



Universiteit
Leiden
The Netherlands

Multiple moderator meta-analysis using the R-package Meta-CART

Li, X.; Dusseldorp, E.M.L.; Su, X.; Meulman, J.J.

Citation

Li, X., Dusseldorp, E. M. L., Su, X., & Meulman, J. J. (2020). Multiple moderator meta-analysis using the R-package Meta-CART. *Behavior Research Methods*, 52(6), 2657-2673. doi:10.3758/s13428-020-01360-0

Version: Publisher's Version

License: [Creative Commons CC BY 4.0 license](https://creativecommons.org/licenses/by/4.0/)

Downloaded from: <https://hdl.handle.net/1887/3160874>

Note: To cite this publication please use the final published version (if applicable).



Multiple moderator meta-analysis using the R-package Meta-CART

Xinru Li¹ · Elise Dusseldorp² · Xiaogang Su³ · Jacqueline J. Meulman^{1,4}

© The Author(s) 2020

Abstract

In meta-analysis, heterogeneity often exists between studies. Knowledge about study features (i.e., moderators) that can explain the heterogeneity in effect sizes can be useful for researchers to assess the effectiveness of existing interventions and design new potentially effective interventions. When there are multiple moderators, they may amplify or attenuate each other's effect on treatment effectiveness. However, in most meta-analysis studies, interaction effects are neglected due to the lack of appropriate methods. The method meta-CART was recently proposed to identify interactions between multiple moderators. The analysis result is a tree model in which the studies are partitioned into more homogeneous subgroups by combinations of moderators. This paper describes the R-package **metacart**, which provides user-friendly functions to conduct meta-CART analyses in R. This package can fit both fixed- and random-effects meta-CART, and can handle dichotomous, categorical, ordinal and continuous moderators. In addition, a new look ahead procedure is presented. The application of the package is illustrated step-by-step using diverse examples.

Keywords Meta-analysis · Heterogeneity · Interaction between moderators · CART · Fixed effect · Random effects · Computer software

Introduction

Methodology for synthesizing findings from multiple studies addressing the same research question has a long history (Hedges, 1981; Hedges & Olkin, 1985). The typical goals of meta-analysis are to estimate the overall effect size (i.e., a weighted average of study effect sizes), to quantify the heterogeneity in the study effect sizes, and to investigate the study characteristics that explain the

heterogeneity (i.e., moderators). The relationship between moderators and the study effect sizes can be of high interest for behavioral scientists to evaluate existing interventions and to design new potentially effective interventions. In the meta-analysis framework, moderator analysis can be conducted with several stand-alone software programs such as **Comprehensive Meta-Analysis** (Borenstein et al., 2009), **Meta-Analyst** (Wallace et al., 2009), and the Cochrane Collaboration's **RevMan** (Collaboration, 2014). Also, there are add-ins/macros/packages that can be used to conduct moderator analysis in many software languages such as **Stata** (StataCorp 2017, for details, see Sterne et al. 2008), **SAS** (Inc, 2002, for details, see Wang & Bushman, 1999; Arthur et al., 2001, and Sheu & Suzuki, 2001), **SPSS** (IBM Corp, 2013, for details, see Lipsey & Wilson, 2001), and **R** (Team, 2017), e.g., package **rmeta** by Lumley, 2012, **mvmeta** Gasparrini et al. (2012), and **metafor** by Viechtbauer (2010)). Most of these programs/add-ins are based on meta-regression, which is the most commonly used method for moderator analysis in meta-analysis.

In practice, behavioral scientific meta-analyses often have multiple moderators to be examined. For example, interventions to change health-related behavior generally include various behavior change techniques (BCTs), and researchers are interested in investigating the influence of

Electronic supplementary material The online version of this article (<https://doi.org/10.3758/s13428-020-01360-0>) contains supplementary material, which is available to authorized users.

✉ Elise Dusseldorp
elise.dusseldorp@fsw.leidenuniv.nl

- ¹ Mathematical Institute, Leiden University, P.O. Box 9512, 2300 RA Leiden, The Netherlands
- ² Institute of Psychology, Leiden University, P.O. Box 9555, 2300 RB, Leiden, The Netherlands
- ³ Department of Mathematical Sciences, University of Texas at El Paso, Austin, TX 78712, USA
- ⁴ Department of Statistics, Stanford University, Stanford, CA 94305, USA

BCTs on the effectiveness of interventions (see Michie et al. 2009). However, meta-regression has several limitations in such cases. First, when the number of included studies is small, meta-regression suffers from low statistical power to examine all moderators simultaneously (Tanner-Smith & Grant, 2018). Second, meta-regression has difficulties in exploring interaction effects among moderators since it requires moderators and their interactions to be specified beforehand. When there are no a priori hypotheses available, the number of all possible interaction terms are usually too large to be included in one model. The interaction effects, however, can provide valuable information to answer questions such like “do these intervention components amplify or attenuate each other’s effectiveness?” and “which combination of study characteristics results in the highest effectiveness?”.

To overcome the aforementioned limitations, tree-based models can be integrated into the framework of meta-analysis. Tree-based methods were introduced for the first time by Morgan and Sonquist (1963) in a method called automatic interaction detection (AID), and were fully developed in classification and regression trees (CART) by Breiman et al. (1985). Trees are good at dealing with many predictor variables that may interact, and produce results that can be easily interpreted. Tree-based methods have been used in the field of behavioral and medical sciences (for details, see Finch et al., 2011; Leach et al., 2016; Trujillano, Badia, Serviá, March, & Rodriguez-Pozo, 2009), but the idea of using trees in the meta-analysis framework is relatively new. For individual patient data (IPD) meta-analyses, Mistry et al. (2018) proposed a recursive partitioning method called IPD-SIDES that identifies patient subgroups by individual characteristics that may be related to the response to intervention. For aggregated data meta-analyses, a method called meta-CART was proposed to identify interaction effects among study-level characteristics (Dusseldorp et al., 2014; Li et al., 2017; 2019). This paper focuses on moderator analyses on aggregated meta-analysis by meta-CART. Compared to meta-regression, meta-CART has several advantages: first, it excels at dealing with interaction effects, and the interactions can be easily interpreted; second, it has automatic variable selection and does not require model selection; third, it is able to handle non-linear associations between moderators and effect size. Furthermore, since tree models are invariant to monotone transformation of predictors, meta-CART can keep the ordering information of ordinal moderators, whereas meta-regression usually codes ordinal variables the same as categorical variables. Li et al. (2019) showed via a simulation study that with a sufficiently large sample size (i.e., $n \geq 40$ for simple interactions and $n \geq 80$ for complex interactions), meta-CART has a good control of type I error rate (≤ 0.05), and

can achieve satisfactory power and recovery rates (i.e., ≥ 0.80). The meta-CART method has been acknowledged as a potential alternative statistical method for meta-regression to understand the combined effects of moderators (O’Brien et al., 2015; Michie et al., 2015; Tipton et al., 2018), and it has been applied in several meta-analytic studies (e.g., Bull et al., 2018; van Genugten et al. 2018).

In the present paper, we introduce the R package **metacart**, which implements the meta-CART method. This package provides user-friendly functions to perform meta-CART analysis for various types of moderators (i.e., continuous, ordinal, and categorical variables), and includes various additional options such as tuning the pruning of the tree model, restricting the minimum number of studies in a subgroup, and so on. In addition, we developed a new option to apply a “look-ahead” strategy specifically focusing on interaction detection. This paper aims to provide a general overview of the capabilities of the **metacart** package for conducting multivariate moderator meta-analysis with R. “Meta-CART method” introduces the meta-CART method. “The metacart package” describes the main functions in the **metacart** package, and “Examples” illustrates their practical usage with examples of real meta-analyses. “Conclusions” contains concluding remarks.

Meta-CART method

The goal of meta-CART analysis

The underlying goal of meta-CART analysis is to identify subgroups defined by the moderators that can explain the heterogeneity in the study effect sizes. Appropriate data for meta-CART include an outcome variable of interest (i.e., study effect size), the within-study sample variance of the effect size, and the potential moderator variables that may influence the effect size. Depending on the type of study, there is a variety of different effect-size measures, including the odds ratio, the relative risk, the correlation coefficient, and (the standardized) mean difference. Meta-CART can deal with all these effect-size measures as long as all the studies use the same measure¹. To identify influential moderators, meta-CART partitions the studies into subgroups that are more homogeneous with respect to their study effect sizes. The result is a tree model with the terminal nodes as the identified subgroups and the splitting variables as the influential moderators. Figure 1 shows an example of a tree model fitted by the package **metacart**. The root node (i.e., the node at the top) represents all the studies that are included in the analysis. From the root node,

¹Most effect-size measures can be converted to one another (see Hedges & Olkin 1985).

