HHS Public Access

Author manuscript

Eur Neuropsychopharmacol. Author manuscript; available in PMC 2016 October 01.

Published in final edited form as:

Eur Neuropsychopharmacol. 2015 October ; 25(10): 1586–1599. doi:10.1016/j.euroneuro.2015.06.006.

An Analysis of Moderators in the COMBINE Study: Identifying Subgroups of Patients Who Benefit from Acamprosate

Ralitza Gueorguieva, PhD[1], **Ran Wu, MS**[2], **Wan-Min Tsai, PhD**[1], **Patrick G O'Connor, MD, MPH**[3], **Lisa Fucito, PhD**[2], **Heping Zhang, PhD**[1], and **Stephanie S. O'Malley, PhD**[2]

[1]Department of Biostatistics, Yale University School of Public Health and School of Medicine, New Haven, CT 06520, USA

[2]Department of Psychiatry, Yale University School of Medicine, New Haven, CT 06519, USA

[3]Department of Internal Medicine, Yale University School of Medicine, New Haven, CT 06519, USA

Abstract

The goal of the current study was to use tree-based methods to identify moderators of acamprosate effect on abstinence from heavy drinking in COMBINE, the largest study of pharmacotherapy for alcoholism in the United States to date. We used three different tree-based methods for identification of subgroups with enhanced treatment response on acamprosate based on over 100 predictors measured at baseline in COMBINE. No heavy drinking during the last two months of treatment was the considered outcome. All three methods identified consecutive days of abstinence prior to treatment as the most important moderator of treatment effect. Acamprosate was beneficial for participants with shorter abstinence (1 week or less) especially when body mass index was low or normal. In this group, 46% of participants receiving active acamprosate abstained from heavy drinking compared to 23% of those receiving placebo acamprosate. Prior treatment, age, drinking goal and cognitive inefficiency were identified as moderators of acamprosate effects by one of the three methods. In conclusion, acamprosate may be beneficial for participants with shorter abstinence who are not overweight or obese. One hypothesis for this finding is that this subgroup may have greater glutamatergic hyperactivity, a target of acamprosate, and may achieve better drug plasma levels based on their lower BMI. In contrast,

Corresponding author: Ralitza Gueorguieva, 60 College St., New Haven, CT 06520-8034, Tel: 203 285 8155, Fax: 203 785 6912, ralitza.gueorguieva@yale.edu.

Dr. Stephanie O'Malley disclosed the following financial interests: member American Society of Clinical Psychopharmacology workgroup, the Alcohol Clinical Trial Initiative, sponsored byAbott Laboratories, Eli Lilly & Company Lundbeck, Pfizer, and Ethypharma; medication donations, Pfizer, Inc.; Advisory Board, Alkermes and the Hazelden Betty Ford Foundation. All co-authors have no conflicts of interest to disclose.

Drs. O'Malley and Gueorguieva designed the study and wrote the majority of the manuscript. Dr. Tsai, Ms. Wu and Dr. Gueorguieva performed the statistical analyses. Drs. O'Connor and Fucito were involved with literature searches, contributed to variable selection, interpretation of results and drafted parts of the manuscript. Dr. Zhang supervised statistical analyses and was involved in manuscript preparation. All authors contributed to and have approved the final manuscript.

those with extended pretreatment abstinence who have an otherwise good prognosis did not benefit from acamprosate. Further validation of the results in independent data sets is necessary.

Keywords

alcohol dependence; moderator effects; classification and regression trees; subgroups with enhanced treatment effect; clinical trials

Introduction

The primary objective of randomized clinical trials is to assess average treatment effects: that is, how much the treatment effects differ on average across participants within each condition. However, due to between-subject heterogeneity, treatments may work well in one subset of the population and may be less effective in another subset. For such treatments, it is hard to show an average beneficial effect and hence these treatments may be underutilized in a population for which they might provide significant benefit. This is a particularly troublesome issue in clinical trials of treatments for alcohol dependence characterized by high patient heterogeneity and where treatment effects are typically in the small to medium range. To address this issue it has become necessary to explore moderator effects, i.e. to identify specific baseline covariates that stratify the population into subgroups for which treatment has differential effects (Kraemer et al, 2002). However, the usual approach has been to consider baseline predictors one at a time (e.g., Ray and Hutchison, 2007) or to test treatment effects among predefined endophenotypes (e.g., Mann et al, 2009). In COMBINE, the largest clinical trial of treatments for alcoholism to date in the United States (Anton et al. 2006) only individual predictors/moderators of treatment effects (naltrexone, acamprosate, CBI) have been considered (e.g. Anton et al., 2008) or “unsupervised” clustering methods have been applied (Bogenschutz et al., 2008). Since covariates are often related to each other and subpopulations are defined by combinations of predictor variables, it is of limited use to consider only main effects of predictors. Furthermore, an easy interpretation is essential for translating findings from clinical trials into clinical practice.

Tree-based and forest-based methods address the limitations of considering predictors one at a time and are considered “supervised learning” approaches. Classical decision trees (Breiman et al. 1984; Zhang and Singer, 2010) identify combinations of patient characteristics associated with good outcome overall, i.e. they identify which variables interact with one another to produce a certain classification. This is done via recursive partitioning by dividing the study sample recursively into groups that are most homogeneous with respect to the outcome and most distinct from one another. Different versions of the algorithm incorporate different statistical criteria for splitting the sample and determining the optimal size of the tree.

Tree-based methods are appealing alternatives to standard linear model techniques when assumptions of additivity of the effects of explanatory variables, normality and linearity are untenable. Tree-based and forest-based methods are nonparametric computationally intensive algorithms that can be applied to large data sets and are resistant to outliers. They allow consideration of a large pool of predictor variables and can discover predictors that

even experienced investigators may have overlooked (Zhang et al., 2010). These methods are most useful for identification of variable interactions and may be easier to use in clinical settings because they require evaluation of simple decision rules rather than mathematical equations (Zhang and Singer, 2010).

Prior analysis of the COMBINE data using classical tree-based approaches (Gueorguieva et al., 2014) identified longer abstinence, drinking goal of total abstinence and older age as predictive of lower probability of heavy drinking during the last two months of double-blind treatment irrespective of treatment. However, the tree-based methods did not identify interactions involving treatment and thus did not consider moderating effects of the various treatments. Several distinct methods which represent modifications of decision trees have been proposed in recent years (Zhang et al, 2010; Foster et al., 2011, Lipkovich and Dmitrienko, 2014). Each of these methods allows identification of subgroups of participants for whom there are significant differences in effectiveness of treatments and thus could be useful in identifying moderators. In the current study we apply each of these three different methods to identification of moderators of acamprosate effects and evaluate the consistency of the conclusions from these three approaches.

The COMBINE Study evaluated the benefits of combining pharmacotherapy treatment (naltrexone, acamprosate) and behavioral interventions (Medication Management (MM) (Pettinati et al., 2004), Combined Behavioral Intervention (CBI), (Miller, 2004)) in alcohol dependent patients. In the primary analyses of the study, naltrexone (+MM) and CBI (+MM) were associated with improved outcome. However, participants on acamprosate did not have significantly better outcome than participants on placebo (Anton et al., 2006). Despite the absence of an average treatment effect of acamprosate, it is possible that there are subgroups of patients for whom acamprosate is beneficial. In particular, acamprosate is hypothesized to affect negative reinforcement of addictive behavior (Littleton, 1995, Mann et al., 2008) and hence pretreatment commitment to abstinence (Hall et al., 1990) could be an important moderator of treatment response. Consistent with this, acamprosate has been found to be effective among those who were committed to abstinence (Mason et al., 2006). There is also evidence that acamprosate may be helpful for alleviating withdrawal symptoms during initial alcohol abstinence such as sleep disturbance (Perney et al., 2012; Staner et al., 2006). In previous analyses by our group, acamprosate appeared to “rescue” early non-compliers to CBI (Gueorguieva et al., 2014) and baseline trajectories of drinking moderated acamprosate response (Gueorguieva et al., 2011) such that acamprosate was counter-therapeutic for daily drinkers who achieved a longer period of abstinence prior to treatment.

Research findings on potential moderators of response in COMBINE are accruing but most of the results are focused on the effects of naltrexone. For example, naltrexone response in COMBINE has been shown to be moderated by smoking status (Fucito et al., 2012), alcoholism typology (Bogenschutz et al., 2008), craving (Subbaraman et al., 2013), social network (Worley et al., 2015) and OPRM1 genotype (Anton et al., 2008), but not by family history of alcoholism (Capone et al., 2011). Studies evaluating moderating effects of acamprosate are fewer. Baseline trajectories of drinking have been shown to moderate acamprosate response in COMBINE (Gueorguieva et al., 2011). In a different study, the minor allele of GRIN2B rs2058878 was found to be associated with shorter abstinence in

acamprosate treated participants (Karpyak et al., 2014) but there was no control sample of participants on placebo or comparator treatment.

