

Original Paper

# Exploratory Data Mining Techniques (Decision Tree Models) for Examining the Impact of Internet-Based Cognitive Behavioral Therapy for Tinnitus: Machine Learning Approach

Hansapani Rodrigo<sup>1,2</sup>, PhD; Eldré W Beukes<sup>2,3,4</sup>, PhD; Gerhard Andersson<sup>5,6</sup>, PhD; Vinaya Manchaiah<sup>2,3,7,8</sup>, MBA, AUD, PhD

<sup>1</sup>School of Mathematical and Statistical Sciences, University of Texas Rio Grande Valley, Edinburg, TX, United States

<sup>2</sup>Virtual Hearing Lab, Beaumont, TX, United States

<sup>3</sup>Department of Speech and Hearing Sciences, Lamar University, Beaumont, TX, United States

<sup>4</sup>Department of Vision and Hearing Sciences, School of Psychology and Sport Science, Anglia Ruskin University, Cambridge, United Kingdom

<sup>5</sup>Department of Behavioral Sciences and Learning, Department of Biomedical and Clinical Sciences, Linköping University, Linköping, Sweden

<sup>6</sup>Department of Clinical Neuroscience, Division of Psychiatry, Karolinska Institute, Stockholm, Sweden

<sup>7</sup>Department of Speech and Hearing, School of Allied Health Sciences, Manipal Academy of Higher Education, Karnataka, India

<sup>8</sup>Department of Speech-Language Pathology and Audiology, University of Pretoria, Gauteng, South Africa

**Corresponding Author:**

Hansapani Rodrigo, PhD

School of Mathematical and Statistical Sciences

University of Texas Rio Grande Valley

1201 W University Drive

Edinburg, TX, 78539

United States

Phone: 1 9566652313

Email: [hansapani.rodrigo@utrgv.edu](mailto:hansapani.rodrigo@utrgv.edu)

## Abstract

**Background:** There is huge variability in the way that individuals with tinnitus respond to interventions. These experiential variations, together with a range of associated etiologies, contribute to tinnitus being a highly heterogeneous condition. Despite this heterogeneity, a “one size fits all” approach is taken when making management recommendations. Although there are various management approaches, not all are equally effective. Psychological approaches such as cognitive behavioral therapy have the most evidence base. Managing tinnitus is challenging due to the significant variations in tinnitus experiences and treatment successes. Tailored interventions based on individual tinnitus profiles may improve outcomes. Predictive models of treatment success are, however, lacking.

**Objective:** This study aimed to use exploratory data mining techniques (ie, decision tree models) to identify the variables associated with the treatment success of internet-based cognitive behavioral therapy (ICBT) for tinnitus.

**Methods:** Individuals (N=228) who underwent ICBT in 3 separate clinical trials were included in this analysis. The primary outcome variable was a reduction of 13 points in tinnitus severity, which was measured by using the Tinnitus Functional Index following the intervention. The predictor variables included demographic characteristics, tinnitus and hearing-related variables, and clinical factors (ie, anxiety, depression, insomnia, hyperacusis, hearing disability, cognitive function, and life satisfaction). Analyses were undertaken by using various exploratory machine learning algorithms to identify the most influencing variables. In total, 6 decision tree models were implemented, namely the classification and regression tree (CART), C5.0, GB, XGBoost, AdaBoost algorithm and random forest models. The Shapley additive explanations framework was applied to the two optimal decision tree models to determine relative predictor importance.

**Results:** Among the six decision tree models, the CART (accuracy: mean 70.7%, SD 2.4%; sensitivity: mean 74%, SD 5.5%; specificity: mean 64%, SD 3.7%; area under the receiver operating characteristic curve [AUC]: mean 0.69, SD 0.001) and gradient boosting (accuracy: mean 71.8%, SD 1.5%; sensitivity: mean 78.3%, SD 2.8%; specificity: 58.7%, SD 4.2%; AUC: mean 0.68, SD 0.02) models were found to be the best predictive models. Although the other models had acceptable accuracy (range 56.3%-66.7%) and sensitivity (range 68.6%-77.9%), they all had relatively weak specificity (range 31.1%-50%) and AUCs (range

0.52-0.62). A higher education level was the most influencing factor for ICBT outcomes. The CART decision tree model identified 3 participant groups who had at least an 85% success probability following the undertaking of ICBT.

**Conclusions:** Decision tree models, especially the CART and gradient boosting models, appeared to be promising in predicting ICBT outcomes. Their predictive power may be improved by using larger sample sizes and including a wider range of predictive factors in future studies.

(*J Med Internet Res* 2021;23(11):e28999) doi: [10.2196/28999](https://doi.org/10.2196/28999)

## KEYWORDS

tinnitus; internet interventions; digital therapeutics; cognitive behavioral therapy; artificial intelligence; machine learning; data mining; decision tree; random forest

## Introduction

### Background

Tinnitus is the perception of a sound in the ears or head in the absence of a corresponding external sound source. It is very prevalent; it has been estimated that 10% to 15% of the adult population experiences tinnitus [1]. Various conditions are associated with developing tinnitus, such as ear disorders [2], exposure to loud noise, the presence of hearing loss, and increasing age [3]. Tinnitus experiences are highly heterogeneous in terms of how it manifests (eg, the types of sounds experienced, how individuals react to their sounds, the associated comorbidities, etc) and how individuals with tinnitus respond to treatment [4]. Although a majority of those with tinnitus are not bothered by their tinnitus, a significant number of people with tinnitus experience distressing tinnitus that affects their quality of life [5]. Nevertheless, tinnitus can affect people in different ways; the most common complaints include annoyance, irritability, fatigue, stress, sleep problems, and trouble concentrating [6]. Moreover, distressing tinnitus is often associated with an increased risk of anxiety and depression [7,8]. Various management strategies are used to help persons with tinnitus, including sound therapy (eg, hearing aids and masking); informational counseling for aiding people's understanding of tinnitus; and psychological approaches that address unhelpful thought patterns and reactions to tinnitus, such as cognitive behavioral therapy (CBT). Of these, CBT has the highest level of research evidence on reducing tinnitus distress [9,10].