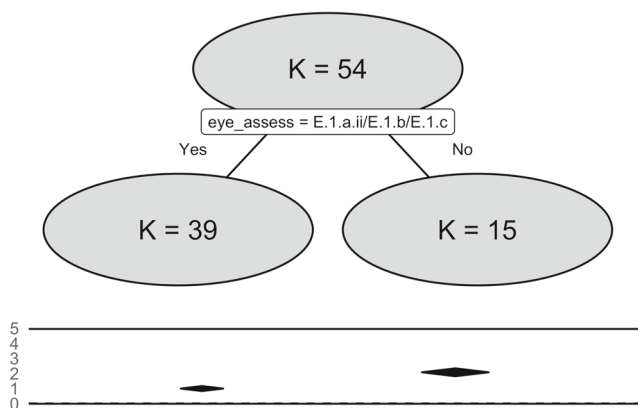


Fig. 1 The meta-CART analysis results of 54 samples examining the influence of moderators on the association between handedness and eye-dominance. The figure shows the main effect of the method to assess eye-dominance, which partitions the samples in two subgroups. The *two solid lines* show the range of the effect sizes of all the studies. The *diamonds* between the solid lines present the 95% confidence intervals of the summary effect sizes

the studies are partitioned into two subgroups by applying a threshold on the values of a moderator. For example, the root node in Fig. 1 is partitioned into two child nodes based on the moderator “eye_assess”. This moderator is categorical, and the studies of which “eye_assess” equals to “E.1.a.ii”, “E.1.b”, or “E.1.c” fall into the left child node and the other studies fall into the right child node. If the moderator is an ordinal or a continuous variable, a binary question such as “is the value of the moderator smaller than the split point?” will be asked to introduce a split. Note that the moderators and the corresponding split points are automatically selected by the algorithm. How the moderators and split points are selected will be explained in “[Meta-CART algorithm](#)”.

The resulting subgroup memberships are defined by a hierarchically nested set of decision rules such as “does the intervention include certain treatment components?” or “is the average age of the patients smaller than 25?”. In the identified subgroups, the summary effect sizes and their confidence intervals are estimated.² The analysis results can be used by researchers to predict the effect size given the study characteristics, or to identify the combination of study characteristics that results in the highest/lowest effect size. Note that meta-CART has an exploratory nature, since the subgroups are not predefined but identified from the data. Thus, meta-CART should be used as a hypothesis-generating tool, and we recommend testing the identified moderators by confirmatory methodology in further studies.

²the estimated confidence intervals are normal-theory 95% CIs. The `metacart` package always provides 95% CIs for the estimated effect sizes. If user is interested in CIs of other confidence levels, they can be easily computed based on the standard error of the estimates.

Meta-CART algorithm

There are two types of meta-CART algorithm: fixed-effect (FE) meta-CART, a partitioning algorithm that ignores the residual heterogeneity unexplained by the moderators, and random-effects (RE) meta-CART, a partitioning algorithm that takes into account the residual heterogeneity. As in standard meta-analysis, the choice between FE and RE assumptions should be based on a researcher’s prior belief; in other words, a researcher should decide on which assumption to use before analyzing the data. In general, if a researcher believes that given the study characteristics the variance of the study effect sizes is merely due to the sampling variance³, FE meta-CART can be chosen. If there is no a priori information about the residual heterogeneity, RE meta-CART is recommended. A more general discussion about FE model and RE model in meta-analysis can be found in Borenstein et al. (2010) and Schmidt et al. (2009).

In this section, we describe the FE and RE meta-CART algorithms, which underly the `metacart` package.

Fixed effect meta-CART

FE meta-CART splits the studies into more homogeneous subgroups under the FE assumption; that is, moderator effects and the within-study sampling variance are the only two sources of the variation in study effect sizes. Denote the true effect size in the k^{th} study by δ_k , and denote the observed effect size in the k^{th} study by d_k .⁴ Under the FE assumption, the observed effect size is given by

$$d_k = \delta_k + \epsilon_k = \beta_0 + \beta_1 x_{1k} + \beta_2 x_{2k} + \dots + \beta_M x_{Mk} + \epsilon_k, \quad (1)$$

where x_{mk} ($m = 1, \dots, M$) specify the values on the M moderators of the k^{th} study, and the β s are the corresponding coefficients. The sampling error ϵ_k is assumed to be distributed as $\mathcal{N}(0, \sigma_{\epsilon_k}^2)$, where $\sigma_{\epsilon_k}^2$ is the sampling variance.

The sampling variance $\sigma_{\epsilon_k}^2$ depends on the within-study sample sizes and the type of effect size. Using Hedges’ g as the measure of effect size, the sampling variance of the k^{th} study is estimated by

$$\hat{\sigma}_{\epsilon_k}^2 = \frac{n_k^T + n_k^C}{n_k^T n_k^C} + \frac{d_k^2}{2(n_k^T + n_k^C)}, \quad (2)$$

³This implies that differences between the studies are all accounted for by moderators, and sampling variance accounts for the variance of the study effect sizes adjusted for moderator effects.

⁴Note that we focus in the description on the d -family of effect sizes, that is, measures of the standardized mean outcome difference between treatment and control groups (e.g., Cohen’s d or Hedges’ g). However, the package `metacart` can also be used for other effect size measures such as log odds ratio (see the example in “[New approach meta-CART compared to meta-regression](#)”), log relative risk, and Fisher’s z .

where n_k^T, n_k^C are the treatment and control group sample sizes of the k th study, respectively (Hedges & Olkin, 1985, p. 86).

The summary effect size is computed as the weighted mean, with weights $w_k = 1/\sigma_{\epsilon_k}^2$:

$$d_+ = \frac{\sum d_k/\sigma_{\epsilon_k}^2}{\sum 1/\sigma_{\epsilon_k}^2}. \quad (3)$$

The measure of heterogeneity, the Q -statistic, is given by

$$Q = \sum_{k=1}^K \frac{(d_k - d_+)^2}{\sigma_{\epsilon_k}^2}. \quad (4)$$

Starting from one group including all the studies (i.e., the root node), FE meta-CART partitions the root node into two subgroups (i.e., offspring nodes), by searching through all possible splits and finds the moderator with corresponding split point that maximizes the between-subgroups heterogeneity. Denote the summary effect size of the t^{th} subgroup (i.e., t^{th} node in the tree) by d_{t+} , the between-subgroups heterogeneity measure is computed as

$$Q_B = \sum_t^{|T|} \sum_{k \in t} \frac{(d_{t+} - d_{++})^2}{\sigma_{\epsilon_k}^2}, \quad (5)$$

where $|T|$ is the total number of subgroups⁵, and d_{++} is the weighted grand mean of the parent node.

To grow a tree, FE meta-CART recursively searches the split that maximizes the heterogeneity Q_B between the left and right child nodes. After each split, the algorithm partitions a parent node into two child nodes. The tree-growing process is a recursive partitioning procedure since the same operation can be applied to any child node itself without taking the other nodes in the current tree into account. A flowchart of the tree growing process of FE meta-CART is shown in Fig. 2. The splitting process continues until all terminal nodes contain fewer than K_0 studies, where K_0 is a user-specified threshold (see “The metacart package”). To prevent overfitting, the initial tree will be iteratively pruned to a nested sequence of subtrees,

from which a best-sized tree is chosen via cross-validation. Different rules can be applied to determine the best-sized tree. To generalize the rules, a parameter c is introduced to select the “optimal” tree by using the $c \cdot SE$ rule (Dusseldorp et al., 2010). The $c \cdot SE$ rule selects the smallest tree with cross-validation error within the minimum cross-validation error plus its standard error multiplied by c .