The goal of the current study is to use tree-based approaches to identify important moderators of acamprosate effect in order to inform providers about the likelihood of success with this particular treatment based on the characteristics of the individual patient. The COMBINE baseline assessments were selected to capture demographic characteristics and domains that were thought to be predictors or moderators of treatment efficacy based on prior findings, hypothesized mechanisms, and theory (The COMBINE Study Research Group, 2003). Thus this data set provides an opportunity to consider a large number of theoretically-derived potentially important predictors and use a data-driven approach to identify potential moderators of acamprosate effect.

2. Experimental Procedures

2.1. The study sample

In COMBINE, eight groups received medical management (MM) with 16 weeks of naltrexone (100 mg/day) or acamprosate (3 g/day), both, and/or both placebos, with or without CBI. Our analysis focused on participants who had any drinking data during treatment (N=1220). A small percentage of participants had received inpatient treatment in the 30 days prior to enrollment (7.7%) and the majority was recruited from the community.

2.2. Drinking outcome

The outcome measure was no heavy drinking during the last eight weeks of double-blind treatment. This measure is recommended for clinical trials because it is associated with reduced risk of alcohol related consequences while allowing for improvements in drinking short of abstinence (Falk et al., 2010). It is also convenient to use with the classification approach as it provides an easily ascertained outcome and an easily interpreted decision. Missing heavy drinking data were coded as heavy drinking.

2.3. Predictors

We considered over one hundred baseline predictors in COMBINE that had less than 15% missing values. Categorical predictors with missing values had an additional missing category created. Continuous predictors had missing values imputed using PROC MI in SAS. Ordinal and continuous predictors with more than 5 levels were categorized in 4–5 categories in order to avoid over-representation as splitters and to improve interpretability. For example, all laboratory measures were categories as 1=below the lower limit of the normal range (if applicable), 2=lower third of the normal range, 3=middle third of the normal range, 4=upper third of the normal range, 5=above the upper limit of the normal range. Predictors are shown by domain in Table 1 and described in detail in Appendices 1 and 2.

2.4. Trees and tree construction methods

Each tree consists of a root node, a number of internal nodes (denoted by ovals in the figures, e.g. Figure 1) and a number of terminal nodes (denoted by rectangles). The entire

sample is represented in the root node of the tree. Then the sample is split recursively according to different criteria into two daughter nodes as described below in each of the three methods so that the participants on one treatment in each daughter node have as different outcome as possible compared to the participants on the other treatment in the same daughter node. All predictor variables and possible levels of these predictor variables at which the sample can be split are considered and at each stage the best possible splitting variable is selected. Splitting variables and cutoffs are shown underneath each node of the tree. Sample sizes and relevant measures of difference in outcome proportions for participants on different treatments in the nodes are provided in the nodes. The methods are now explained in more detail.

2.4.1. Zhang et al. (2010) method—The algorithm proceeds in two steps: tree growing and tree pruning.

Tree growing: Each node is split into 2 daughter nodes based on maximizing the difference between the proportions of participants with no heavy drinking during the last 8 weeks of treatment on acamprosate and on placebo in the parent and daughter nodes (Zhang et al., 2010). All predictors and all possible values at which the predictors can be split are considered. Splitting proceeded recursively until no further splits were possible. Restriction of at least 100 participants in each node (50 on each treatment) is imposed in order to avoid splits based on small samples that might be difficult to validate.

Tree pruning: Sibling nodes that favor the same treatment are pruned from the bottom-up using an algorithm implemented in R (R core team, 2013). In the final tree terminal nodes that are associated with better outcome on acamprosate than on placebo are colored in orange while terminal nodes that are associated with worse outcome on acamprosate than on placebo are colored in blue.

2.4.2. Foster et al. (2011) method—Prior to building a tree, a modified outcome variable (the estimated causal effect of treatment) is generated for each subject. The outcome variable is the difference in estimated probabilities of no heavy drinking during the last 8 weeks of treatment on acamprosate and on placebo and it is generated as described below. For each subject we observe directly only one of these two outcomes depending on whether they are randomized to acamprosate or to placebo. We use random forests of 1000 trees each to estimate the probabilities of the outcome on the actually received treatment and on the alternative (counterfactual) treatment for each subject using the R code provided by Foster et al. (2011). Then we calculate the difference of these two probabilities for each subject that represents the modified outcome variable for each individual. Then a classical regression tree is built using the *rtree* function in R. We also calculate and report variable importance scores based on the random forests constructed for the modified outcome variable which indicate which variables occur most often as moderators of acamprosate in random forests. Variable importance scores reflect the impact of removing a variable from the set of covariates on the predictive performance of the tree. The higher the score, the more information is lost when the variable is removed and the greater the change in predictive ability of the tree. We report the variables with top 10 variable importance scores

according to percent increase in mean squared error and percent increase in node impurity after random permutation. These variables represent the strongest moderators that can be identified in the data set.

2.4.3. SIDESscreen method (Lipkovich and Dmitrienko, 2014)—The goal of this method is to identify subgroups of participants with enhanced treatment effect. Rather than building complete trees like the methods of Foster et al. (2011) and Zhang et al. (2010), it focuses on the maximal treatment effect only in subgroups of participants in terminal nodes of trees of up to pre-specified number of layers. There are three steps of the algorithm. In the first step, the subgroups with enhanced treatment effect are identified recursively subject to the constraints of maximum number of covariates defining subgroups and minimal node size. We used the maximal treatment effect splitting criterion (which is based on the one-sided p-value for the treatment effect in the subgroup), the recommended cutoff of three for the number of covariates defining a subgroup and the default minimum node size (30). In the second step, only subgroups in which the treatment effect is below a pre-specified threshold value are retained which is equivalent to the pruning step in the full tree approaches. In the third step, a multiplicity adjustment is applied to correct the p-values for the treatment effect within the identified subgroups for the extensive data-mining inherent in the algorithm. The default settings in the SIDESscreen Excel Macro for the second two steps provided by Lipkovich and Dmitrienko (2014) were used in the current analysis.

3. Results

The final tree built using the approach of Zhang et al. (2010) is shown in Figure 1. This method splits the nodes in the tree based on the difference of probabilities of the outcome in the two treatment groups. Terminal nodes in the tree are color-coded so orange nodes correspond to better outcome on acamprosate and blue nodes correspond to worse outcome on acamprosate compared to placebo. The root node (node 1) of this tree shows the entire sample of 1220 participants, of which 604 received active acamprosate and 616 received placebo acamprosate. Among those who received active acamprosate, 44% did not have any heavy drinking days in the last 8 weeks of treatment and among those who received placebo acamprosate, 40% did not have any drinking days in the last 8 weeks of treatment. This difference was not statistically significant (p=0.16).

Among all considered predictor variables, consecutive abstinence prior to randomization was the best moderator variable. While those with more than 1 week of abstinence prior to treatment had overall better outcomes than those with shorter abstinence, acamprosate showed benefit among the group with less pretreatment abstinence. Specifically, those with up to 1 week of abstinence prior to treatment had a better outcome on acamprosate (39% had no heavy drinking days) compared to placebo (30% had no heavy drinking days, node 2). In addition, the advantage of acamprosate over placebo for participants with shorter abstinence was most pronounced for participants with low or normal Body Mass Index (BMI, node 4). In this group, twice as many patients on acamprosate (46%) compared to patients on placebo (23%) abstained from heavy drinking in the last 8 weeks of treatment (node 4). Acamprosate also appeared beneficial for 1) participants with less than 1 week of abstinence who had above normal BMI and who acknowledged some mood-induced cognitive inefficiency on

the Profile of Mood States Confusion subscale (node 9, 37% with good outcome on acamprosate vs. 24% on placebo). Within those who had shorter abstinence, however, there was also a subgroup composed of patients who were overweight or obese and who did not report confusion for whom acamprosate was associated with poorer outcome (node 8; 34% acamprosate vs. 48% placebo).

Among those with pretreatment abstinence of more than 1 week, participants had poorer outcome on acamprosate (52%) than on placebo (60%) (node 3). Furthermore, this negative effect of acamprosate among those with greater pretreatment abstinence was larger in participants who also had GGT above the normal reference range (node 7, 49% acamprosate with good outcome vs. 67% placebo). However, a benefit of acamprosate was found for the subgroup of individuals with longer abstinence and GGT within the normal range (node 6, 57% with good outcome on acamprosate vs. 50% on placebo).