Although the use of CBT is recommended in many tinnitus practice guidelines [11], it is seldomly provided. This is partly due to a lack of trained professionals who can offer CBT for tinnitus in an in-person format. To overcome this barrier, internet-based CBT (ICBT) was developed in the late 1990s [12]. In ICBT, the treatment strategies are offered to individuals with tinnitus as self-help materials that are provided over the internet together with professional guidance [13]. The feasibility and efficacy of such an approach have been demonstrated among several populations in Sweden, Germany, Australia, the United Kingdom [14], and, more recently, the United States [15,16]. In general, studies have shown that nearly 50% to 60% of those who undergo ICBT will experience a clinically significant reduction in tinnitus distress [17,18]. To date, no strong predictors of ICBT outcomes have been identified to indicate who is likely to benefit from ICBT interventions. The predictors of outcomes that were identified when examining the long-term

(1 year) outcomes of ICBT in the United Kingdom were higher baseline tinnitus severity, more engagement with the ICBT program (ie, more modules opened), and higher self-reported satisfaction with the intervention [18]. To further explore predictors of outcomes, various univariate and multivariate (ie, logistic and linear) regression models were applied to a combined data set of multiple ICBT studies [19]. These linear and logistic regression models identified education level (linear regression:  $P=.01$ ; logistic regression:  $P<.001$ ) and baseline tinnitus severity (linear regression:  $P<.001$ ; logistic regression:  $P<.001$ ) to be significant predictor variables that contribute to reductions in tinnitus severity following an ICBT intervention. As per the linear regression model, participants who had received disability allowance showed a 25.30-point less (95% CI  $-46.35$  to  $-4.24$ ) Tinnitus Functional Index (TFI) score reduction when compared to those who did not have to work less due to tinnitus after adjusting for baseline tinnitus severity and participants' education levels. Although many other predictors, including age and tinnitus duration [19], were not identified to be significant under these linear models, these variables might have a nonlinear association with the response.

In the last 2 decades, various artificial intelligence and machine learning techniques have been developed and applied to hearing health data. Such approaches have mainly been used for disease profiling, although some studies have focused on the prediction of treatment outcomes [20-24]. It is noteworthy that the intervention trials in audiology and tinnitus research usually involve a few hundred participants and the collection of generally extensive data regarding demographic characteristics and clinical variables. Such a data set with many predictor variables may be best handled by exploratory data mining techniques, such as the use of tree-based models (eg, the random forest [RF] model). Such models tend to perform well even in the presence of multicollinearity among a large number of predictor variables, as these models decorrelate the variables [25]. For example, a recent study that examined various machine learning algorithms for predicting CBT treatment outcomes in the tinnitus population suggested that gradient boosted trees (area under the receiver operating characteristic curve [AUC]=0.89) have the best predictive power [21]. This study found that subjectively perceived tinnitus-related impairment, depression, sleep problems, physical health-related impairments in quality of life, the time spent on completing questionnaires, and educational level highly contributed to the model's predictions. However, no previous studies have examined the

application of artificial intelligence and machine learning techniques to ICBT outcomes in tinnitus research.

## Objectives

To further explore outcome predictors for ICBT, this study aimed to examine the applications of various exploratory data mining techniques based on decision tree models. In particular, we wanted to (1) investigate which types of decision tree models were the most applicable to ICBT outcome prediction (ie, models with the best accuracy and predictive power) and (2) identify the most relevant predictive factors of ICBT outcomes by using the most appropriate decision tree models.

## Methods

### Study Design and Ethical Considerations

We included 228 participants who previously underwent ICBT for tinnitus and whose data were collected as a part of 3 separate ICBT trials [17,18,26] that were conducted from 2016 to 2018. This study was a secondary analysis of these ICBT intervention studies. Ethical clearance was obtained from the Faculty of Science and Technology Research Ethics Panel of Anglia Ruskin University (reference numbers: FST/FREP/14/478 and FST/FREP/14/478) and the East of England–Cambridge South Research Ethics Committee (reference number: 16/EE/0148) and Health Research Authority (Integrated Research Application System project ID: 195565).

### Participant Characteristics

A heterogeneous sample of individuals with tinnitus was obtained; thus, the sample represented typical tinnitus populations, as seen in [Multimedia Appendix 1](#). The average age was 55.14 years (SD 12.92 years), and 98 out of the 228 (43%) participants were females. The majority of participants (154/228, 67.5%) had long-standing tinnitus with a mean duration of 17.68 years (SD 19.42 years). Of the 228 participants, 59 (25.9%) had completed high school education, 61 (26.8%) had an undergraduate degree, and only 30 (13.2%) had a postgraduate degree. Approximately 48% (109/228, 47.8%) of the participants experienced tinnitus in both ears, 26.8% (61/228) of them experienced tinnitus in 1 ear, and the others reported experiencing tinnitus in their head or in other locations. The majority (159/228, 69.7%) of participants did not wear hearing aids, and 25.4% (58/228) of them had sought tinnitus treatment previously.

### Intervention

The study participants completed an 8-week ICBT intervention that was presented in a self-help format [13,27]. The intervention was administered by using a secure e-platform [28,29]. During this 8-week period, participants were presented with 2 to 3 learning modules that contained various elements of CBT that were specifically adapted for tinnitus, including applied relaxation, cognitive restructuring, and imagery. The digital materials were presented by using text, images, and videos. In addition, various exercises were presented in these learning modules to improve engagement.

## Data Collection

The baseline data collection included an extensive questionnaire that focused on demographics and tinnitus-related and treatment-related information. Outcome data were gathered by using standardized primary and secondary self-reported questionnaires, which were administered before (baseline), during (weekly), and after the intervention. The primary outcome was a change in tinnitus severity, as measured by the TFI [30]. The secondary outcome measures included the Insomnia Severity Index [31] (a measure of insomnia), the Generalized Anxiety Disorder-7 [32] (a measure of anxiety); the Patient Health Questionnaire-9 [33] (a measure of depressive symptoms); the Hearing Handicap Inventory for Adults Screening version [34] (a measure of self-reported hearing disability); the Hyperacusis Questionnaire [35], which was used to assess the presence hyperacusis (ie, reduced tolerance to everyday sounds); the Cognitive Failures Questionnaire [36], which was used to assess cognitive functions; and the Satisfaction with Life Scales [37], which were used to assess global life satisfaction.

## Data Analyses

### Variables

The primary outcome variable (the dependent variable) in this study was a change in tinnitus severity. A 13-point reduction in TFI scores following the ICBT intervention was regarded as a clinically significant (successful) treatment outcome [30]. Significant differences in scores were assessed by using paired sample *t* tests. All tests were two-tailed, and significance was set to  $P=.05$ . There were 33 predictor variables selected, as outlined in [Multimedia Appendix 2](#). These included the following:

- 7 demographic variables (ie, age, gender, education level, employment type, noise exposure, the presence of psychological conditions, and tinnitus that affects the ability to work)
- 15 tinnitus and hearing-related variables (ie, baseline tinnitus severity, tinnitus duration, how often tinnitus is heard, tinnitus location, 9 different types of tinnitus, tinnitus in which multiple tones are heard, and the presence of hearing loss)
- 4 treatment-related variables (ie, past treatment sought, tinnitus maskability, hearing aid use, and medication use)
- 7 clinical factors (ie, anxiety, depression, insomnia, hyperacusis, hearing disability, cognitive functions, and life satisfaction).

