021	37.875	0.000	72.570	80.490	***
4	87	90.627	1.900	47.701	0.000	86.904	94.351	***
5	88	70.764	1.739	40.700	0.000	67.357	74.172	***

Visualization tree stability

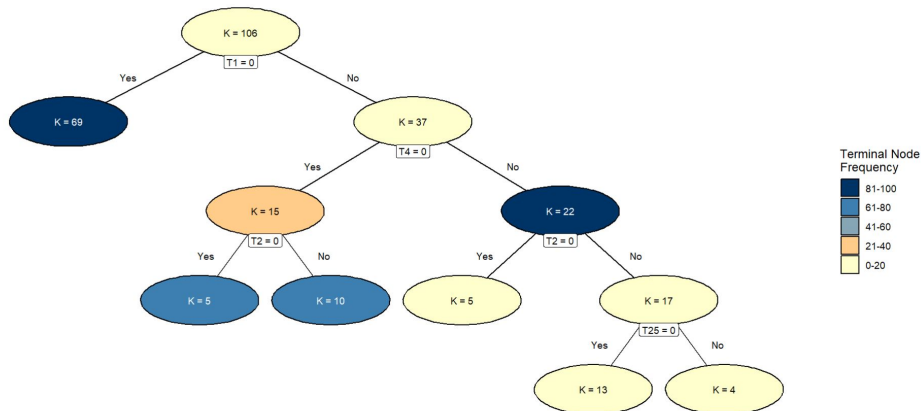
Check stability of the tree using Cross Validation:

```
PLOT_PV<-plotPV(REtree, iter = 100, c.pruning = 0)
```



Visualization tree stability

Check the stability of the tree using Cross Validation:



Thank you for your attention!

<https://cran.r-project.org/package=metacart>