The final regression tree built using the approach of Foster et al. (2011) is shown in Figure 2. The outcome for this analysis was the difference in estimated probabilities of no heavy drinking during the last 8 weeks of treatment on acamprosate and on placebo. Each node of the tree in Figure 2 shows the total number of participants in this node and the average difference in the estimated probabilities of the outcome on active and on placebo for the participants in the node. Positive numbers indicate that the node favors acamprosate, negative numbers indicate that the node favors placebo. The same color-coding scheme is used as in the approach of Zhang et al. (2010). In the entire sample of 1220 participants (node 1), the difference in estimated probabilities of good outcome on acamprosate vs. placebo was 0.03 (or 3%). The top splitting variable in this approach also appeared to be consecutive abstinence prior to treatment. Among those with abstinence of 1 week or less (n=815), the estimated probability of good outcome on acamprosate was on average 6% higher than on placebo (node 2). The difference was even larger for older participants (>45 years old) within this group (node 5, n=398, 8%) than for younger participants (node 4, n=417, 4%). On the other hand, for participants with more than one week of abstinence, placebo was associated with higher probability of good outcome than acamprosate, especially if the drinking goal was controlled drinking (node 6, n=126, average difference in probabilities 6%) rather than other goals (node 7, n=279, diff = 2%).

The three splitters in this tree, specifically duration of baseline abstinence, age and drinking goal, were also the three most important predictors of the estimated causal effect of treatment identified by the Foster et al. (2011) variable importance approach (Table 2). The other top moderators of acamprosate effect were the Alcohol Abstinence Self-Efficacy (AASE) confidence total score and several subscale scores (social, negative affect, withdrawal/urge), uric acid (which has been shown to increase linearly with increased levels of alcohol consumption (Oliveira et al, 2010), perhaps especially so with beer and liquor consumptions as opposed to wine consumption (Choi and Curhan, 2004)), measures of heavy drinking (heavy drinking days per week, SCID alcohol dependence symptoms, BAC peak) and any drinking (days abstinent per week).

The SIDESscreen approach (Lipkovich and Dmitrienko, 2014) identified two subgroups with most pronounced benefit of acamprosate compared to placebo (Table 3). Participants

who were abstinent for four or fewer days prior to treatment, who had prior treatment and who were not obese (n=168) had poor outcome on placebo (percent with no heavy drinking days of only 13%) but much better outcome on acamprosate (42%). The second subgroup was slightly different and included participants who had up to one week of abstinence, had prior treatment and were not overweight or obese (n=137). These participants also had much better outcome on acamprosate than on placebo (52% vs. 19% with no heavy drinking during the last 8 weeks of treatment). The unadjusted p-values for acamprosate effects in both of these subgroups were <.0001. After conservative adjustment for multiple testing the comparison in the first group was still statistically significant at experiment-wise level of 0.05 (p=0.04) while the second comparison was a trend (p=0.06). No other subgroups had statistically significant or close to significant treatment effects after correction for multiple testing.

4. Discussion

All three tree-based approaches identified pre-treatment abstinence as an important moderator of acamprosate effect. Two of the approaches also showed that among those with shorter abstinence acamprosate had significant benefit for those with low or normal BMI. Better response when BMI is low could be related to better drug levels as individuals with lower BMI may be receiving a greater dose on a mg/kg basis. Consistent with this hypothesis, earlier European studies observed dose-dependent effects of acamprosate on abstinence and positive results were obtained where dosing was adjusted by body weight (see Mason et al., 2001 for a review). A recent Japanese study also found significant effect of acamprosate on abstinence (Higuchi, 2015). Participants in the European studies and in the Japanese study were also on average lighter than participants in COMBINE. For example, the average reported weight was 69.3kg in Paille et al. (1995), approximately 60kg in Higuchi (2015), 73.1kg in Sass et al. (1996) while it was 81.5kg in COMBINE. Moreover, concurrent administration with food also lowers acamprosate absorption (Saivin et al., 1998); an effect that might be important in overweight and obese individuals. Differences in body weight and resulting drug plasma levels might explain the absence of overall acamprosate effects in COMBINE, however COMBINE used a higher dose of acamprosate because of the evidence of dose-dependent effects in the earlier European studies. In our analysis body weight was particularly important as a moderator of acamprosate effect only for participants with shorter abstinence who were expected to have worse outcome. Of note, those who received placebo in this subgroup had particularly poor outcomes, so it is possible that body weight influenced acamprosate response through a mechanism other than drug levels. Nonetheless, our data suggest that acamprosate may have some benefit for those who are not overweight and achieve less abstinence when abstinence is required prior to starting treatment.

Among those with greater abstinence prior to treatment who had a better prognosis overall, acamprosate was associated with poorer outcome compared to placebo. In this subgroup, adverse events or other effects of active drug may have undermined the expected improvement of these good prognosis patients. These results are consistent with previous exploratory analyses of COMBINE that reported negative effects of acamprosate in daily drinkers at baseline who were able to maintain longer abstinence prior to treatment

(Gueorguieva et al, 2011). Possible implications of this result for the design of future efficacy studies is to set upper limits on pre-treatment abstinence in order to minimize the placebo response rate and avoid exposure to possible adverse effects of medications to patients with an otherwise good prognosis.

Prior treatment, cognitive inefficiency (POMS confusion scores), age and drinking goal were each identified as potential moderators of acamprosate effects in one of the approaches we used. According to the SIDESscreen method (which is the only method with built-in adjustment for multiple testing) the effect of acamprosate for participants with shorter abstinence and low or normal BMI was most pronounced if they also had prior treatment. Prior treatment, combined with shorter duration of abstinence, may be selecting a subgroup with abstinence induced glutamatergic activity that might benefit from acamprosate especially if adequate levels are achieved. While this finding requires further validation, if confirmed, it suggests a specific group with very poor outcome on placebo who might benefit from acamprosate. For participants with shorter abstinence but who had above normal BMI, acamprosate appeared to be beneficial for those who acknowledged some cognitive inefficiency on the POMS confusion scale. This finding also needs further confirmation before implications can be discussed.

We previously observed that among those with shorter abstinence prior to treatment, older participants compared to younger participants had better outcome regardless of treatment (Gueorguieva et al, 2014). The current study suggests that there is potential benefit of acamprosate in the same subgroup (patients older than 45 years with less abstinence). However, younger participants with less abstinence benefited from naltrexone while acamprosate was not associated with a significant advantage. Thus naltrexone and/or more intensive interventions may be more appropriate for this subgroup (younger patients with less abstinence).

We also previously found that goal of total abstinence was associated with better outcome regardless of treatment (Gueorguieva et al, 2014). In the current study drinking goal appeared as a potential moderator variable only in the Foster et al. (2011) approach and only among those with longer pre-treatment abstinence. A controlled drinking goal among those with longer abstinence was associated with better outcome on placebo compared to acamprosate. Thus acamprosate may be counterproductive for good prognosis patients with a controlled drinking goal and alternative treatments may be preferred for such individuals.

Despite evidence in the literature that pre-treatment commitment to abstinence could be an important moderator of treatment effect (Hall et al., 1990), we did not find evidence that this is a moderator of acamprosate effect. Only actual abstinence prior to treatment was related to acamprosate effectiveness. Likewise, sleep prior to treatment was not a significant acamprosate treatment modifier despite prior evidence that acamprosate improves sleep disturbance common during early alcohol abstinence (Perney et al., 2012; Staner et al., 2006). Genetic samples were available only a subset of the COMBINE population hence we were unable to test hypotheses related to potential moderating effects of genotype on acamprosate effects.

While the three approaches identified potential subgroups with enhanced or negative treatment effect, the effect sizes for acamprosate effects were either relatively small even in those carefully selected groups (e.g. the Foster's approach identified difference as large as 8% in the estimated probabilities of NHDD) or the groups themselves were a small proportion of the entire sample (e.g. in the SIDESscreen approach the subgroups are only 10–15% of the entire sample). For comparison we calculated the average naltrexone and CBI effects in the entire sample and in some subsamples. The differences in probabilities of no heavy drinking in the entire sample were of similar magnitude to the largest difference observed for acamprosate: 54% vs. 46% for NHDD on naltrexone compared to placebo naltrexone, and 53% vs. 43% on CBI compared to no CBI placebo. Within certain subgroups, the naltrexone effect was even larger. For example, in the subgroup of younger participants with shorter abstinence 30% had NHDD on naltrexone compared to 19% on placebo naltrexone leading to a difference of probabilities of 0.11.

A potential caveat of the statistical approach used is that tree-based and forest-based methods are prone to idiosyncratic results that may fail to cross-validate in the same sample or replicate in other samples. By considering three different conceptual approaches for moderator effects and focusing on the congruent results, we minimize the probability of chance findings due to a particular method. However, without replication on another sample, we cannot claim generalizability of our results to other samples or target populations. The COMBINE sample is not necessarily representative of the patient populations treated for alcohol dependence. Programs may vary in the requirements for abstinence, and patients may have comorbidities, such as drug dependence, that were exclusionary criteria in COMBINE.