### Decision Tree Models (Classifiers)

The data analysis focused on decision tree–based models, as they play an essential role in exploratory data mining and facilitate human decision-making by providing decision rules [38]. Despite their simplicity, decision trees usually exhibit high variance in their predictions and are not consistently robust. Given these issues, their powerful counterparts, such as the RF [39], gradient boosting (GB) [40], and extreme GB (XGBoost) [41], models were selected. For comparison, 6 decision tree models were used, namely the classification and regression tree (CART) [42], C5.0 [43], GB, XGBoost, AdaBoost algorithm

[44], and RF models. As the CART, C5.0, and RF decision tree models involve stratifying or segmenting the predictor space into a number of nonoverlapping regions [38], recursive binary splitting for classification via the Gini index was performed [45]. Many of these decision tree types have been applied to audiological data, and they were found to provide good results in previous studies [20-24].

## Data Analysis Steps

### Summary of Data Analysis Steps

The analyses were performed in 4 stages. First, the data were split into training and testing data, and the classifier models were trained on a data set before testing. Second, the six classifiers were applied to the test data to identify the most suitable models based on their performance evaluation. Third, the two best models were used to determine the predictors of ICBT outcomes. Fourth, the optimal CART decision tree model was used to identify the participants who were more (or less) likely to benefit from ICBT treatment. The steps are described in more detail below.

### Step 1: Classifier Training

Prior to applying the decision tree classifiers, the entire data set was divided into the training (183/228, 80.3%) and testing (45/228, 19.7%) data sets. The training data set was used to develop the corresponding data mining model, while the testing data set was used to evaluate the model predictions. As the training data set was relatively small ( $n=183$ ), a repeated 3-fold cross-validation was performed. This was not done for the CART model, for which the full training data set was used for model training. With this approach, each fold was given a chance to act as their own validation set to minimize the propensity of model overfitting. In total, 10 different models were created with several random initializations for each data mining method. Hyperparameter tuning for each of these decision models was performed, as required. For instance, when training the RF models, we explored a range of different numbers of predictors for the splitting at each tree node and their impact on the models' performance.

### Step 2: Classifier Performance Evaluation

The trained models were evaluated (ie, by using the testing data set) in terms of their mean predictive accuracy, sensitivity (true positive rate), specificity (true negative rate), and AUCs. These were presented as means and SDs and based on the 10 replicated models for each data mining technique. The AUC is used as a measurement of model discrimination power. The optimal decision tree models were selected based on having the highest AUC value. In general, models with an AUC of 0.5 have no discriminatory power, models with an AUC of 0.7 to 0.8 are considered acceptable, models with an AUC of 0.8 to 0.9 are considered excellent, and models with an AUC of  $>0.9$  are deemed to have outstanding discriminatory power (ie,

the ability to identify patients with and without a disease or condition based on a new set of data).

### Step 3: Predictors of ICBT Outcomes

Decision tree-based classifiers provide insights on different participant groups who show promising results following ICBT. After identifying the two most optimal models, the model-agnostic posthoc framework Shapley additive explanations (SHAP) was used for analyzing ICBT outcome predictors [46,47]. This framework facilitates model interpretations and identifies the most influential factors that result in successful ICBT outcomes (ie, a reduction in TFI scores following the ICBT intervention). SHAP measure the impact of variables and take into account variables' interactions with other variables. SHAP values indicate the importance of a feature and are calculated by comparing model predictions that account and do not account for a given feature. However, since the order in which a model sees features can affect its predictions, this comparison is done in every possible order, so that the features are compared in a fair manner.

### Step 4: Identification of Participants Who Are Most Likely to Benefit From ICBT

The CART decision tree model was used to identify the participants who were the most (or least) likely to benefit from ICBT. In training, a minimum split of 20 and a max depth of 10 were used as the control parameters for the CART decision tree models. Tree pruning was conducted to reduce the overfitting in the CART decision tree models, although the best decision tree model remains the same even after pruning.

The data analysis was performed with R version 4.0.3 (R Foundation for Statistical Computing) software. The code is available in the GitHub repository for this study [48]. The data can be made available upon reasonable request.

## Results

### ICBT Effects

Undertaking ICBT significantly reduced tinnitus severity scores ( $t_{227}=16.37$ ;  $P<.001$ ) from a mean baseline severity score of 57.93 (SD 19.17) to a mean post-ICBT severity score of 34.22 (SD 22.78), as measured by the TFI. A clinically significant 13-point change in TFI scores was achieved by 150 of the 228 participants (65.8%) after the intervention.

### Decision Tree Model Performance Evaluations

Table 1 contains the model evaluation information of all 6 decision tree classifiers that were based on the test data. Following training via the 3-fold cross-validation method, the mean accuracies of the six decision tree classifiers ranged between a minimum of 56.3% (the C5.0 model) to a maximum of 71.8% (the GB model). Model predictions showed variations in their sensitivity (range 68.6%-78.3%), specificity (range 31.1%-64%), and AUC values (range 0.52-0.69).

**Table 1.** Decision tree model evaluations.

Classification model	Accuracy (%), mean (SD)	Sensitivity (%; true positive rate), mean (SD)	Specificity (%; true negative rate), mean (SD)	AUC <sup>a</sup> , mean (SD)
Classification and regression decision tree	70.7 (2.4)	74 (5.5)	64 (3.7)	0.69 (0.001)
C5.0	56.3 (1.1)	68.6 (1.9)	31.1 (6.3)	0.52 (0.001)
Gradient boosting	71.8 (1.5)	78.3 (2.8)	58.7 (4.2)	0.68 (0.02)
Extreme gradient boosting	65 (4.1)	77.9 (8.7)	39.2 (6.6)	0.62 (0.08)
AdaBoost algorithm	63.6 (3.2)	73.3 (5.2)	44 (7.8)	0.58 (0.05)
Random forest	66.7 (3)	75 (6.1)	50 (7.2)	0.60 (0.01)

<sup>a</sup>AUC: area under the receiver operating characteristic curve.