The final analysis results consist of the selected tree model that represents the interactions between moderators, the estimates of the summary effect sizes d_{t+} within each subgroup, and the between-subgroups Q_B with $df = |T| - 1$. It should be noted that the significance test based on the between-subgroups Q -statistic is an overoptimistic pseudo Q -test. Given pre-defined subgroup membership, the between-subgroups Q -statistic follows a Chi-square distribution with $df = |T| - 1$ under the null hypothesis that there is no significant heterogeneity between the subgroups. However, in meta-CART the subgroup membership is not pre-defined, but identified by the tree-growing and cross-validation procedures. Therefore, over-optimism exists in the Q -test, and the p value should not be interpreted as the face value. It only gives information when the moderator effects are not significant. In other words, a non-significant p value indicates that the influence of the identified moderator(s) on study effect sizes is not significant, but a significant p value does not confirm the significance of the moderator effects. It is also worth to note that despite the over-optimism in the Q -test, the type I error rate of meta-CART (i.e., defined as the rate of finding a nontrivial tree with a significant between-subgroups Q -statistic while there is no moderator effect in the true structure underlying the data) is not inflated, because the type I error of moderator effects is mainly controlled by the pruning procedure (for details, see Li et al., 2019).

RE meta-CART

RE meta-CART takes the residual heterogeneity unexplained by moderators into account. The RE model is expressed as below:

$$d_k = \beta_0 + \beta_1 x_{1k} + \beta_2 x_{2k} + \dots + \beta_M x_{Mk} + \tau_k + \epsilon_k, \quad (6)$$

where τ_k is the variation introduced by the residual heterogeneity. Denote the residual heterogeneity by σ_τ^2 , with τ_k distributed as $\mathcal{N}(0, \sigma_\tau^2)$. Note that there are two random components in the RE model: τ_k and ϵ_k , where τ_k is the difference between the true effect size of the k th

⁵ $|T| = 2$ when only considering the heterogeneity between left and right child nodes after a split of the parent node.

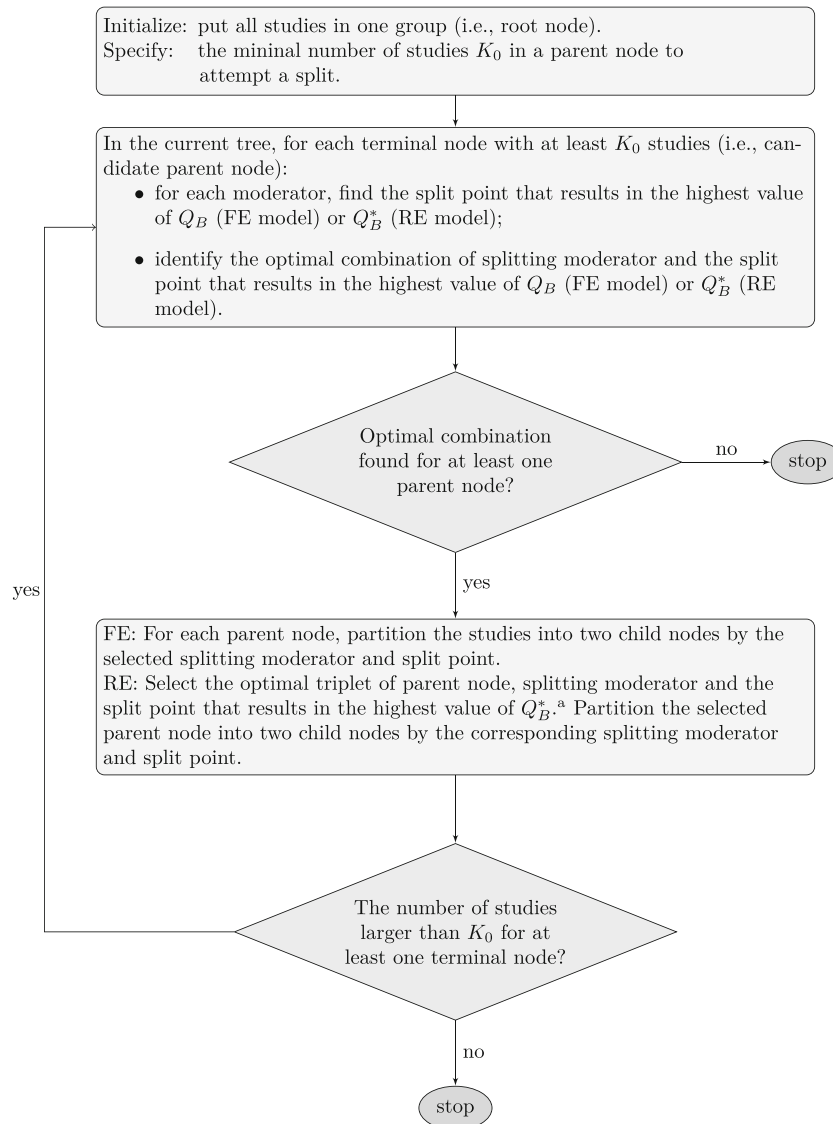


Fig. 2 Flowchart of tree growing process of meta-CART algorithms for fixed-effect (FE) model and random-effects (RE) model. ^a Note that this substep is not needed for recursive partitioning method like ordinary CART and FE meta-CART, but necessary for RE meta-CART due to the re-estimation of σ_{τ}^2

study and the mean of the population from which all study effect sizes are sampled, and ϵ_k is the sampling error of the observed effect size d_k as an estimate of the true effect size of the k th study (Hedges & Vevea, 1998). These two sources of variance are estimated in a hierarchical manner (Erez et al., 1996). The first level accounts for the within-study variation (or the sampling variance) $\sigma_{\epsilon_k}^2$, which can be estimated as [2]. The second level accounts for the between-study variance (or the residual heterogeneity) σ_{τ}^2 . There are various estimators for σ_{τ}^2 , including the Hunter–Schmidt estimator (Schmidt & Hunter, 2014), the Hedges estimator (Hedges & Olkin, 1985), the DerSimonian–Laird estimator (DerSimonian & Laird, 1986), the Sidik–Jonkman estimator

(Sidik & Jonkman, 2005), and the maximum-likelihood or restricted maximum-likelihood estimator (Viechtbauer, 2005). In the **metacart** package, we use the DerSimonian–Laird estimator for its lower computational cost.⁶

With the estimated residual heterogeneity, the summary effect size can be computed with the RE weights $w_k^* = 1/(\sigma_{\epsilon_k}^2 + \sigma_{\tau}^2)$:

$$d_+^* = \frac{\sum d_k / (\sigma_{\epsilon_k}^2 + \sigma_{\tau}^2)}{\sum 1 / (\sigma_{\epsilon_k}^2 + \sigma_{\tau}^2)}. \quad (7)$$

⁶In our pilot simulation study, it was found that the different choices of the estimator led to similar trees.

The RE heterogeneity is given by

$$Q^* = \sum_{k=1}^K \frac{(d_k - d_+^*)^2}{\sigma_{\epsilon_k}^2 + \sigma_\tau^2}. \quad (8)$$

Similar to FE meta-CART, RE meta-CART starts from the root node, and searches for the split that maximizes the between-subgroups heterogeneity. The difference is that before searching for the split, RE meta-CART needs to estimate the residual heterogeneity (σ_τ^2) first, since the RE between-subgroups heterogeneity is given by

$$Q_B^* = \sum_t^{|T|} \sum_{k \in t} \frac{(d_{t+}^* - d_{++}^*)^2}{\sigma_{\epsilon_k}^2 + \sigma_\tau^2}. \quad (9)$$

To continue the splitting process, RE meta-CART updates the estimate for σ_τ^2 and searches for the new split maximizing the partitioning criterion. A flowchart of the tree-growing process of RE meta-CART is also given in Fig. 2. Note that a split of a node will globally affect the estimation of σ_τ^2 and the value of Q_B^* . Thus, in contrast to FE meta-CART, RE meta-CART considers heterogeneity between all the terminal nodes rather than only between the resulting left and right child nodes after a split. As a result, the tree-growing process of RE meta-CART is not fully recursive, since the algorithm needs to take all terminal nodes of the current tree into account to introduce a new split. In other words, the estimate of σ_τ^2 and the optimal choice for a new split depend on the sequence of previous splits. Instead, RE meta-CART applies a sequential partitioning algorithm.⁷

As in FE meta-CART, the splitting process of RE meta-CART continues until a large tree is grown. Then, an optimally sized subtree is selected using cross-validation with the $c \cdot SE$ rule. The associated between-subgroups Q_B^* , the estimates for residual heterogeneity σ_τ^2 , and the within-subgroup summary effect sizes d_{j+}^* are obtained as the final tree is selected.