Although external validation of our results is necessary, this study demonstrates the usefulness of the tree-based approach for identification of subsamples with differential treatment effects. Tree-based methods have advantages over classical statistical methods because they rely on fewer assumptions and are useful for identification of interactions. Although logistic regression can also be used to test interactions, usually the number of predictors and the order of the considered interactions is limited (only up to two-way or three-way) thus they are difficult to use for systematic exploration of interactive effects. In contrast, trees automatically present in the form of simple decision rules and can be easier to adapt for use in clinical settings.

Our focus in the current study was on the simple binary outcome of no heavy drinking after a grace period. Trees for binary outcomes are easy to interpret and to incorporate in clinical practice. However, other types of outcomes such as percent heavy drinking days or time to relapse to drinking are also of interest and further research is necessary to identify moderators of treatment effects for such outcomes. Of the methods presented in this paper only the SIDESscreen approach can be directly applied to continuous outcomes. There are other methods that have been recently developed that can be used for continuous (Su et al., 2009; Dusseldorp et al., 2010; and Dusseldorp and Mechelen, 2014) and censored continuous outcomes (Negassa et al., 2005 and Loh et al., 2014).

The constructed trees in this study can be directly used by clinicians to identify for a particular patient the terminal node to which this subject belongs and provide an immediate estimate of the expected treatment outcome on the alternative treatment and thus guide clinical decision making. Pending replication, our results using these methods suggest that patients who are not overweight or obese and who achieve less abstinence (1 week or less) prior to treatment may benefit from acamprosate, whereas those who are able to achieve more than 1 week of abstinence, a predictor of good outcome overall, do not.

Acknowledgments

We would like to acknowledge Dr. Eugenia Buta for assistance with statistical analysis.

The project described was supported by Grants R01AA017173, K05 AA014715, and K23 AA020000 from the National Institute on Alcohol Abuse and Alcoholism. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute on Alcohol Abuse and Alcoholism or the National Institutes of Health. The sponsor *had no role in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the paper for publication.*

Appendix 1: Baseline predictors in COMBINE

We considered the three **treatments** (naltrexone, acamprosate, CBI) and **demographic** variables (age, gender, race, marital status, years of education, employment, family income).

Pre-treatment alcohol consumption was assessed on the Form-90 (Miller and Del Boca, 1994; Tonigan et al., 1997). The Form-90 was developed for Project MATCH and is a standardized 90-day retrospective interview about daily alcohol consumption. It uses a combination of Timeline Follow-Back (TLFB, Sobell and Sobell, 1995) and grid averaging assessment strategies to obtain accurate assessments of alcohol consumption. We used percent heavy days, percent abstinent days, consecutive days of abstinence and peak BAC level (from Form-90, averaged over the two heaviest drinking episodes in the 90 days prior to intake).

Alcohol severity was assessed by SCID symptom count (Spitzer et al., 1992), total CIWA score (the Clinical Withdrawal Assessment Scale- AR, Sullivan et al., 1989), the total score on the Drinker Inventory of Consequences (DrInC; Miller et al., 1995), the Alcohol Dependence Scale (ADS; Skinner and Allen, 1982) and age of onset. The Alcohol Dependence Scale is a 25-item scale that assessed dependence symptoms, including withdrawal and increased alcohol tolerance in the 12 months before assessment. The Drinker Inventory of Consequences assessed negative consequences of alcohol abuse in the 90 days before treatment.

Prior Alcohol Treatment was assessed based on single, dichotomous items whether or not the subjects had ever participated in any other alcohol treatment (Treatment Experiences and Expectations (TEE), Donovan D., unpublished instrument), whether they have been previously detoxified and whether they had ever attended Alcoholics Anonymous (Baseline Form 90, Miller, 1996).

Prior to treatment, a question from the Thoughts About Abstinence Scale (Hall et al., 1990) assessed **drinking goal** as part of the Treatment Experiences and Expectancies

questionnaire. The item read, “We would like to know what GOAL you have chosen for yourself about using alcohol at this time.” Participants were categorized into 3 groups: (1) *controlled drinking* (CD), assessed by positive responses to any of the following items, “I want to use alcohol in a controlled manner – to be in control of how often I use and how much I use” and “I don’t want using alcohol to be a habit for me anymore, but I would like to occasionally use alcohol when I really have an urge”; (2) *total abstinence goal* (TA), with a positive response to the following item, “I want to quit using alcohol once and for all, to be totally abstinent, and never use alcohol ever again for the rest of my life”; (3) *conditional abstinence* (CA), assessed by the following items, “I want to be totally abstinent from all alcohol use for a period of time, after which I will make a new decision about whether or not I will use alcohol again in any way” and “I want to quit using alcohol once and for all, even though I realize I may slip up and use alcohol once in a while.” The remaining questions were combined into an “Other” category.

Family history of known alcohol dependence and smoking were considered based on first degree relatives: paternal history, maternal history, or history in two or more first degree relatives (yes, and no otherwise).

Craving was assessed using the Obsessive Compulsive Drinking Scale (OCDS; Anton et al., 1995). The OCDS is a self-administered, 14-item scale with items to assess the obsessive and compulsive characteristics related to thoughts about drinking and the ability to resist drinking-related thoughts and urges. Item scores were combined to create a total scale score.

Cigarette smoking and **cannabis use** were dichotomous variables based on current use reported on the Form-90 (Miller, 1996; Tonigan et al., 1997) at baseline. Participants were dichotomized into smokers and nonsmokers for cigarette use and for cannabis use.

Physical exam measurements included BMI, pulse and blood pressure. Thirty three **laboratory** tests were also considered (e.g., liver and kidney function tests). For a complete listing see Appendix 2.

The **Alcohol Abstinence Self-Efficacy Scale** (AASE; DiClemente et al., 1994) assessed participants’ self-efficacy to abstain from drinking in situations that correspond to typical drinking cues. The types of situations include negative affect, social, physical, withdrawal/urges. The total confidence and temptation score and the four subscales for temptation and confidence were used for these analyses.

The University of Rhode Island Change Assessment (URICA; DiClemente & Hughes, 1990) assessed participants’ stages of **readiness to change**. The URICA is a 28-item scale that assesses the four stages of change: pre-contemplation, contemplation, action, and maintenance. It yields four subscales and an overall readiness score.

Quality of Life was assessed on four domains (physical, psychological, social relationships and environment) using the WHO Quality of Life Scale (Szabo, 1996). **General health** was assessed with a single item from the Short-Form-12, Version 2 (Ware et al., 2002). The item assessed perceived physical health and was scored on a 1 (Poor) to 5 (Excellent) Likert

scale. This item was selected because self-ratings of physical health are related to mortality, even when modeled with other health indices (Idler and Benyamini, 1997; Jylha, 2009).

Profile of Mood States (McNair et al., 1981) was used to assess current mood and included six subscales: tension, depression, anger, vigor, fatigue, and confusion.

The **Perceived Stress** Scale - Short Form (PSS; Cohen et al., 1983; Cohen and Williamson, 1988) assessed the degree to which participants perceived their life situations to be stressful. The PSS is a 4-item instrument scored on a 0 (never) to 4 (very often) Likert scale; items assess the degree to which participants perceive their lives to be controllable and predictable. The total score was created by summing the items.

Sleep problem (yes, no, missing) was defined as any symptom of insomnia, sleep disturbance, problems of sleep, and decreased sleep based on SAFTEE (Johnson et al., 2005) general inquiry at week 0.

The **Important People** Interview (IPI; Longabaugh and Zywiak, 2002) assessed the composition of participants' social networks. The IPI is a structured interview that includes questions about participants' perceptions of people who are most important to them and with whom they have had contact in the previous 4 months. Each participant can list up to 10 network members, specifying various aspects of each relationship including the nature of the relationship, level of supportiveness of drinking in the relationship, drinking status, and frequency of network member drinking. Total number of in-network daily drinkers (Longabaugh et al., 2010) was used in these analyses.

Legal problems were assessed by any history of arrest dichotomized into yes or no (Form-90, Miller, 1996).