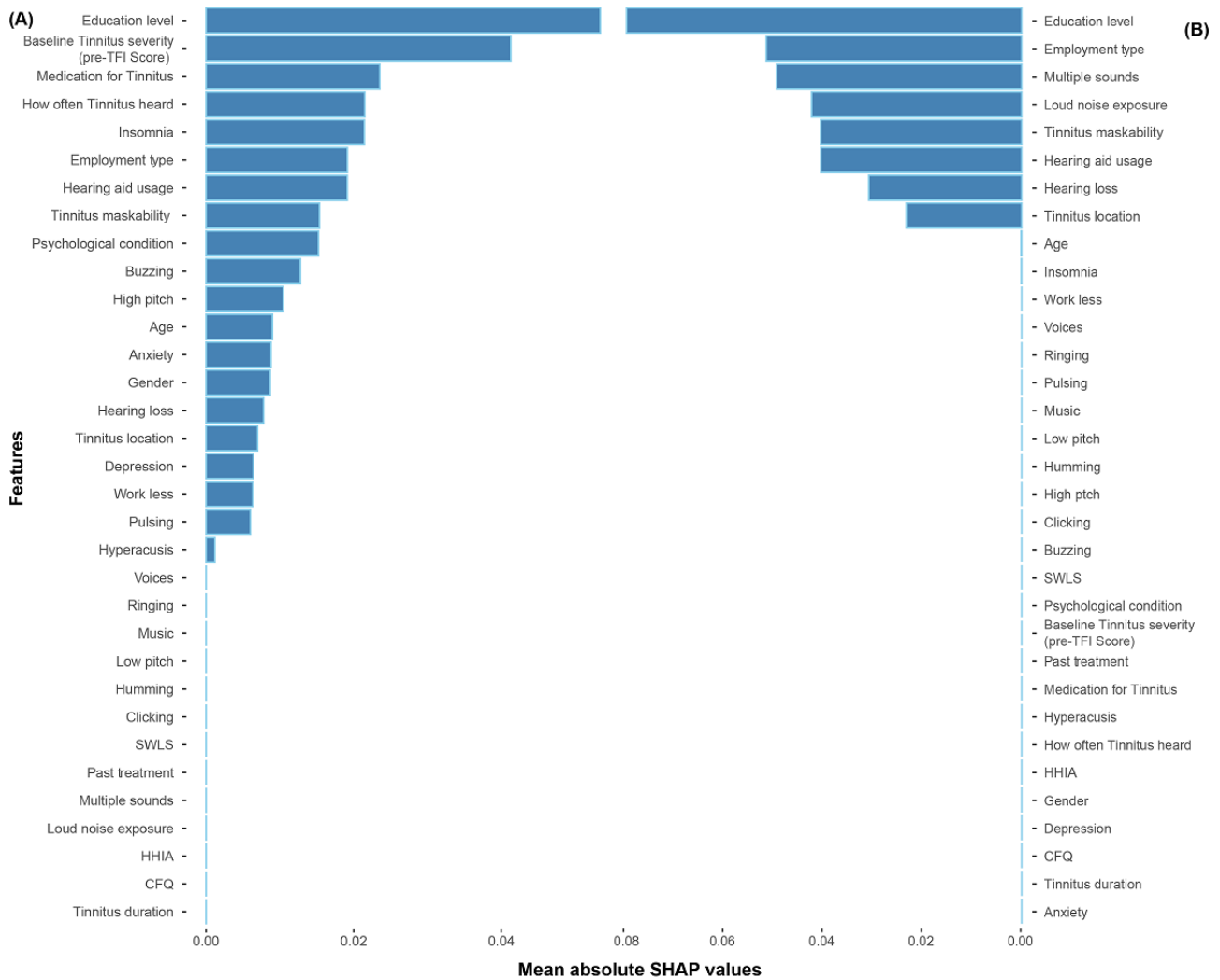
None of the six models were considered robust, as their AUC values were below 0.80. As the CART and GB classifiers were found to be superior compared to the other four models when considering all evaluation measurements as a whole (accuracy, sensitivity, specificity, and the AUC), these two models were further examined.

### Feature Importance

The SHAP framework was applied to the CART and GB classifiers to estimate each predictor variable's importance in predicting ICBT outcomes (Figure 1). Variables with larger SHAP values are relatively more important in terms of their contributions (ie, feature contribution) to a model prediction. Education level (average SHAP values: GB model=0.053; CART model=0.079) was identified as the most important influencing factor in both models. Although they were not

ranked in the same order, the other features that ranked within the top 10 features for both models were employment type (average SHAP values: GB model=0.019; CART model=0.051), hearing aid usage (average SHAP values: GB model=0.019; CART model=0.040), and tinnitus maskability (average SHAP value: GB model=0.015; CART model=0.040). The differences between these models were that the GB model ranked baseline tinnitus severity (average SHAP value=0.041), how often tinnitus is heard (average SHAP value=0.022), insomnia (average SHAP value=0.021), the use of medication for tinnitus (average SHAP value=0.024), and the presence of a psychological condition (average SHAP value=0.015) among the top 10 features, whereas the CART model ranked the presence of multiple sounds (average SHAP value=0.049), loud noise exposure (average SHAP value=0.042), and tinnitus location (average SHAP value=0.023) as key features.

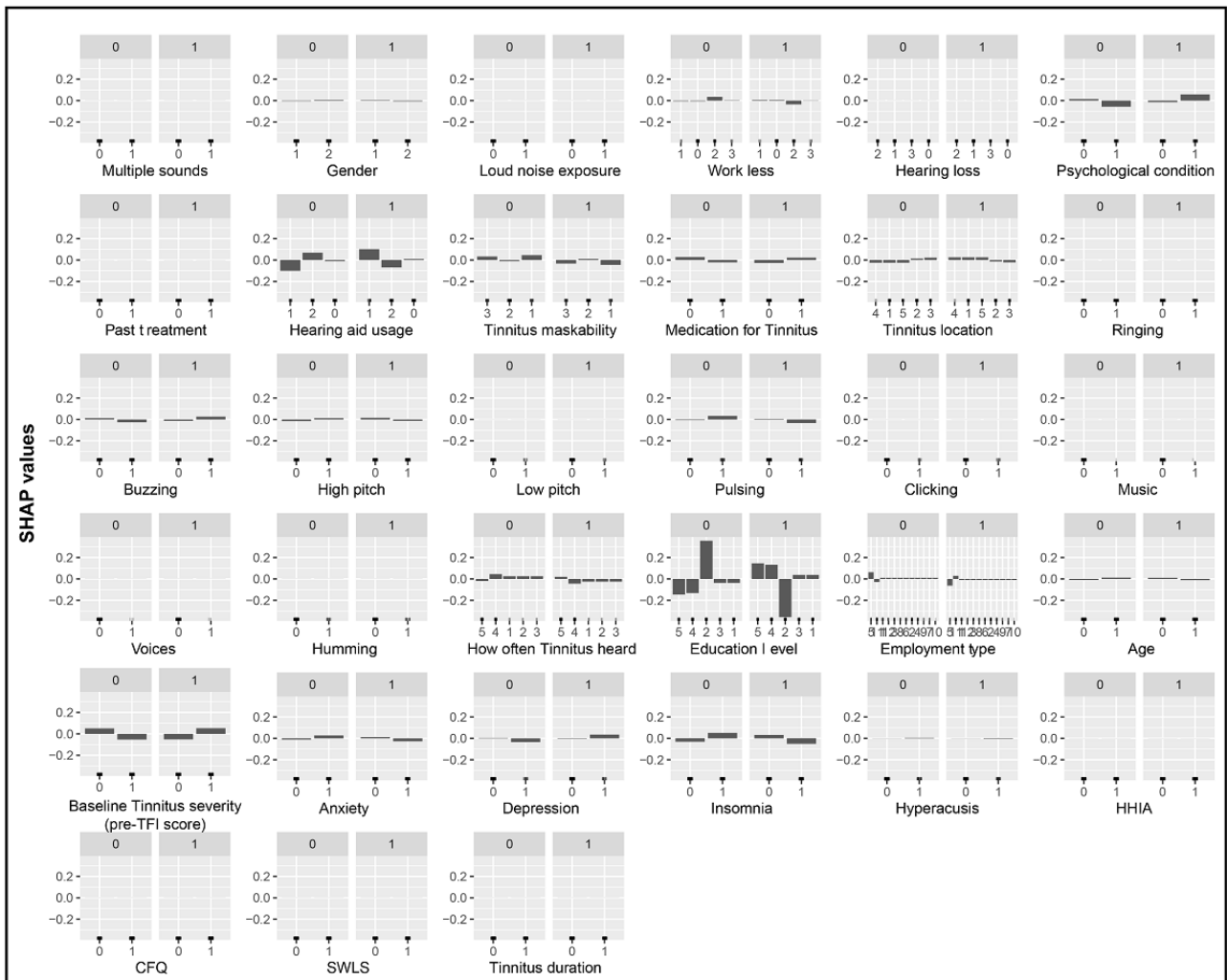
**Figure 1.** Feature importance based on the mean absolute SHAP values from (A) the best gradient model and (B) from the best classification and regression tree decision tree model. These SHAP values represent the absolute change in log odds. Relatively higher importance is indicated with larger SHAP values. CFQ: Cognitive Failures Questionnaire; HHIA: Hearing Handicap Inventory for Adults; SHAP: Shapley additive explanations; SWLS: Satisfaction with Life Scales; TFI: Tinnitus Functional Index.