⁷A sequential tree-growing algorithm is defined as an algorithm that only splits a node in the tree if the reduction in the partitioning criterion at the proposed node exceeds some value, regardless of what might happen lower in the tree (Hand, 1997, Chapter 4). Recursive partitioning can be seen as a restricted form of sequential partitioning where at each node, exactly the same splitting procedure is repeated.

Look-ahead strategy for RE meta-CART

The growing algorithms of both FE and RE meta-CART are fully greedy. Although the approach guarantees a locally optimal solution at each split, it does not guarantee a globally optimal solution for the whole tree. That is, the tree-growing procedure chooses each optimal split with no regard of future splits. Because the estimated residual heterogeneity is influenced by the sequence of partitioning, the RE meta-CART is more sensitive to this local optimization problem. To alleviate this problem, we propose a look-ahead strategy for RE meta-CART, which examines two steps ahead instead of one at the split of the root node. This look-ahead strategy is applied only to the top level of a tree.

Starting the algorithm at the root node, a standard RE meta-tree chooses the single split (i.e., one split point on one moderator) that maximizes the between-subgroups heterogeneity. In contrast, a look-ahead strategy searches for the combination of two splits (i.e., single split points on two moderators or two split points on one moderator) that maximizes the partitioning criterion. As a result, the entire growing procedure that applies a look-ahead strategy consists of two sub-procedures. The first sub-procedure grows a tree with two splits, which searches for the optimal combination of a split of the root node and a split of one of its child nodes. The combination that maximizes Q_B^* is chosen. In the second sub-procedure, the resulting offspring nodes of the first two splits are split following a fully greedy procedure for maximizing Q_B^* at each split. This complete splitting procedure grows a large tree, which will be pruned using the pruning procedure mentioned above.

It should be noted that such a strategy does not guarantee a globally optimal solution, and only partially addresses the local optimization issue. Although more steps to look ahead can be beneficial, we chose the number of steps as two because the computational burden of a look-ahead procedure grows exponentially as the number of steps increases (Esmeir & Markovitch, 2007).

The metacart package

The **metacart** package provides functions to perform both FE and RE meta-CART analysis. The package is available via the Comprehensive R Archive Network (CRAN) at <https://cran.r-project.org/package=metacart>, and can be directly installed within R by typing `install.packages("metacart")`. The current version is 2.0-0.

Functions for FE meta-CART

The function `FEMrt` is the main function to perform a fixed effect meta-CART analysis. Both the tree growing and pruning processes in the function `FEMrt()` are based on the function `rpart()` in the `rpart` package (Therneau et al., 2017), with adaptations to fit a tree model applying appropriate weights on meta-analytic data (see “Examples”) and automatically compute subgroup meta-analysis results within the function. We now describe the different arguments of the function.

```
FEMrt(formula, data, vi, subset, c = 1,
control = rpart.control(minsplit = 6,
minbucket = 3, cp = 1e-04, xval = 10),
...)
```

The argument `formula` defines the outcome variable (i.e., effect size) and the predictor variables (i.e., moderators).

The argument `data` specifies the name of the data set to be analyzed. The argument `vi` requires the column name of the sampling variance of the effect size.

The argument `subset` is an optional expression to select a subset of the data to fit the model. The subset argument can either be a logical or a numeric vector indicating the indices of rows (i.e., studies) to be included.

The argument `c` refers to the pruning parameter (described in “Fixed effect meta-CART”) to be used for the analysis. The default value $c = 1$ corresponds to the one-standard-error rule recommended by Breiman et al. (1985). For meta-CART analysis, the recommended value of c depends on the type of research at hand and the number of studies. In general, if a strict control of the type I error (less than or equal to 0.05) is required, a pruning rule using $c = 1$ can be applied when the number of studies $K < 80$, and $c = 0.5$ when $K \geq 80$ (for more guidelines about the value of the pruning parameter c , (see Li et al., 2019).

The argument `control` specifies the options to control the growing and pruning processes. As mentioned above, both the tree growing and pruning processes are based on the function `rpart`, and therefore the control argument should be an `rpart.control` object (see details in Therneau et al., 2017). Within the `rpart.control` object, `minsplit` specifies the minimum number of studies that must exist in a parent node for a split to be attempted; `minbucket` specifies the minimum number of studies in any terminal node;

`cp` specifies the minimal improve of complexity parameter (i.e., Q_B divided by Q) to make a split in the growing process; `xval` specifies the number of cross-validations in the pruning process. The default value of `minbucket` is chosen as 3, because a sample size of one or two is too small to produce reliable subgroup meta-analysis results. Consequently, the default value of `minsplit` is chosen as 6. The default value `xval = 10` corresponds to the ten-fold cross-validation recommended by Breiman et al. (1985).

The output of the `FEMrt` function is an S3 object of class “`FEMrt`”. The corresponding print and summary methods can be used to display and inspect the elements of the object. The plot method can be used to present the main effects of identified moderators and interaction effects between them in a tree model. The predict method of the S3 object allows the user to predict effect size given the value of moderators. Examples to apply these methods will be described in “Examples”.

Functions for RE meta-CART

The function `REmrt` is the main function to perform random effects meta-CART analysis. Both the tree growing and pruning processes in the function `REmrt` are entirely new R-code combined with C++ that implements the sequential partitioning algorithm described in “RE meta-CART”. The different arguments of the function `REmrt` are described below.

```
REmrt(formula, data, vi, c = 1,
maxL = 5, minsplit = 6, minbucket = 3,
cp = 1e-4, xval = 10, lookahead = FALSE,
...)
```

The arguments `formula`, `data`, `vi`, and `c` are similar to the corresponding arguments in the function `FEMrt`. The argument `maxL` is an option to define the maximum number of splits in the tree growing process. The arguments `minsplit`, `minbucket`, `cp`, and `xval` are other options to control the growing and pruning processes. These options are not given in an `rpart.control` object, because the `REmrt` is an entirely new function and does not depend on the `rpart` function. The specifications of these options are the same as those from the function `FEMrt`.

The argument `lookahead` is a logical indicator to specify whether to apply the look-ahead strategy described in “Examples”.

The output of the `REmrt` function is an S3 object of class "REmrt". Similar to `FEmrt`, there are corresponding print, summary, plot, and predict methods for "REmrt" objects. Examples will be described in "Examples".

Examples

New approach meta-CART compared to meta-regression

In this example, we re-analyze the data "dat.bourassa1996" included in the `metafor` package (Viechtbauer, 2010). The data set contains the meta-analytic data from Bourassa et al. (1996), including results from 47 studies on the association between handed-

ness and eye-dominance. Some studies included multiple (independent) samples, resulting in 54 samples in total. Furthermore, for some studies, the combined data of the males and females were further broken down into the two sub-groups. As a result, the data set contains 96 (sub)samples in total. We only selected the independent samples with the combined data of both males and females ($K = 54$). The results of each of these samples were given in terms of the number of left-handed left-eyed, left-handed right-eyed, right-handed left-eyed, and right-handed right-eyed individuals. We use the log odds-ratio as the measure of effect size, which is the same as in Bourassa et al. (1996). A higher log odds-ratio indicates a higher association between handedness and eye dominance. First, we compute the effect size and sampling variance by using the `escalc` function in the `metafor` package.

```
R> library(metafor)
R> data("dat.bourassa1996")
R> dat.handedness <- escalc(measure = "OR", subset = (sex == "combined"),
+                           ai = lh.le, bi = lh.re, ci = rh.le, di = rh.re,
+                           data = dat.bourassa1996, add = 1/2, to = "all")
R> head(dat.handedness)
```

study	sample	author	year	selection	investigator	hand_assess	eye_assess	mage	lh.le	lh.re	rh.le	rh.re	sex	yi	vi
1	1	1 Mills	1925	no	other	questionnaire	E.1.d	NA	93	17	130	760	combined	3.4384	0.0768
2	2	2 Downey	1927	yes	psychologist	questionnaire	E.1.b	37	140	91	305	697	combined	1.2544	0.0228
5	3	3 Miles	1930	no	psychologist	questionnaire	E.1.a.ii	22	16	14	43	114	combined	1.0970	0.1613
6	4	4 Quinan	1931	no	other	performance	E.1.b	25	102	97	597	1898	combined	1.2061	0.0222
7	5	5 Jasper	1932	yes	psychologist	questionnaire	E.1.a.ii	20	17	14	38	80	combined	0.9257	0.1645
8	6	6 Lund	1932	no	psychologist	performance	E.1.a.ii	20	10	2	52	170	combined	2.6130	0.5202

The computed effect size ("yi") and variance ("vi") are added to the data set. The data set provides information about the following moderators: the publication year ("year"), whether the selection of subjects was based on eye-dominance or handedness ("selection", with two categories), the type of investigator ("investigator", with three categories⁸), the method to assess handedness ("hand_assess", with two categories), the methods to assess eye-dominance ("eye_assess", with six categories), the average age ("mage").