Domain	Predictor	Categories	Source, including reference for the instrument as necessary
Demographics	Age	0–24, 25–34, 35–44, 45–54, 55–64, ≥65	Baseline demographics
	Gender	male, female	
	Race	White, Black, Hispanic, Other	
	Marital status	married, not married	
	Years of education	<12, 12, 13–16, >17	
	Current work	employed, unemployed, homemaker/student/retired, disabled/other	
	Family income	$0-$15,000, $15001–30000, $30001–50000, $50001-75000, $75001–100000, more than $100000	
Alcohol Consumption	% heavy drinking days	0=almost no heavy drinking, daily	Derived from baseline Time-line Follow-Back (TLFB, Sobell & Sobell, 1995)

Domain	Predictor	Categories	Source, including reference for the instrument as necessary
	% abstinent days	0=almost no drinking, daily	
	Consecutive days of abstinence prior to randomization	0–4, 5–7, 8–14, 15–21, >22	
	Peak BAC	<.06,≥.06 to <.11, .11 to<.21, .21 to<.3, .3 to<.4, ≥.5	Form-90 (Miller and Del Boca, 1994; Tonigan et al., 1997)
Alcohol Severity	SCID symptom count	count 3–7	Derived from SCID-IV Module E (Spitzer et al.,1992)
	DRINC Total	men: 0–38, 39–59, >59; women: 0–35, 36–52, >52	Drinker Inventory of Consequences (DrlnC; Miller et al., 1995)
	CIWA	0=0, 0–7, 8–14, >14	CIWA (The Clinical Withdrawal Assessment Scale-AR, Sullivan et al., 1989)
	ADS Alcohol Dependence Score	0–13, 14–21, 22–30, 31–46; high=non dependent	Alcohol Dependence Scale (ADS; Skinner and Allen, 1982)
	Age of onset	<25, 25–44, ≥45	Structured Clinical Interview and Diagnosis (SCID-IV Module E, Spitzer et al., 1992)
Prior Alcohol Treatment	Detoxification	yes, no	Baseline Form 90 (Miller, 1996)
	AA attendance	yes, no	
	Any treatment	yes, no	Treatment Experiences and Expectations (TEE, Donovan D., unpublished instrument)
Drinking Goal	Drinking goal	complete abstinence, conditional abstinence, controlled drinking, other	Derived from Treatment Experiences and Expectations (TEE, Donovan D., unpublished instrument, Bujarski et al., 2013; Hall et al., 1990)
Family History	Alcohol	yes, no	Family History
	Smoking	yes, no	
Craving	OCDS total score	≤10, 11–20, 21–30, ≥31	Obsessive Compulsive Drinking Scale (OCDS; Anton et al., 1995)
Smoking	Current smoker	yes, no	Baseline Form 90 (Miller, 1996)
Drug use	Cannabis use	yes, no	

Domain	Predictor	Categories	Source, including reference for the instrument as necessary
Physical Exam	BMI	<18.5, 18.5–24.9, 25–29.9, ≥30	Physical Exam
	Pulse rate per minute	<60, 60–100, ≥102	
	Blood pressure (sitting) Systolic (mmHg)	<120, 120–139, 140–159, ≥161	
	Blood pressure (sitting) Diastolic (mmHg)	<80, 80–89, 90–99, 100–120, ≥122	
Laboratory Analysis	AST(SGOT, IU/L)	<0, 0–11.66, 11.67–23.32, 23.33–35, >35	Baseline Lab
	ALT(SGPT, IU/L)	<0, 0–11.66, 11.67–23.32, 23.33–35, >35	
	GGT(IU/L)	<0, 0–9.9, 10–19.9, 20–30, >30	
	Total Bilirubin(mg/dL)	<0.3, 0.3–0.59, 0.6–0.89, 0.9–1.2, >1.2	
	Magnesium (mg/dL)	<1.5, 1.5–1.79, 1.8–2.09, 2.10–2.4, >2.4	
	Sodium (mEq/L)	<136/ 136–138, 139–141, 142–145, >145	
	Calcium(mg/dL)	<9, 9–9.49, 9.5–9.99, 10–10.5, >10.5	
	Potassium(mEq/L)	<3.5, 3.5–3.9, 4–4.49, 4.5–5, >5	
	Phosphorus(mg/dL)	<3, 3–3.49, 3.5–3.99, 4–4.5, >4.5	
	Bicarbonate(mEq/L)	<23, 23–24.66, 24.67–26.32, 26.33–28, >28	
	Creatinine(mg/dL)	<0.7, 0.7–0.89,0.9–1.09, 1.1–1.3, >1.3	
	BUN(mg/dL)	<8, 8–11, 12–15, 16–20, >20	
	Glucose(mg/dL)	<70, 70–79, 80–89, 90–100, >100	
	Uric Acid(mg/dL)	<2.5, 2.5–4.32,4.33–6.16, 6.17–8, >8	
	Alkaline Phosphatase(IU/L)	<36, 36–54.66, 54.67–73.32, 73.33–92, >92	
	Lactate Dehydrogenase(IU/L)	<60, 60–73.32, 73.33–86.66, 86.67–100, >100	
	Total Protein(g/dL)	<6, 6–6.59, 6.6–7.19, 7.2–7.8, >7.8	
	Albumin(g/dL)	<3.5, 3.5–4.16, 4.17–4.82, 4.83–5.5, >5.5	
	Hemoglobin(g/dL)	Male: <14, 14–15, 15–16, 16–17, >17; Female: <12, 12–12.33, 12.33–14.66, 14.67–16, >16	
	Hematocrit(percent)	Male: <41, 41–44.33, 44.33–47.66, 47.67–51, >51;	

Domain	Predictor	Categories	Source, including reference for the instrument as necessary
		Female: <36, 36–39.66, 39.67–43.33, 43.33–47, >47	
	RBC(x10^6/uL)	<4.2, 4.2–4.76, 4.77–5.32, 5.33–5.9, >5.9	
	WBC(x10^3/uL)	<4, 4–5.99, 6–7.99, 8–10, >10	
	Platelet Count(x10^3/uL)	<150, 150–216.66, 216.67–283.32, 283.33–350, >350	
	MCV(fL)	<80, 80–86.66, 86.67–93.32, 93.33–100, >100	
	MCH(pg/cell)	<28, 28–29.32, 29.33–30.66, 30.67–32, >32	
Alcohol Abstinence Self-Efficacy	AA1 Total Confidence Score	not at all, not very, moderately, very, extremely	Alcohol Abstinence Self-Efficacy Scale (AASE; DiClemente et al., 1994)
	AA1 Confidence: Negative Affect	not at all, not very, moderately, very, extremely	
	AA1 Confidence: Social	not at all, not very, moderately, very, extremely	
	AA1 Confidence: Physical	not at all, not very, moderately, very, extremely	
	AA1 Confidence: Withdrawal/Urge	not at all, not very, moderately, very, extremely	
	AA2 Total Temptation Score	not at all, not very, moderately, very, extremely	
	AA2 Temptation: Negative Affect	not at all, not very, moderately, very, extremely	
	AA2 Temptation: Social	not at all, not very, moderately, very, extremely	
	AA2 Temptation: Physical	not at all, not very, moderately, very, extremely	
	AA2 Temptation: Withdrawal/Urge	not at all, not very, moderately, very, extremely	
WHO Quality of Life	WHO Physical Health Domain	Higher is better in quality of life, mean score 1–5.	The World Health Organization Quality of Life assessment (Szabo, 1996)
	WHO Psychological Domain	Higher is better in quality of life, mean score 1–5.	
	WHO Social Relationships Domain	Higher is better in quality of life, mean score 1–5.	
	WHO Environment Domain	Higher is better in quality of life, mean score 1–5.	
University of Rhode Island Change Assessment Scale (URICA)	URA Overall Readiness Score	<9, 9–11, 12–14, >14	University of Rhode Island Change Assessment (URICA; DiClemente & Hughes, 1990)

Domain	Predictor	Categories	Source, including reference for the instrument as necessary
General health	In general, would you say your health is?	excellent, very good, good, fair, poor	SFA scale (SF-12, Ware et al. 2002)
Perceived Stress	PSS Perceived Stress Score	never, almost never, sometimes, fairly often, very often	Perceived Stress Scale - Short Form (PSS; Cohen et al., 1983; Cohen and Williamson, 1988)
Sleep problems	Any sleep problems	yes, no, missing	Systematic Assessment for Treatment Emergent Events (SAFTEE) General Inquiry (Johnson et al, 2005)
Important People	Important persons	0=missing and 0, 1, 2, 3, ≥4	Important People Interview (IPI; Longabaugh and Zywiak, 2002)
Legal Problems	Legal problems	yes, no	Baseline Form 90 (Miller, 1996)
Profile of Mood States	POM Tension Subscale	not at all, a little, moderately, quite a bit, extremely	Profile of mood states (POMS, McNair et al., 1981)
	POM Depression Subscale	not at all, a little, moderately, quite a bit, extremely	
	POM Anger Subscale	not at all, a little, moderately, quite a bit, extremely	
	POM Vigor Subscale	not at all, a little, moderately, quite a bit, extremely	
	POM Fatigue Subscale	not at all, a little, moderately, quite a bit, extremely	
	POM Confusion Subscale	not at all, a little, moderately, quite a bit, extremely	
Treatment Condition	Acamprosate	placebo, acamprosate (3gm)	COMBINE treatment assignments
	Naltrexone	placebo, naltrexone (100mg)	
	COMBINE Behavioral Intervention	no CBI,CBI	

References

Anton RF, Moak DH, Latham P. The Obsessive Compulsive Drinking Scale: A self-rated instrument for the quantification of thoughts about alcohol and drinking behavior. Alcohol Clin Exp Res. 1995; 19(1):92–99. [PubMed: 7771669]

Bujarski S, O'Malley SS, Lunny K, Ray LA. The effects of drinking goal on treatment outcome for alcoholism. Journal of Consulting and Clinical Psychology. 2013; 81(1):13–22. [PubMed: 23231573]

Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Social Behavior. 1983; 24:385–396.