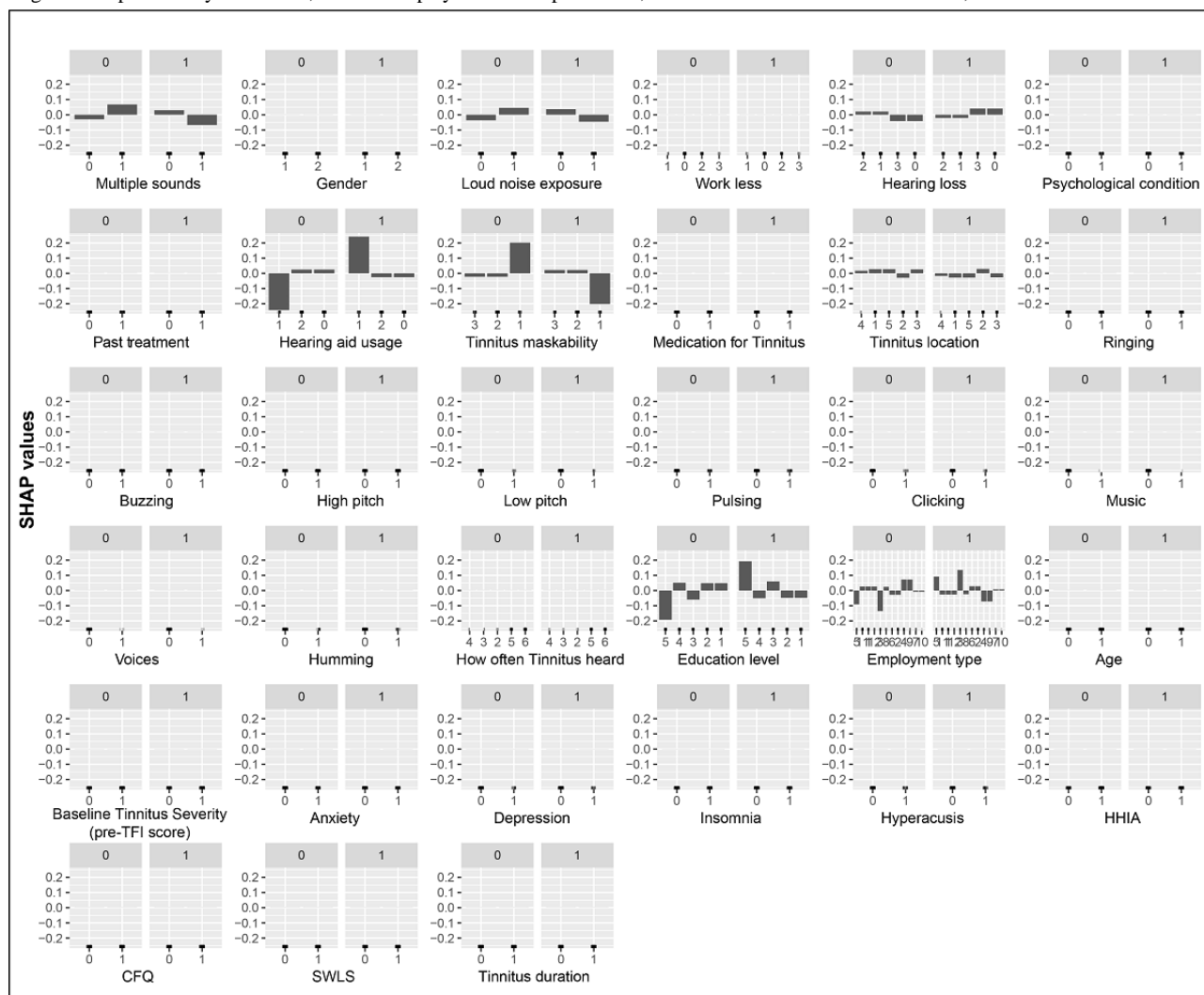
Figures 2 and 3 present the effect that each feature category had on the outcome variable, as decided by the best GB and CART decision tree models. Features' impacts on the two classes are presented in 2 separate plots for each feature ("1" indicates the effect on the successful treatment class and "0" indicates the effect on the unsuccessful treatment class). Positive SHAP values in each successful treatment group indicate a higher log odd of achieving a 13-point or more tinnitus severity score reduction (ie, on the TFI for a given category) for a feature and vice versa. This log odd is relative to the training set average.

Figures 2 and 3 both depict positive SHAP values for the participants who had a vocational training degree or a master's degree (or above), those with higher levels of education (postgraduate degrees), and those who were using a hearing aid in only 1 ear. As per the GB model, this reduction was more likely to occur for participants with insomnia (scores of 14 or less on the Insomnia Severity Index), a psychological condition, and tinnitus that can be described as a buzzing sound and for participants who had median baseline tinnitus severity scores (ie, pre-TFI assessment) of >55.2.