Both meta-regression and meta-CART are applied to re-analyze this data set. The analyses results are compared to illustrate the different research questions that can be answered by these approaches. For both analyses, the moderators "selection", "investigator", "hand_assess", and "eye_assess" are selected. The random effects model is chosen and the DerSimonian-Laird estimator is used to estimate the residual heterogeneity. We use "set.seed(2018)" so that the analysis results described in this paper can be replicated.

First, to answer the question "which moderators affect hand-eye association?", we fit a meta-regression model with the main effects of the selected moderators.

⁸The three categories are "psychologist", "educationalist", and "other"

```
R> set.seed(2018)
R> regHAND.re <- rma(yi = yi, vi = vi, method = "DL",
+                   mods = ~ eye_assess + investigator + hand_assess +
+                   selection, data = dat.handedness)
R> regHAND.re
Mixed-Effects Model (k = 54; tau^2 estimator: DL)

tau^2 (estimated amount of residual heterogeneity):      0.1268 (SE = 0.0726)
tau (square root of estimated tau^2 value):              0.3561
I^2 (residual heterogeneity / unaccounted variability): 60.39%
H^2 (unaccounted variability / sampling variability):    2.52
R^2 (amount of heterogeneity accounted for):             58.26%
```

```
Test for Residual Heterogeneity:
QE(df = 44) = 111.0699, p-val < .0001
```

```
Test of Moderators (coefficient(s) 2:10):
QM(df = 9) = 51.3955, p-val < .0001
```

Model Results:

	estimate	se	zval	pval	ci.lb	ci.ub
intrcpt	1.3564	0.4468	3.0356	0.0024	0.4806	
eye_assessE.1.a.ii	-0.6513	0.3222	-2.0214	0.0432	-1.2828	
eye_assessE.1.b	-0.9618	0.3585	-2.6829	0.0073	-1.6645	
eye_assessE.1.c	-0.8967	0.3486	-2.5718	0.0101	-1.5800	
eye_assessE.1.d	0.6733	0.4381	1.5369	0.1243	-0.1854	
eye_assessE.2.a	0.2246	0.3898	0.5762	0.5645	-0.5394	
investigatorother	0.4302	0.3774	1.1400	0.2543	-0.3094	
investigatorpsychologist	0.4222	0.3510	1.2027	0.2291	-0.2658	
hand_assessquestionnaire	0.0595	0.1913	0.3108	0.7559	-0.3156	
selectionyes	0.0117	0.2392	0.0491	0.9609	-0.4571	
intrcpt	2.2322					**
eye_assessE.1.a.ii	-0.0198					*
eye_assessE.1.b	-0.2592					**
eye_assessE.1.c	-0.2133					*
eye_assessE.1.d	1.5320					
eye_assessE.2.a	0.9887					
investigatorother	1.1698					
investigatorpsychologist	1.1101					
hand_assessquestionnaire	0.4345					
selectionyes	0.4806					

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The output shows that three contrasts derived from the method to assess eye-dominance are significant: contrasted with the eye-dominance assessment method “E.1.a.i” (i.e., the reference method), “E.1.a.ii”, “E.1.b”, and “E.1.c” result in significantly lower hand-eye association. The remaining two methods, “E.2.a” and “E.1.d” do not differ significantly from the reference method. Note that the interpretation of these contrasts depends highly on the choice of reference method.

Further analyses are needed if we are interested in whether some categories of a moderator can be combined and whether interaction effects between the moderators are present. Alternatively, meta-CART can be performed to answer the question “which combinations of (categories of) moderators are influential?”. The combinations can be either combinations of multiple moderators, or combinations of multiple categories of one moderator. We fit a RE meta-CART model by

```
R> library(metacart)
Loading required package: rpart
Loading required package: ggplot2
Loading required package: gridExtra
R> modHAND.re <- REMrt(yi ~ eye_assess + investigator + hand_assess
+                       + selection, data = dat.handedness, vi = vi,
+                       c = 0.5, maxL = 10L, minsplit = 10L, cp = 1e-04,
+                       minbucket = 3, xval = 10, lookahead = FALSE)
R> summary(modHAND.re)
Random Effects meta-tree (K = 54 studies);
metacartv2::REMrt(formula = yi ~ eye_assess + investigator +
hand_assess + selection, data = dat.handedness, vi = vi,
c = 0.5, maxL = 10L, minsplit = 10L, cp = 1e-04, minbucket = 3,
lookahead = FALSE, xval = 10)

A tree with 2 terminal nodes was detected
Moderators were detected as: eye_assess

Test for Between-Subgroups Heterogeneity under RE assumption:
Qb = 46.293 (df = 1), p-value < 1e-04;
The estimate for the residual heterogeneity tau2 = 0.120;

Subgroup Meta-analysis Results:
  K    g    se   zval  pval  ci.lb  ci.ub
2 39 0.988 0.086 11.553 0.000 0.821 1.156 ***
3 15 2.071 0.134 15.438 0.000 1.808 2.334 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The plot function can be used to inspect the influence of eye-dominance assessment methods on the association between handedness and eye-dominance.

```
R> plot(modHAND.re)
```

The plot can be found in Fig. 1, which shows a tree with one split and two terminal nodes. Note that this tree results from pruning an initially large tree based on cross-validation (see 4). According to the summary results and the plot, the moderator “eye_assess” has a strong influence on the association between handedness and eye-dominance ($Q_B^* = 46.29$, $df = 1$, p value < 0.0001). When the eye-dominance is assessed using

the methods “E.1.a.ii”, “E.1.b”, or “E.1.c”, the observed association between handedness and eye-dominance is lower (log odds-ratio = 0.988, 95% CI: 0.821 to 1.156). When the eye-dominance is assessed using the methods “E.1.a.i”, “E.1.d”, or “E.2.a”, the observed association is generally higher (log odds-ratio = 2.071, 95% CI: 1.808 to 2.334). This is in accordance with the claims in Bourassa et al. (1996) that the methods to assess eye-dominance can be partitioned into two subgroups: (1) unbiased methods including “E.1.a.ii: monocular procedure with object/instrument held in both hands”, “E.1.b: binocular procedure”, and “E.1.c: a combination of the previous methods”, (2) biased methods including “E.1.a.i: monocular

procedure with object/instrument held in one hand”, “E.2.a: assessment based on a questionnaire”, and “E.1.d: some other method”. The methods “E.1.a.ii”, “E.1.b”, or “E.1.c” were symmetric in the sense that both hands are used equivalently during the process of measurement, and therefore are less likely to have measurement bias.

Comparing to meta-regression, which estimates the coefficients and tests the significance for the contrasts derived from “eye_assess”, meta-CART partitions the multi-categorical moderator “eye_assess” into subgroups, and tests the heterogeneity between the resulting subgroups. The subgrouping membership can be used to verify prior hypotheses⁹, or to generate hypotheses if no a priori hypotheses exist.

Identify interaction effects using FE/RE assumptions

In this example, we perform a meta-analysis to identify the most effective combination of treatment components by exploring the interactions between them. The **metacart** package provides the data object `dat.BCT2009` as a subset of the meta-analytic data from Michie et al. (2009). The meta-analysis by Michie et al. (2009) aimed to assess the effectiveness of interventions designed to promote physical activity and health eating, and investigated whether theoretically specified behavior change techniques (BCTs) improve the effectiveness. The subset used in this example consists of 106 interventions that included at least one of the motivation-enhancing BCTs.