Cohen, S.; Williamson, GM. Perceived stress in a probability sample of the United States. In: Spacapan, S.; Oskamp, S., editors. The social psychology of health. Newbury Park, CA: Sage; 1988. p. 31-67.

DiClemente CC, Carbonari JP, Montgomery RPG, Hughes SO. The Alcohol Abstinence Self Efficacy Scale. Journal of Studies on Alcohol. 1994; 55:141–148. [PubMed: 8189734]

DiClemente CC, Hughes SO. Stages of change profiles in outpatient alcoholism treatment. Journal of Substance Abuse. 1990; 2:217–235. [PubMed: 2136111]

Hall SM, Havassy BE, Wasserman DA. Commitment to abstinence and acute stress in relapse to alcohol, opiates, and nicotine. Journal of Consulting and Clinical Psychology. 1990; 58:175–181. [PubMed: 2335634]

Idler EL, Benyamini Y. Self-rated health and mortality: a review of twenty-seven community studies. Journal of Health and Social Behavior. 1997; 38(1):21–37. [PubMed: 9097506]

Johnson B, Ait-Daoud N, Roache J. The COMBINE SAFTEE: A structured instrument for collecting adverse events adapted for clinical studies in the alcoholism field. Journal of Studies on Alcohol. 2005 Supplement 15.

Jylha. What is self-rated health and why does it predict mortality? Towards a unified conceptual model. Social Sciences and Medicine. 2009; 69(3):307–316.

Longabaugh, R.; Zywiak, W. Project COMBINE: A manual for the administration for the Important People Instrument. Providence, RI: Adapted for use of Project COMBINE, Center for Alcohol and Addiction Studies, Brown University; 2002.

Longabaugh R, Wirtz PW, Zywiak WH, O'Malley SS. Network support as a prognostic indicator of drinking outcomes: the COMBINE study. J Stud Alcohol Drugs. 2010; 71(6):837–846. [PubMed: 20946740]

McNair, DM.; Lorr, M.; Droppleman, LF. Profile of Mood States. San Diego, CA: Educational and Industrial Testing Service; 1981.

Miller WR, Del Boca FK. Measurement of drinking behavior using the Form 90 family of instruments. Journal of Studies on Alcohol Supplement. 1994; 12:112–118. [PubMed: 7722987]

Miller WR, Tonigan JS, Longabaugh R. The Drinking Inventory of Consequences (DrInC): An instrument for assessing adverse consequences of alcohol abuse. Test manual. 1995; 4 Project MATCH Monograph Series.

Miller, WR. NIAAA Project MATCH Monograph Series. Washington: Government Printing Office; 1996. Form 90: A structured assessment interview for drinking and related behaviors.

Skinner HA, Allen BA. Alcohol dependence syndrome: measurement and validation. Journal of Abnormal Psychology. 1982; 91(3):199–209. [PubMed: 7096790]

Sobell, LC.; Sobell, MB. Alcohol consumption measures. In: Allen, JP.; Columbus, M., editors. Assessing alcohol problems: A guide for clinician and researchers. Rockville, MD: National Institute on Alcohol Abuse and Alcoholism; 1995. p. 55-73.

Spitzer RL, Williams JB, Gibbon M, First MB. The structured Clinical Interview for DSM-III-R (SCID): I. History, rational and description. Arch Gen Psychiatry. 1992; 49:624–629. [PubMed: 1637252]

Sullivan JT, Sykora K, Schneiderman J, Naranjo CA, Sellers EM. Assessment of alcohol withdrawal: The revised Clinical Institute Withdrawal Assessment for Alcohol scale (CIWA-AR). Br J Addict. 1989; 84:1353–1357. [PubMed: 2597811]

Szabo, S. The World Health Organization Quality of Life (WHOQOL) Assessment Instrument, in Quality of Life and Pharmacoeconomics in Clinical Trials. 2nd ed.. Spiker, B., editor. Philadelphia: Lippincott-Raven Publishers; 1996. p. 355-362.

Tonnigan JS, Miller WR, Brown JM. The reliability of Form 90: an instrument for assessing alcohol treatment outcome. J Stud Alcohol. 1997; 58:358–364. [PubMed: 9203116]

Ware, JE.; Kosinski, M.; Turner-Bowker, DM.; Gandek, B. How to Score Version 2 of the SF-12 Health Survey. Lincoln, RI: Quality-Metric. World Health Organization. The World Health Report 2001 – Mental; 2002.

References

Anton RF, O'Malley SS, Ciraulo D, et al. Combined pharmacotherapies and behavioral interventions for alcohol dependence. The COMBINE study: a randomized controlled trial. JAMA. 2006; 295:2003–2017. [PubMed: 16670409]

Anton RF, Oroszi G, O'Malley S, Couper D, Swift R, Pettinati H, et al. An evaluation of mu-opioid receptor (OPRM1) as a predictor of naltrexone response in the treatment of alcohol dependence: results from the Combined Pharmacotherapies and Behavioral Interventions for Alcohol Dependence (COMBINE) study. Arch Gen Psychiatry. 2008; 65(2):135–144. [PubMed: 18250251]

Bogenschutz MP, Tonigan S, Pettinati HM. Effects of alcoholism typology on response to naltrexone in the COMBINE study. ACER. 2008; 33(1):10–18.

Breiman, L.; Friedman, J.; Stone, CJ.; Olshen, RA. Classification and Regression Trees. New York: Chapman & Hall/CRC; 1984.

Breiman L. Random forests. Machine Learning. 2001; 45:5–32.

Capone C, Kahler CW, Swift RM, O'Malley S. Does family history of alcoholism moderate naltrexone's effects on alcohol use? Journal of Studies on Alcohol and Drugs. 2011; 72:135–140. [PubMed: 21138720]

Choi HK, Curhan G. Beer, liquor, and wine consumption and serum uric acid level: The Third National Health and Nutrition Examination Survey. Arthritis Care and Research. 2004; 51(6): 1023–1029. [PubMed: 15593346]

Dusseldorp E, Conversano C, Van Os BJ. Combining an additive and tree-based regression model simultaneously: STIMA. Journal of Computational and Graphical and Statistics. 2010; 19:514–530.

Dusseldorp E, van Mechelen I. Qualitative interaction trees: a tool to identify qualitative treatment-subgroup interactions. Statistics in Medicine. 2014; 33:219–237. [PubMed: 23922224]

Falk D, Wang XQ, Liu L, Fertig J, Mattson M, Ryan M, Johnson B, Stout R, Litten RZ. Percentage of subjects with no heavy drinking days: evaluation as an efficacy endpoint for alcohol clinical trials. Alcohol Clin Exp Res. 2010; 34(12):2022–2034. [PubMed: 20659066]

Foster J, Taylor JMG, Ruberg SJ. Subgroup identification from randomized clinical trial data. Statistics in Medicine. 2011; 30:2867–2880. [PubMed: 21815180]

Fucito L, Park A, Gulliver SB, Mattson ME, Gueorguieva R, O'Malley S. Cigarette Smoking Predicts Differential Benefit from Naltrexone for Alcohol Dependence. Biological Psychiatry. 2012; 72(10):832–838. [PubMed: 22541040]

Gueorguieva R, Wu R, Donovan D, Rounsaville B, Couper D, Krystal J, O'Malley S. "Baseline Trajectories of Drinking Moderate Acamprosate and Naltrexone Effects in the COMBINE study". Alcohol Clin Exp Res. 2011; 35(3):523–531. [PubMed: 21143249]

Gueorguieva R, Wu R, O'Connor P, Weissner C, Fucito L, Hoffmann S, Mann K, O'Malley SS. Predictors of abstinence from heavy drinking during treatment in COMBINE and external validation in PREDICT. Alcohol Clin Exp Res. 2014; 38(10):2647–2656. [PubMed: 25346505]