**Figure 2.** The best gradient boosting model–based feature effects. Each graph represents a feature and their corresponding SHAP values. Plots labeled with “1” illustrate the impact that each feature has on achieving a successful treatment outcome (a 13-point or more reduction in TFI score). CFQ: Cognitive Failures Questionnaire; HHIA: Hearing Handicap Inventory for Adults; SHAP: Shapley additive explanations; SWLS: Satisfaction with Life Scales; TFI: Tinnitus Functional Index.



**Figure 3.** The best classification and regression tree decision tree model feature effects. Each graph represents a feature vs. corresponding SHAP value. Plots with “1” represent the effect that each feature has on achieving a successful treatment outcome. CFQ: Cognitive Failures Questionnaire; HHIA: Hearing Handicap Inventory for Adults; SHAP: Shapley additive explanations; SWLS: Satisfaction with Life Scales; TFI: Tinnitus Functional Index.



### Identification of Participants Who Are Likely to Benefit From ICBT

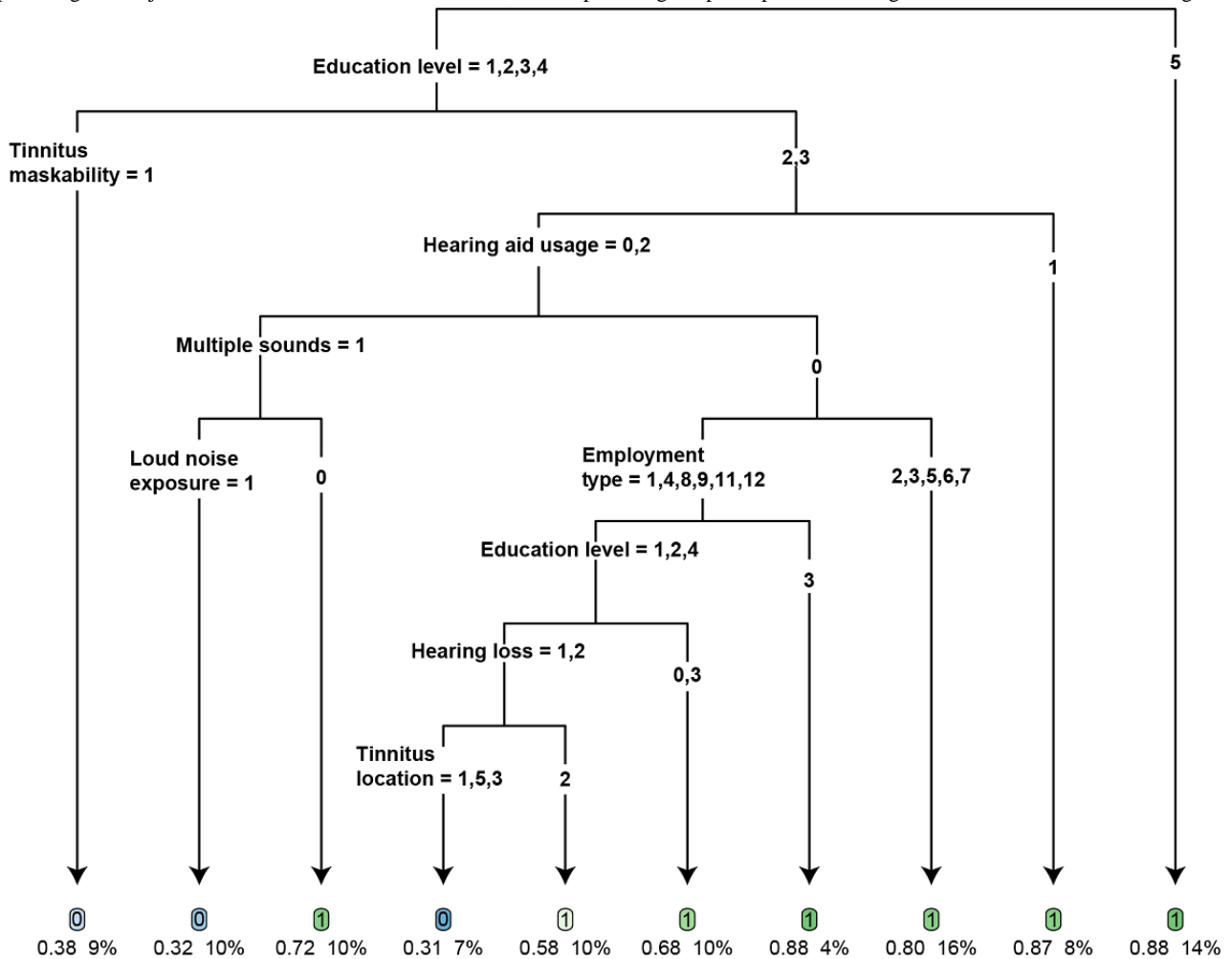
Figure 4 presents the final decision tree model, which has 10 nodes. A detailed explanation is provided in Multimedia Appendix 3. In this model, homogenous groups were formed by creating binary splits at each node. The decision nodes represent the treatment groups (either 0 or 1) that were the most likely to achieve tinnitus severity reduction, and the decisions in the tree were based on the characteristics of each group. These characteristics were represented by each branch (based on their corresponding feature values) of the tree. This model showed that higher education level, tinnitus maskability, hearing aid usage, the presence of multiple tinnitus sounds, loud noise exposure, employment type, the presence of hearing loss, and tinnitus location were important factors for determining

treatment outcomes. The following participant groups had at least an 85% chance of achieving a TFI score reduction of 13 points or more:

- Participants with postgraduate education (master’s degree or higher)
- Participants with an education level other than a master’s level of education and poor tinnitus maskability (or only partial tinnitus maskability) and those who wore a hearing aid in 1 ear
- Participants with no tinnitus maskability or only partial tinnitus maskability; those who did not wear a hearing aid or used hearing aids bilaterally; those who did not hear multiple tinnitus sounds; and those with an occupation that could be described as *professional, technical, skills based* (ie, skilled tradesman), *service related*, or *medical*



**Figure 4.** The best classification and regression tree decision tree model. The fitted tree has 10 terminal nodes (denote the decision criteria). The variable categories (please refer to Multimedia Appendix 2 for variable category labels) corresponding to each split are given at the top of each branch. Each terminal node contains the predicted treatment class ("1" indicates the successful treatment class and "0" indicates the unsuccessful treatment class), the percentage of subjects with successful treatment outcomes, and the percentage of participants with the given characteristics in the training set.



## Discussion

The aim of this study was to explore predictors of outcomes in ICBT for tinnitus by applying 6 types of decision tree models to a combined data set of participants from 3 clinical trials. The key findings are discussed below.