```
R> data("dat.BCT2009")
R> head(dat.BCT2009, n = 5)
```

	study	g	vi	T1	T2	T3	T4	T25
1	ALDANA 2005 (PA)	0.60751	0.02469612	1	1	0	1	0
2	ANDERSON 2006	0.74651	0.05116192	0	0	0	1	0
3	ARAO 2007 (PA)	0.73090	0.06694639	0	0	0	1	0
4	BABAZONO 2007 (PA)	0.88197	0.09911163	0	0	0	1	0
5	BAKER 2008	0.73575	0.05308416	0	1	0	0	0

```
R> summary(dat.BCT2009)
```

study	g	vi	T1
ALDANA 2005 (PA) : 1	Min. : -0.1696	Min. : 0.00129	0:69
ANDERSON 2006 : 1	1st Qu.: 0.1291	1st Qu.: 0.01533	1:37
ARAO 2007 (PA) : 1	Median : 0.3033	Median : 0.02796	
BABAZONO 2007 (PA) : 1	Mean : 0.3308	Mean : 0.04985	
BAKER 2008 : 1	3rd Qu.: 0.4932	3rd Qu.: 0.06188	
BENNETT 2008 : 1	Max. : 1.2831	Max. : 0.45788	
(Other) : 100			
T2	T3	T4	T25
0:42	0:106	0:33	0:89
1:64		1:73	1:17

As displayed above, the data set contains information about the name and the publication year, the estimated effect size (g = Hedges' g , which denotes the standardized mean difference in treatment outcome), the sampling variance of the effect size, and whether specific BCTs were applied or not in a study (“0” for absent and “1” for present). In this example, we will re-analyze this data set focusing on the motivation-enhancing BCTs that may explain the

heterogeneity in the effect sizes of the interventions. We use the pruning parameter $c = 0.5$ for both FE and RE meta-CART analyses. Four moderators “T1: Provide information about behavior-health link”, “T2: Provide information on consequences”, “T4: Prompt intention formation”, and “T25: Motivational interviewing” are included in the meta-CART analysis. The moderator “T3: Provide information about other’s approval” is excluded since none of the studies applied this BCT.

If we have prior knowledge that the FE assumption is reasonable here, a FE meta-CART model can be fitted using the following code

⁹If the subgroup membership is tested as a categorical moderator, meta-regression yields the same results as meta-CART.

```
R> modBCT.fe <- FEmrt(g ~ T1 + T2 + T4 + T25, data = dat.BCT2009,
+                   vi = vi, c = 0.5, control = rpart.control(xval = 10,
+                   minbucket = 5, minsplit = 10, cp = 1e-04))
```

To obtain the summary output, we use

```
R> summary(modBCT.fe)
Fixed Effects meta-tree (K = 106 studies);
FEmrt(formula = g ~ T1 + T2 + T4 + T25, data = dat.BCT2009, vi = vi,
c = 0.5, control = rpart.control(xval = 10, minbucket = 5,
minsplit = 10, cp = 1e-04))

A tree with 3 terminal nodes was detected
Moderators were detected as: T1, T4

Test for Between-Subgroups Heterogeneity under FE assumption:
Qb = 40.589 (df = 2), p-value < 1e-04;

Subgroup Meta-analysis Results:
  K      Qw      g      se      zval      pval      ci.lb      ci.ub
2 69 168.397 0.202 0.014 14.521 0.000 0.175 0.229 ***
6 15  20.195 0.191 0.037  5.115 0.000 0.118 0.264 ***
7 22  23.855 0.438 0.035 12.534 0.000 0.369 0.506 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

This output shows that a tree with three terminal nodes was detected. The studies are partitioned into three subgroups based on two influential moderators “T1” and “T4”. The between-subgroups heterogeneity is significant ($Q_B = 40.589$, $df = 2$, p value < 0.0001). The subgroup analysis results show the number of studies (K), the within-subgroup Q -statistic, the summary effect sizes (g) in each subgroup, the standard errors of the summary effect sizes (se), Z -test statistics of the summary effect sizes, and the confidence intervals of the summary effect sizes. The `plot` function can be used to inspect the final tree, in this case, the interaction between the two moderators “T1” and “T4”.

```
R> plot(modBCT.fe)
```

The plot is shown in Fig. 3. If an intervention does not include “T1” (i.e., $T_1 = 0$ is true), then the intervention ends up in the left terminal node (with $K = 69$). For those that include “T1” but not include “T4”, they end up in the middle terminal node (with $K = 15$). Interventions including both “T1” and “T4” end up in the right terminal node (with $K = 22$). Combined with the other results, we can see that the summary effect size is the highest when “T1” and “T4” are both present in the intervention ($g = 0.438$, 95% CI: 0.369 to 0.506).

If a new intervention is designed with three BCTs “T1”, “T2” and “T25” (thus without “T4”), the prediction of its effect size based on this tree (`modBCT.fe`) can be obtained by the `predict` function.

```
R> newStudy <- data.frame(T1 = factor(1, levels = 0:1),
+                         T2 = factor(1, levels = 0:1),
+                         T4 = factor(0, levels = 0:1),
+                         T25 = factor(1, levels = 0:1))
R> predict(modBCT.fe, newStudy)
      g      ci.lb      ci.ub T1 T2 T4 T25
1 0.1906789 0.1176188 0.263739 1  1  0  1
```

When interpreting the FE meta-CART analysis results, it is important to realize that the FE assumption ignores the uncertainty introduced by the residual heterogeneity. As a result, the confidence intervals of the summary effect sizes

are more narrow than those estimated by using the RE model.

If we would like to take into account the residual heterogeneity, a RE meta-CART model can be fitted using the following commands

```
R> modBCT.re <- REmrt(g ~ T1 + T2 + T4 + T25, data = dat.BCT2009, vi = vi,
+                   c = 0.5, maxL = 10L, minsplit = 10L, cp = 1e-04,
+                   minbucket = 3, xval = 10, lookahead = FALSE)
Warning message:
In REmrt(g ~ T1 + T2 + T4 + T25, data = dat.BCT2009, vi = vi, c = 0.5, :
no moderator effect was detected
R> summary(modBCT.re)
Random Effects meta-tree (K = 106 studies);
REmrt(formula = g ~ T1 + T2 + T4 + T25, data = dat.BCT2009, vi = vi,
c = 0.5, maxL = 10L, minsplit = 10L, cp = 1e-04, minbucket = 3,
xval = 10, lookahead = FALSE)
```

No moderator effect was detected

Test for Heterogeneity

$Q = 253.0357$ (df = 105), p-value < 1e-04;

The estimate for the residual heterogeneity $\tau^2 = 0.023$;

Random Effects Meta-analysis Results:

K	g	se	zval	pval	ci.lb	ci.ub	
106	0.270	0.022	12.344	0.000	0.227	0.313	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

The message “no moderator effect was detected” indicates that the cross-validation procedure selected the best-sized subtree with no splits, which means that no influential moderators were identified. In this case, the summary method shows the standard RE meta-analysis results instead of the subgroup analysis results. The heterogeneity among the studies is significant ($Q_B^* = 253.036$, $df = 105$, p value < 0.0001). The estimated summary effect size for all studies ($K = 106$) is 0.270 (95% CI: 0.227 to 0.313).