Hall SM, Havassy BE, Wasserman DA. Commitment to abstinence and acute stress in relapse to alcohol, opiates, and nicotine. J Consult Clin Psychol. 1990; 58:175–181. [PubMed: 2335634]

Higuchi S. for the Japanese Acamprosate Study Group. E. fficacy of acamprosate for the treatment of alcohol dependence long after recovery from withdrawal syndrome: A randomized, double-blind, placebo-controlled study conducted in Japan (Sunrise Study). J Clin Psychiatry. 76(2):181–188. [PubMed: 25742205]

Karpyak VM, Biernacka JM, Geske JR, et al. Genetic markers associated with abstinence length in alcohol-dependent subjects treated with acamprosate. Translational Psychiatry. 2014; 4:e462. [PubMed: 25290263]

Kraemer HC, Wilson GT, Fairburn CG, Agras WS. Mediators and moderators of treatment effects in randomized clinical trials. Archives of General Psychiatry. 2002; 59:877–883. [PubMed: 12365874]

Lipkovich I, Dmitrienko A. Strategies for identifying predictive biomarkers and subgroups with enhanced treatment effect in clinical trials using SIDES. Journal of Biopharmaceutical Statistics. 2014; 24:130–153. [PubMed: 24392982]

Littleton J. Acamprosate in alcohol dependence: how does it work? Addiction. 1995; 90:1179–1188. [PubMed: 7580816]

Loh WY, He X, Man M. A regression tree approach to identifying subgroups with differential treatment effects. 2014 arXiv:1410.1932v1.

Mann K, Kiefer F, Spanagel R, Littleton J. Acamprosate: recent findings and future research directions. Alcohol Clin Exp Res. 2008; 32:1105–1110. [PubMed: 18540918]

Mann K, Kiefer F, Smolka M, Gann H, Wellek S, Heinz A. Searching for responders to acamprosate and naltrexone in alcoholism treatment: Rational and design of the PREDICT study. ACER. 2009; 33(4):674–683.

Mason BJ. Treatment of alcohol-dependent outpatients with acamprosate: a clinical review. J Clin Psychiatry. 2001; 62(S20):42–48. [PubMed: 11584875]

Mason BJ, Goodman AM, Chabac S, Lehert P. Effect of oral acamprosate on abstinence in patients with alcohol dependence in a double-blind, placebo -controlled trail: the role of patient motivation. Journal of Psychiatry Research. 2006; 40:383–393.

Miller, WR., editor. Combined Behavioral Intervention manual: A clinical research guide for therapists treating people with alcohol abuse and dependence. Vol. 1. Bethesda, MD: National Institute on Alcohol Abuse and Alcoholism; 2004.

Negassa A, Ciampi A, Abrahamowicz M, Shapiro S, Boivin J-F. Tree-structured subgroup analysis for censored survival data: Validation of computationally inexpensive model selection criteria. Statistics and Computing. 2005; 15:231–239.

Oliveira A, Rodriguez-Artalejo F, Lopes C. Alcohol intake and systemic markers of inflammation - shape of the association according to sex and body mass index. Alcohol and Alcoholism. 2010; 45(2):119–125. [PubMed: 20083478]

Paille FM, Guelfi JD, Perkins AC, Royer RJ, Steru L, Parot P. Double-blind randomized multicenter trial of acamprosate in maintaining abstinence from alcohol. Alcohol and Alcoholism. 1995; 30:239–247. [PubMed: 7662044]

Perney P, Lehert P, Mason BJ. Sleep disturbance in alcoholism: proposal of a simple measurement, and results from a 24-week randomized controlled study of alcohol-dependent patients assessing acamprosate efficacy. Alcohol Alcohol. 2012; 47(2):133–139. [PubMed: 22218671]

Pettinati, HM.; Weiss, RD.; Miller, WR.; Donovan, DM.; Ernst, DB.; Rounsaville, BJ. Medical Management (MM) treatment manual: A clinical research guide for medically trained clinicians providing pharmacotherapy as part of the treatment for alcohol dependence. Vol. 2. Bethesda, MD: National Institute on Alcohol Abuse and Alcoholism; 2004.

R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2013. URL http://www.R-project.org/ ISBN 3-900051-07-0

Ray LA, Hutchison KE. Effects of naltrexone on alcohol sensitivity and moderators of medication response. Arch Gen Psychiatry. 2007; 64(9):1069–1077. [PubMed: 17768272]

Saivin S, Hulot T, Chabac S, Potgieter A, Durbin P, Houin G. Clinical pharmacokinetics of acamprosate. Clinical Pharacokinetics. 1998; 35:331–345.

Sass H, Soyka M, Mann K, Zieglgänsberger W. Relapse prevention by acamprosate results from a placebo-controlled study on alcohol dependence. Arch Gen Psychiatry. 1996; 53:673–680. [PubMed: 8694680]

Staner L, Boeijinga P, Danel T, Gendre I, Muzet M, Landron F, Luthringer R. Effects of acamprosate on sleep during alcohol withdrawal: A double-blind placebo-controlled polysomnographic study in alcohol-dependent patients. Alcohol Clin Exp Res. 2006; 30(9):1492–1499. [PubMed: 16930211]

Subbaraman MS, Lendle S, van der Laan M, Kaskutas LA, Ahern L. Cravings as a mediator and moderator of drinking outcomes in the COMBINE study. Addiction. 2013; 108:1737–1744. [PubMed: 23668624]

Su X, Meneses K, McNees P, Johnson WO. Interaction trees: exploring the differential effects of an interaction program for breast cancer survivors. Applied Statistics. 2011; 60:457–474.

The COMBINE Study Research Group. Testing Combined Pharmacotherapies and Behavioral Interventions in Alcohol Dependence: Rationale and Methods. Vol. 27. Alcoholism: Clinical Exp Research; 2003. p. 1107-1122.

Worley MJ, Witkiewitz K, Brown SA, Kivlahan DR, Longabaugh R. Social network predictors of naltrexone and behavioral treatment effects on heavy drinking in the COMBINE study. Alcohol Clin Exp Res. 2014; 39:93–100. [PubMed: 25623409]

Zhang, H.; Singer, B. Recursive Partitioning and Applications. New York: Springer; 2010.

Zhang H, Legro RS, Zhang J, Zhang L, Chen X, Huang H, Casson PR, et al. Decision trees for identifying predictors of treatment effectiveness in clinical trials and its application to ovulation in a study of women with polycystic ovary syndrome. Human Reproduction. 2010; 25:2612–2621. [PubMed: 20716558]