### Best Decision Tree Models

In this study, we applied 6 different decision tree models to ICBT data. Although none of the six models attained an excellent or outstanding status, the CART and GB models' discriminative power can be considered to be satisfactory, given the moderate sample size with just 33 predictive factors (features). This is consistent with a recent study that used 10 decision tree models to predict the outcomes of CBT for patients with tinnitus (N=1416) and found that the GB model with 26 predictive factors had the best predictive power; it had an AUC value of 0.89 [21].

Further work however is needed in this area to determine which models and how many vital factors may result in an optimal predictive model. A larger sample size would likely improve the results. However, we are not sure if just adding factors would be helpful. This is because Niemann et al [21] included 205

factors in their analysis, and of these, only 26 were helpful in achieving optimal results. Moreover, of the 26 factors, only a handful had the most considerable effect. For instance, only 1 factor (ie, tinnitus impairment in terms of loudness, frequency, and distress) resulted in an AUC of 0.79, only 3 features resulted in an AUC of 0.85, and only 8 factors resulted in an AUC of 0.85. These results indicate that including key factors with high predictive power may be a better approach than just adding all of the possible factors.

### Predictors of ICBT Outcomes

Among the best decision tree models (ie, the CART and GB models), various factors were found to be critical predictors of ICBT outcomes. These included demographics (ie, education level, employment type, and the presence of a psychological condition), tinnitus and hearing-related factors (ie, baseline tinnitus severity, tinnitus location, how often tinnitus is heard, a buzzing type of tinnitus, tinnitus maskability, and hearing loss type), treatment-related factors (ie, hearing aid usage), and clinical factors (ie, insomnia). However, education level was the most notable predictor among these.

Participants who had a master's degree or above had an 88% chance of achieving a successful outcome. This is

understandable, as the ability to read, understand, and follow instructions is key for undergoing self-help interventions. However, it is likely that the way in which the materials were written may also have played a role. For instance, UK ICBT materials are written at a ninth-grade reading level [49], which may require higher literacy skills. However, these materials have been rewritten at a sixth-grade reading level and below [49] to ensure accessibility for those with lower education levels. More and more people, including those with lower education, are using the internet and participating in internet-based treatments, particularly due to the constraints that have been placed on health care during the COVID-19 pandemic [50-52]. In our sample, over 85% (198/228, 86.8%) of the participants had a below-master-level education. This highlights the need for making ICBT more accessible to increase the chances of achieving improved outcomes for those with lower education levels.

Baseline tinnitus severity was found to be another critical factor for predicting ICBT outcomes in the GB model. Our previous studies on 1-year outcomes [53] and our previous application of univariate and multivariate analyses to this study's sample [19] identified baseline tinnitus severity as a critical predictive variable. In the Niemann et al [21] study, tinnitus loudness, frequency, and distress, which were measured by using a visual analog scale, were found to be key predictive factors. Further, tinnitus distress, which was measured by using the German version of the Tinnitus Questionnaire (a tool that is comparable to the TFI in this study), was not found to be the key predictive factor. However, based on both clinical experiences and findings from many previous studies, baseline tinnitus severity is an important factor for determining treatment outcomes. The clinical factors depression and anxiety were among the key predictive factors in the GB model. A recent clinical trial conducted by Beukes et al [54], as well as the Niemann et al [21] study, determined that those with high levels of depression had a better chance of achieving success.

Although various other tinnitus and hearing-related variables could have played a role in determining the outcomes of ICBT, the predictive power of our models was relatively low based on the AUC values. Nevertheless, it would be useful for hearing health care professionals to examine these factors when deciding on the candidacy of self-help psychological interventions such as ICBT. Moreover, it would be useful for future studies to examine any additional factors (eg, health literacy) that may have a bearing on ICBT outcomes.

### Study Limitations and Future Directions

Although this study is among the first to apply data mining models to ICBT data, it has several limitations. The sample size was limited, and this may have contributed to the low predictive accuracies of the models. The exploratory decision tree models worked better when including a large number of predictive factors. In this study, we only included 33 predictive factors in our models, and this may have limited the performance of our models. Further, we may have missed some important factors (eg, health literacy) that have a bearing on ICBT outcomes.

The inclusion and exclusion criteria that were used in the three trials from which our data were generated may have resulted in a sample with high tinnitus severity levels that may not be representative of the general tinnitus population. This may have also contributed to our limited key findings. Future studies could include more extensive samples of heterogeneous patients with tinnitus as well as all of the possible predictive factors that could help with improving our models' predictive power. Moreover, developing nonlinear classifiers with artificial neural networks and support vector machines could help with achieving higher prediction accuracies and should be examined in future studies.

In conclusion, tree models, especially the CART and GB models, appear to be promising in predicting ICBT outcomes. Future studies should be undertaken with larger sample sizes and include a more comprehensive range of predictive factors to improve their models' predictive power.

---

### Conflicts of Interest

None declared.

---

### Multimedia Appendix 1

Characteristics of the study's participants.

[\[DOCX File , 28 KB-Multimedia Appendix 1\]](#)

---

### Multimedia Appendix 2

Predictor variables.

[\[DOCX File , 29 KB-Multimedia Appendix 2\]](#)

---

### Multimedia Appendix 3

Detailed explanation of classification and regression tree model.

[\[DOCX File , 25 KB-Multimedia Appendix 3\]](#)

---

### References

1. McCormack A, Edmondson-Jones M, Somerset S, Hall D. A systematic review of the reporting of tinnitus prevalence and severity. *Hear Res* 2016 Jul;337:70-79 [[FREE Full text](#)] [doi: [10.1016/j.heares.2016.05.009](https://doi.org/10.1016/j.heares.2016.05.009)] [Medline: [27246985](https://pubmed.ncbi.nlm.nih.gov/27246985/)]