Look-ahead strategy

The `metacart` package provides a simulated data set `dat.balanced` to illustrate the look-ahead strategy for RE meta-CART analysis. The simulated data set contains $K = 60$ studies with four moderators: x_1, x_2, x_3, x_4 , among which x_1, x_2 , and x_4 are randomly sampled dichotomous variables and x_3 is sampled from a uniform distribution $U(0, 1)$. The sample size n was generated from a normal distribution $\mathcal{N}(160, (\frac{160}{3})^2)$. The residual heterogeneity was set as $\sigma_\tau^2 = 0.01$. The true model used to generate data was

$$d_k = 0.5 \cdot I(x_1 = 0, x_2 = 1) + 0.5 \cdot I(x_2 = 0, x_1 = 1) + \tau_k + \epsilon_k. \quad (10)$$

This model is similar to the simulated example used by Tibshirani and Knight (1999). This model is shown

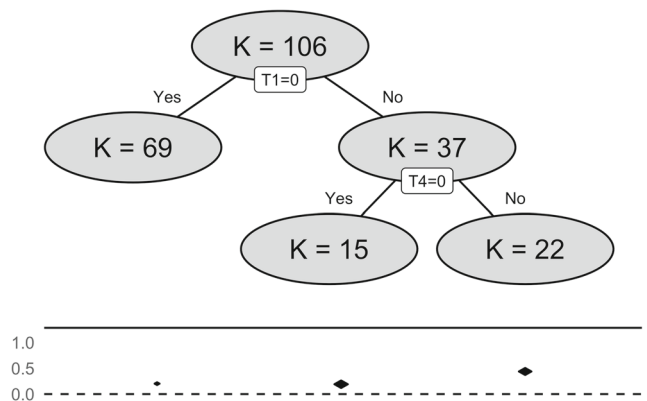


Fig. 3 The meta-CART analysis result of 106 studies that examine the influence of motivation-enhancing BCTs on healthy eating and physical activities. The figure shows the FE meta-CART structure with splitting information at each internal node and the number of studies in each subgroup implied by a terminal node. The two solid lines show the range of the effect sizes of all the studies. The diamonds between the solid lines present the 95% confidence intervals of the summary effect sizes.

by Tibshirani and Knight (1999) to be difficult for greedy search procedures like CART because there is no information on where to split at the top level. Due to the same reason, standard meta-CART is likely to end up with a local optimum solution in such case. In this example, we

would like to show that the look-ahead strategy described in “Examples” can alleviate this problem.

First, we inspect the data and fit a RE meta-CART model without using the look-ahead strategy.

```
R> data("dat.balanced")
R> head(dat.balanced)
      g          vi  x1 x2          x3 x4
1  0.08365950 0.02791743  0  0 0.03468667  1
2 -0.07713184 0.01252533  0  0 0.66439264  1
3  0.65561074 0.07416912  0  1 0.01766560  1
4  0.28081322 0.02305076  0  1 0.29408049  1
5  0.29945301 0.01215463  0  1 0.42494663  1
6  0.02193797 0.01452387  1  1 0.68668759  1
R> res <- REmrt(g ~ x1 + x2 + x3 + x4, vi = vi, data = dat.balanced,
+             c = 1, maxL = 5L, minsplit = 6L, cp = 1e-04,
+             minbucket = 3L, xval = 10, lookahead = FALSE)
Warning message:
In REmrt(g ~ x1 + x2 + x3 + x4, vi = vi, data = dat.balanced, :
no moderator effect was detected
```

The output shows that meta-CART failed to detect any moderator. The reason is that the algorithm got stuck at a local

optimum in the splitting procedure. This can be verified by inspecting the initial tree, unpruned by cross-validation.

```
R> res$initial.tree
      Qb          tau2      split  mod pnode
1  0.000000 0.09243657      <NA> <NA>    NA
2  3.040267 0.08828453  x3 < 0.91  x3     1
3  7.317024 0.08224968  x3 < 0.27  x3     2
4 18.303908 0.06847671  x3 <0.072  x3     4
5 22.521824 0.06555919  x3 <0.035  x3     6
6 24.970538 0.06449154  x4 = 1   x4     3
```

This output shows the between-subgroups Q -statistic, the residual heterogeneity, the chosen split points (by default the values are shown in two decimal places), the chosen moderator, and the parent node to be partitioned at each split, with the first row presenting the root node when no

split occurs. From the output we can see that the algorithm falsely chose the first splits at x_3 and ended up with $Q_B = 24.97$ after five splits.

Then we fit a RE meta-CART model using the look-ahead strategy.

```

R> res2 <- REmrt(g ~ x1 + x2 + x3 + x4, vi = vi, data = dat.balanced,
+               c = 1, maxL = 5L, minsplit = 6L, cp = 1e-04,
+               minbucket = 3L, xval = 10, lookahead = TRUE)
R> summary(res2)
Random Effects meta-tree (K = 60 studies);
metacartv2::REmrt(formula = g ~ x1 + x2 + x3 + x4, data = dat.balanced,
vi = vi, c = 1, maxL = 5L, minsplit = 6L, minbucket = 3L,
lookahead = T, cp = 1e-04, xval = 10)

A tree with 4 terminal nodes was detected
Moderators were detected as: x1, x2

Test for Between-Subgroups Heterogeneity under RE assumption:
Qb = 145.004 (df = 3), p-value < 1e-04;
The estimate for the residual heterogeneity tau2 = 0.017;

Subgroup Meta-analysis Results:
  K      g      se    zval  pval  ci.lb ci.ub
4 16 -0.054 0.044 -1.244 0.214 -0.140 0.031
5 17  0.516 0.045 11.413 0.000  0.427 0.605 ***
6 12 -0.017 0.051 -0.334 0.738 -0.117 0.083
7 15  0.518 0.046 11.358 0.000  0.429 0.607 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> plot(res2)

```

The results show that the RE meta-CART model with look-ahead strategy successfully recovered the true model as in (10). A tree with four terminal nodes was detected. As shown in Fig. 4, the terminal node with $x_1 = 0$ and $x_2 = 1$ ($K = 17$) and the terminal node with $x_1 = 1$ and $x_2 = 0$ ($K = 15$) have higher effect sizes with CIs (presented by

the diamonds) covering 0.5, whereas the terminal node with x_1 and x_2 both equal to 0 ($K = 16$) and the terminal node with x_1 and x_2 both equal to 1 ($K = 12$) have lower effect sizes with CIs covering 0.

The initial tree obtained with look-ahead strategy can be inspected by

```

R> res2$initial.tree

```

	Qb	tau2	split	mod	pnode
1	0.0000000	0.092436572	<NA>	<NA>	NA
2	0.2548616	0.092830849	x1 = 0	x1	1
3	42.0137849	0.049149501	x2 = 0	x2	2
4	145.0037965	0.016652079	x2 = 1	x2	3
5	188.0117000	0.011161345	x3 < 0.26	x3	6
6	213.7011113	0.008802224	x3 < 0.6	x3	4

Comparing this tree to the initial tree obtained from the previous model without the look-ahead strategy, the look-ahead strategy correctly finds the first

two splits, and obtains a solution with much larger between-subgroups Q -statistic ($Q_B = 213.70$ after five splits).

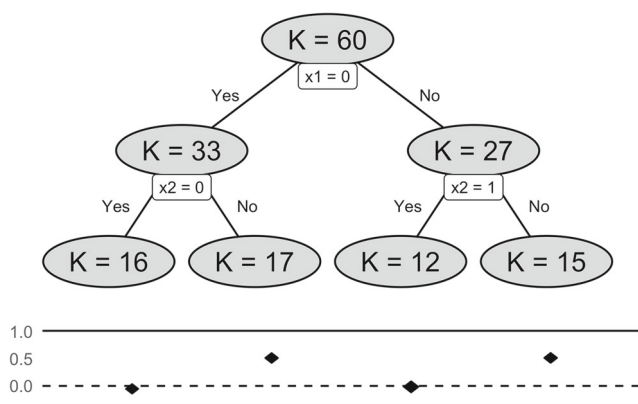


Fig. 4 The analysis results of simulated data generated by fitting a RE meta-CART model with look-ahead strategy. The analysis successfully recovered the true model that was used to generate the simulated data

Conclusions

This paper presents the **metacart** package written in R to perform meta-CART analysis both for fixed effect and random effects models. The algorithms and the main functions of the package **metacart** are described in “Meta-CART method” and “The metacart package”, respectively. Applications of the R-package **metacart** were illustrated through example analyses on two real-world data sets and one simulated data set in “Examples”.

One strength of the package **metacart** is that it can easily explore the interaction effects among multiple moderators using an interpretable tree model. Furthermore, for multi-categorical variables, it creates automatically the contrasts between (combinations of) categories that account for the highest amount of heterogeneity. Another strength is that **metacart** provides researchers various options for the tree growing and pruning processes. For example, researchers can choose the minimum number of studies in parent nodes and the minimum number of studies in terminal nodes based on the total sample size. In general, it is recommended to have at least three studies in each terminal nodes. The pruning parameter can be chosen based on the balance between power and type I error. A detailed guideline for choosing the pruning parameter can be found in Li et al. (2019). The look-ahead strategy is recommended to partially relieve the local optimization problem when fitting a RE meta-CART model.