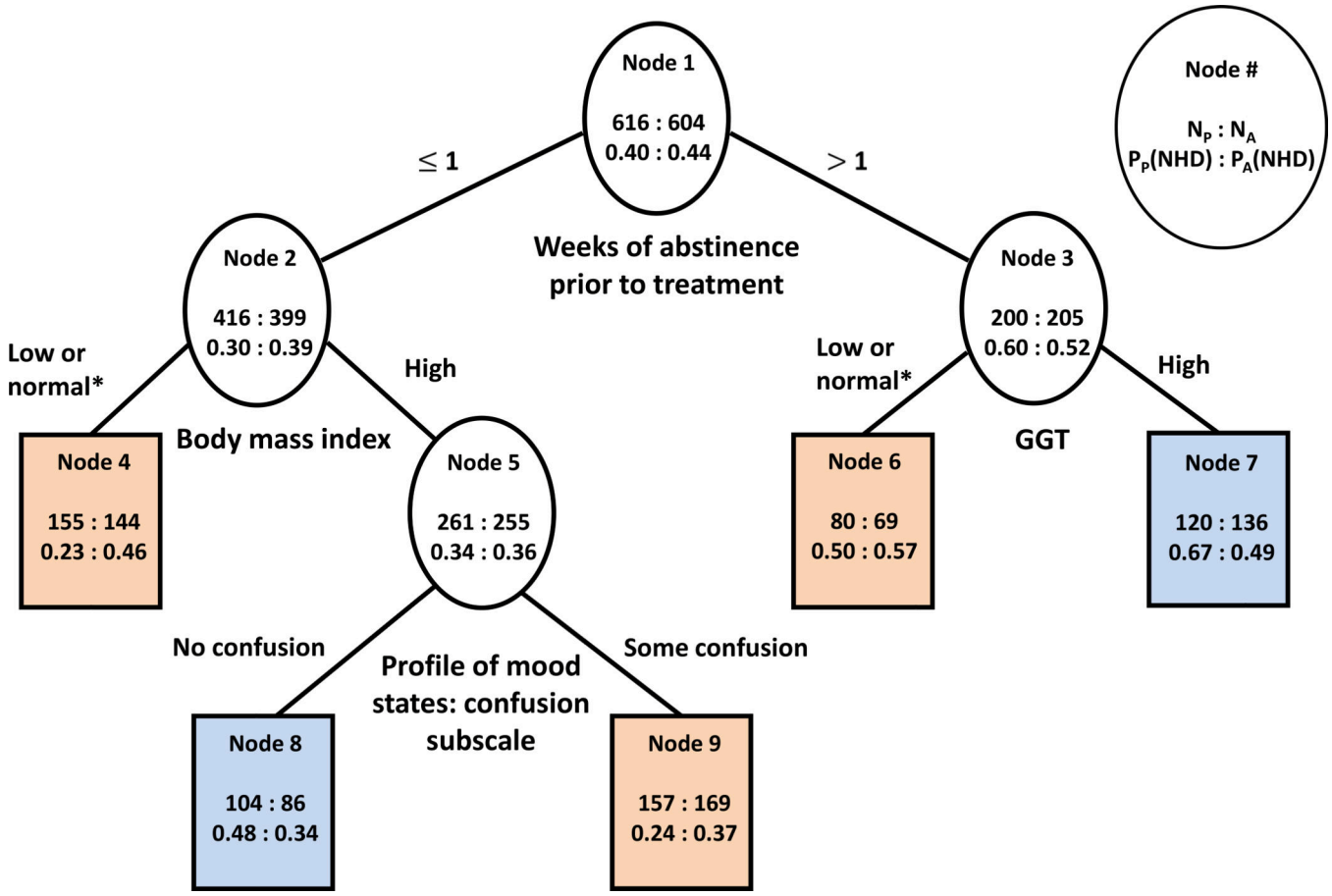

Figure 1.

Classification tree built using the approach of Zhang et al. (2010).

*Low or normal means within the normal range or below the lower limit of the normal range.

Terminal nodes in orange denote better outcome on acamprosate while terminal nodes in blue denote worse outcome on acamprosate than on placebo.

N_A: Number of participants on acamprosate; **N_P**: Number of participants on placebo; **P_A(NHD)**: Proportion of participants with no heavy drinking during the last 8 weeks of treatment on acamprosate; **P_P(NHD)**: Proportion of participants with no heavy drinking during the last 8 weeks of treatment on acamprosate

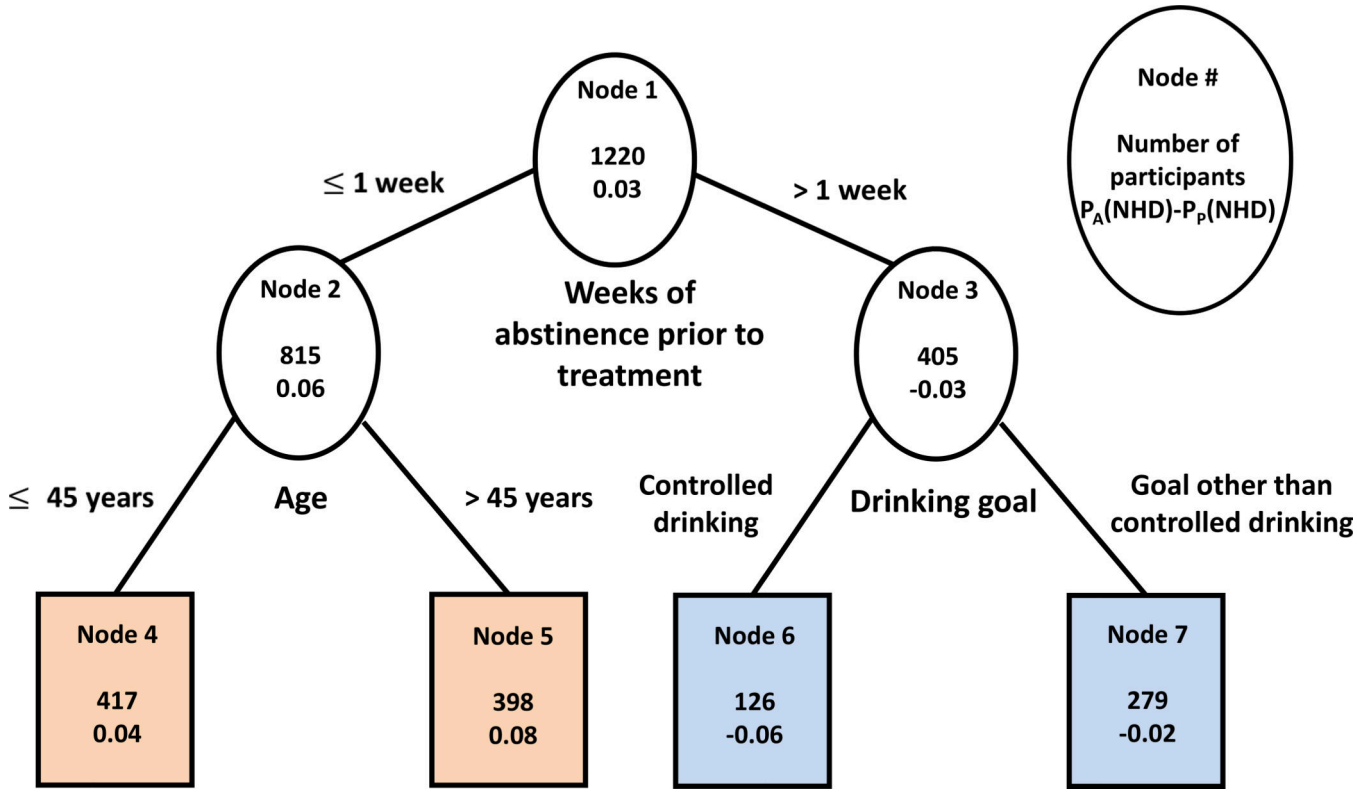

Figure 2.

Regression tree built using the approach of Foster et al. (2011).

Terminal nodes in orange denote better outcome on acamprosate than on placebo while terminal nodes in blue denote worse outcome on acamprosate than on placebo.

$P_A(NHD)-P_P(NHD)$: Difference in average probabilities of no heavy drinking during the last 8 weeks of treatment on acamprosate and on placebo for participants in the node.

Table 1

Predictors in tree analyses for moderator analyses in COMBINE.

DOMAIN	VARIABLES
Demographics	Age, Gender, Race, Marital Status, Education, Employment, Family income
Alcohol Consumption	% heavy days, % abstinent days, Consecutive days abstinent, Peak BAC
Alcohol Severity	SCID symptom count, CIWA, DRINC Total, ADS, Age of onset
Prior Alcohol Treatment Drinking Goal	Detoxification, AA attendance, Any treatment complete abstinence, conditional abstinence, controlled drinking, other
Family History	Alcohol, Smoking
Craving	OCDS Total Score
Smoking	Current smoker
Drug use	Cannabis use
Physical Exam	BMI, Pulse, Blood pressure
Laboratory Analysis	Urine and blood test results
Alcohol Abstinence Self-Efficacy	Total and Subscale Scores
URICA	Overall Readiness and 4 Subscale scores
WHO Quality of Life	Environment, Physical, Psychological, Social Relationships
SF12	Physical health
Profile of Mood States	Tension, Depression, Anger, Vigor, Fatigue, Confusion
Perceived Stress	PSS total
Sleep problems	Any symptom of insomnia, sleep disturbance, problems of sleep, and decreased sleep
Important People	Number of in-network daily drinkers
Legal Problems	Arrested

Table 2

Top ten moderator variables according to two different statistical criteria identified using the Foster et al. (2011) approach.

Rank	Top variables according to percent increase in Mean Squared Error after random permutation	Top variables according to percent increase in node impurity after random permutation
1	Consecutive days of abstinence prior to randomization	Consecutive days of abstinence prior to randomization
2	Age	Age
3	Drinking goal	Drinking goal
4	Uric acid	SCID alcohol dependence symptoms
5	AASE Confidence: Negative affect subscale score	Heavy drinking days per week
6	AASE Confidence: Social subscale score	Family income
7	Alcohol Abstinence Self Efficacy (AASE): Total confidence score	Uric acid
8	AASE Confidence: Withdrawal/urge subscale score	Days abstinent per week
9	Blood Alcohol Concentration (BAC) peak	Self-reported health
10	Self-reported health	Blood Alcohol Concentration (BAC) peak

Table 3

Results from SIDESscreen approach (Lipkovich and Dmitrienko, 2014).

Sample description	N	Percent NHD on acamprosate	Percent NHD on placebo	Unadjusted p-value	Adjusted p-value
All subjects	1220	0.44	0.40	0.08	0.08
Four or fewer days of abstinence AND had prior treatment AND not obese	168	0.42	0.13	<0.0001	0.04
1 week or less of abstinence AND had prior treatment AND not overweight or obese	137	0.52	0.19	<0.0001	0.06