In conclusion, this paper and the developed **metacart** package introduce researchers to the implementation of meta-CART analysis, to facilitate exploring interaction effects between multiple moderators in the framework of meta-analysis.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Arthur, W. J., Bennett, W., & Huffcutt, A. I. (2001). *Conducting Meta-analysis using SAS*. Mahwah: SAS Lawrence Erlbaum Associates Inc.
- Borenstein, M., Hedges, L. V., Higgins, J., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods, 1*(2), 97–111.
- Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Hoboken: Wiley.
- Bourassa, D., McManus, I. C., & Bryden, M. P. (1996). Handedness and eye-dominance: a meta-analysis of their relationship. *Laterality: Asymmetries of Body, Brain and Cognition, 1*(1), 5–34.
- Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1985). *Classification and regression trees*. Belmont: Wadsworth.
- Bull, E. R., McCleary, N., Li, X., Dombrowski, S. U., Dusseldorp, E., & Johnston, M. (2018). Interventions to promote healthy eating, physical activity and smoking in low-income groups: a systematic review with meta-analysis of behavior change techniques and delivery/context. *International Journal of Behavioral Medicine, 25*(6), 605–616.
- Collaboration, T. C. (2014). Review manager (RevMan 5.3) The Cochrane Collaboration, Copenhagen, The Nordic Cochrane Centre.
- IBM Corp. (2013). *IBM SPSS statistics for Windows, version 22.0*. Armonk: IBM Corp.
- DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials, 7*(3), 177–188.
- Dusseldorp, E., Conversano, C., & Van Os, B. J. (2010). Combining an additive and tree-based regression model simultaneously: STIMA. *Journal of Computational and Graphical Statistics, 19*(3), 514–530.
- Dusseldorp, E., van Genugten, L., van Buuren, S., Verheijden, M. W., & van Empelen, P. (2014). Combinations of techniques that effectively change health behavior: evidence from meta-CART analysis. *Health Psychology, 33*(12), 1530–1540.
- Erez, A., Bloom, M. C., & Wells, M. T. (1996). Using random rather than fixed effects models in meta-analysis: implications for situational specificity and validity generalization. *Personnel Psychology, 49*(2), 275–306.
- Esmeir, S., & Markovitch, S. (2007). Anytime learning of decision trees. *Journal of Machine Learning Research, 8*, 891–933.
- Finch, W. H., Chang, M., Davis, A. S., Holden, J. E., Rothlisberg, B. A., & McIntosh, D. E. (2011). The prediction of intelligence in preschool children using alternative models to regression. *Behavior Research Methods, 43*(4), 942–952.

- Gasparri, A., Armstrong, B., & Kenward, M. G. (2012). Multi-variate meta-analysis for non-linear and other multi-parameter associations. *Statistics in Medicine*, *31*(29), 3821–3839.
- Hand, D. J. (1997). *Construction and assessment of classification rules*. Hoboken: Wiley.
- Hedges, L. V. (1981). Distribution theory for glass's estimator of effect size and related estimators. *Journal of Educational and Behavioral Statistics*, *6*(2), 107–128.
- Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando: Academic Press.
- Hedges, L. V., & Vevea, J. L. (1998). Fixed-and random-effects models in meta-analysis. *Psychological Methods*, *3*(4), 486.
- Inc, S. I. (2002). *SAS/STAT Software, version 9.1*. Cary: SAS Institute Inc.
- Leach, H. J., O'Connor, D. P., Simpson, R. J., Rifai, H. S., Mama, S. K., & Lee, R. E. (2016). An exploratory decision tree analysis to predict cardiovascular disease risk in African American women. *Health Psychology*, *35*(4), 397.
- Li, X., Dusseldorp, E., & Meulman, J. J. (2017). Meta-CART: A tool to identify interactions between moderators in meta-analysis. *British Journal of Mathematical and Statistical Psychology*, *70*(1), 118–136.
- Li, X., Dusseldorp, E., & Meulman, J. J. (2019). A flexible approach to identify interaction effects between moderators in meta-analysis. *Research synthesis methods*, *10*(1), 134–152.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical Meta-analysis* Vol. 49. Thousand Oaks: Sage Publications.
- Lumley, T. (2012). *rmeta: Meta-analysis*. R package version 2.16.
- Michie, S., Abraham, C., Whittington, C., McAteer, J., & Gupta, S. (2009). Effective techniques in healthy eating and physical activity interventions: a meta-regression. *Health Psychology*, *28*(6), 690–701.
- Michie, S., Johnson, B. T., & Johnston, M. (2015). Advancing cumulative evidence on behaviour change techniques and interventions: a comment on Peters, de Bruin, and Crutzen. *Health Psychology Review*, *9*(1), 25–29.
- Mistry, D., Stallard, N., & Underwood, M. (2018). A recursive partitioning approach for subgroup identification in individual patient data meta-analysis. *Statistics in Medicine*, *37*(9), 1550–1561.
- Morgan, J. N., & Sonquist, J. A. (1963). Problems in the analysis of survey data, and a proposal. *Journal of the American Statistical Association*, *58*(302), 415–434.
- O'Brien, N., McDonald, S., Araújo-Soares, V., Lara, J., Errington, L., Godfrey, A., . . . , et al. (2015). The features of interventions associated with long-term effectiveness of physical activity interventions in adults aged 55–70 years: a systematic review and meta-analysis. *Health Psychology Review*, *9*(4), 417–433.
- Schmidt, F. L., & Hunter, J. E. (2014). *Methods of meta-analysis: correcting error and bias in research findings*. Newbury Park: Sage Publications.
- Schmidt, F. L., Oh, I.-S., & Hayes, T. L. (2009). Fixed-versus random-effects models in meta-analysis: model properties and an empirical comparison of differences in results. *British Journal of Mathematical and Statistical Psychology*, *62*(1), 97–128.
- Sheu, C.-F., & Suzuki, S. (2001). Meta-analysis using linear mixed models. *Behavior Research Methods, Instruments, & Computers*, *33*(2), 102–107.
- Sidik, K., & Jonkman, J. N. (2005). Simple heterogeneity variance estimation for meta-analysis. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *54*(2), 367–384.
- StataCorp. (2017). *Stata statistical software: release 15*. College Station: StataCorp LP.
- Sterne, J. A., Bradburn, M. J., & Egger, M. (2008). *Meta-analysis in Stata™. Systematic Reviews in Health Care: Meta-Analysis in Context*, Second Edition, 347–369.
- Tanner-Smith, E. E., & Grant, S. (2018). Meta-analysis of complex interventions. *Annual Review of Public Health*, *39*, 135–151.
- Team, R. C. (2017). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.
- Therneau, T., Atkinson, B., & Ripley, B. (2017). *rpart: Recursive Partitioning and Regression Trees*. R package version 4.1-11.
- Tibshirani, R., & Knight, K. (1999). Model search by bootstrap “bumping”. *Journal of Computational and Graphical Statistics*, *8*(4), 671–686.
- Tipton, E., Pustejovsky, J. E., & Ahmadi, H. (2018). A history of meta-regression: technical, conceptual, and practical developments between 1974 and 2018. *Research Synthesis Methods*.
- Trujillano, J., Badia, M., Serviá, L., March, J., & Rodriguez-Pozo, A. (2009). Stratification of the severity of critically ill patients with classification trees. *BMC Medical Research Methodology*, *9*(1), 83.
- van Genugten, L., Dusseldorp, E., Webb, T. L., & van Empelen, P. (2016). Which combinations of techniques and modes of delivery in Internet-based interventions effectively change health behavior? A meta-analysis. *Journal of Medical Internet Research*, *18*, 6.
- Viechtbauer, W. (2005). Bias and efficiency of meta-analytic variance estimators in the random-effects model. *Journal of Educational and Behavioral Statistics*, *30*(3), 261–293.
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48.
- Wallace, B. C., Schmid, C. H., Lau, J., & Trikalinos, T. A. (2009). *Meta-analyst: software for meta-analysis of binary, continuous and diagnostic data*. *BMC Medical Research Methodology*, *9*(1), 80.
- Wang, M. C., & Bushman, B. J. (1999). *Integrating Results through meta-analytic review using SAS*. Cary: Software SAS Institute.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.