2. Kostev K, Alymova S, Kössl M, Jacob L. Risk factors for tinnitus in 37,692 patients followed in general practices in Germany. *Otol neurotol* 2019 Apr;40(4):436-440. [doi: [10.1097/mao.0000000000002161](https://doi.org/10.1097/mao.0000000000002161)] [Medline: [30870350](https://pubmed.ncbi.nlm.nih.gov/30870350/)]
3. Kim HJ, Lee HJ, An SY, Sim S, Park B, Kim SW, et al. Analysis of the prevalence and associated risk factors of tinnitus in adults. *PLoS One* 2015 May 28;10(5):e0127578 [FREE Full text] [doi: [10.1371/journal.pone.0127578](https://doi.org/10.1371/journal.pone.0127578)] [Medline: [26020239](https://pubmed.ncbi.nlm.nih.gov/26020239/)]
4. Beukes EW, Manchaiah V, Allen PM, Andersson G, Baguley DM. Exploring tinnitus heterogeneity. *Prog Brain Res* 2021;260:79-99. [doi: [10.1016/bs.pbr.2020.05.022](https://doi.org/10.1016/bs.pbr.2020.05.022)] [Medline: [33637233](https://pubmed.ncbi.nlm.nih.gov/33637233/)]
5. Manchaiah V, Beukes EW, Granberg S, Durisala N, Baguley DM, Allen PM, et al. Problems and life effects experienced by tinnitus research study volunteers: An exploratory study using the ICF classification. *J Am Acad Audiol* 2018;29(10):936-947. [doi: [10.3766/jaaa.17094](https://doi.org/10.3766/jaaa.17094)] [Medline: [30479266](https://pubmed.ncbi.nlm.nih.gov/30479266/)]
6. Beukes EW, Manchaiah V, Andersson G, Allen PM, Terlizzi PM, Baguley DM. Situationally influenced tinnitus coping strategies: a mixed methods approach. *Disabil Rehabil* 2018 Dec;40(24):2884-2894. [doi: [10.1080/09638288.2017.1362708](https://doi.org/10.1080/09638288.2017.1362708)] [Medline: [28793790](https://pubmed.ncbi.nlm.nih.gov/28793790/)]
7. Salazar JW, Meisel K, Smith ER, Quiggle A, McCoy DB, Amans MR. Depression in patients with tinnitus: A systematic review. *Otolaryngol Head Neck Surg* 2019 Jul;161(1):28-35. [doi: [10.1177/0194599819835178](https://doi.org/10.1177/0194599819835178)] [Medline: [30909841](https://pubmed.ncbi.nlm.nih.gov/30909841/)]
8. Trevis KJ, McLachlan NM, Wilson SJ. A systematic review and meta-analysis of psychological functioning in chronic tinnitus. *Clin Psychol Rev* 2018 Mar;60:62-86. [doi: [10.1016/j.cpr.2017.12.006](https://doi.org/10.1016/j.cpr.2017.12.006)] [Medline: [29366511](https://pubmed.ncbi.nlm.nih.gov/29366511/)]
9. Fuller T, Cima R, Langguth B, Mazurek B, Vlaeyen JW, Hoare DJ. Cognitive behavioural therapy for tinnitus. *Cochrane Database Syst Rev* 2020 Jan 08;1(1):CD012614 [FREE Full text] [doi: [10.1002/14651858.CD012614.pub2](https://doi.org/10.1002/14651858.CD012614.pub2)] [Medline: [31912887](https://pubmed.ncbi.nlm.nih.gov/31912887/)]
10. Landry EC, Sandoval XCR, Simeone CN, Tidball G, Lea J, Westerberg BD. Systematic review and network meta-analysis of cognitive and/or behavioral therapies (CBT) for tinnitus. *Otol Neurotol* 2020 Feb;41(2):153-166. [doi: [10.1097/MAO.0000000000002472](https://doi.org/10.1097/MAO.0000000000002472)] [Medline: [31743297](https://pubmed.ncbi.nlm.nih.gov/31743297/)]
11. Tunkel DE, Bauer CA, Sun GH, Rosenfeld RM, Chandrasekhar SS, Cunningham Jr ER, et al. Clinical practice guideline: tinnitus executive summary. *Otolaryngol Head Neck Surg* 2014 Oct;151(4):533-541. [doi: [10.1177/0194599814547475](https://doi.org/10.1177/0194599814547475)] [Medline: [25274374](https://pubmed.ncbi.nlm.nih.gov/25274374/)]
12. Andersson G, Strömngren T, Ström L, Lyttkens L. Randomized controlled trial of internet-based cognitive behavior therapy for distress associated with tinnitus. *Psychosom Med* 2002;64(5):810-816. [doi: [10.1097/01.psy.0000031577.42041.f8](https://doi.org/10.1097/01.psy.0000031577.42041.f8)] [Medline: [12271112](https://pubmed.ncbi.nlm.nih.gov/12271112/)]
13. Beukes EW, Vlaescu G, Manchaiah V, Baguley DM, Allen PM, Kaldo V, et al. Development and technical functionality of an internet-based intervention for tinnitus in the UK. *Internet Interv* 2016 Aug 14;6:6-15 [FREE Full text] [doi: [10.1016/j.invent.2016.08.002](https://doi.org/10.1016/j.invent.2016.08.002)] [Medline: [30135809](https://pubmed.ncbi.nlm.nih.gov/30135809/)]
14. Beukes EW, Manchaiah V, Allen PM, Baguley DM, Andersson G. Internet-based interventions for adults with hearing loss, tinnitus, and vestibular disorders: A systematic review and meta-analysis. *Trends Hear* 2019;23:2331216519851749 [FREE Full text] [doi: [10.1177/2331216519851749](https://doi.org/10.1177/2331216519851749)] [Medline: [31328660](https://pubmed.ncbi.nlm.nih.gov/31328660/)]
15. Beukes EW, Andersson G, Fagelson M, Manchaiah V. Audiologist-supported internet-based cognitive behavioral therapy for tinnitus in the United States: A pilot trial. *Am J Audiol* 2021 Sep 10;30(3):717-729. [doi: [10.1044/2021\\_AJA-20-00222](https://doi.org/10.1044/2021_AJA-20-00222)] [Medline: [34432984](https://pubmed.ncbi.nlm.nih.gov/34432984/)]
16. Beukes E, Andersson G, Fagelson M, Manchaiah V. Internet-based audiologist-guided cognitive behavioral therapy tinnitus in the United States: A randomized controlled trial. *JMIR Preprints*. Preprint posted online on January 30, 2021 [FREE Full text] [doi: [10.2196/preprints.27584](https://doi.org/10.2196/preprints.27584)]
17. Beukes EW, Baguley DM, Allen PM, Manchaiah V, Andersson G. Audiologist-guided internet-based cognitive behavior therapy for adults with tinnitus in the United Kingdom: A randomized controlled trial. *Ear Hear* 2018;39(3):423-433. [doi: [10.1097/AUD.0000000000000505](https://doi.org/10.1097/AUD.0000000000000505)] [Medline: [29095725](https://pubmed.ncbi.nlm.nih.gov/29095725/)]
18. Beukes EW, Andersson G, Allen PM, Manchaiah V, Baguley DM. Effectiveness of guided internet-based cognitive behavioral therapy vs face-to-face clinical care for treatment of tinnitus: A randomized clinical trial. *JAMA Otolaryngol Head Neck Surg* 2018 Dec 01;144(12):1126-1133 [FREE Full text] [doi: [10.1001/jamaoto.2018.2238](https://doi.org/10.1001/jamaoto.2018.2238)] [Medline: [30286238](https://pubmed.ncbi.nlm.nih.gov/30286238/)]
19. Rodrigo H, Beukes EW, Andersson G, Manchaiah V. Internet-based cognitive-behavioural therapy for tinnitus: secondary analysis to examine predictors of outcomes. *BMJ Open* 2021 Aug 20;11(8):e049384 [FREE Full text] [doi: [10.1136/bmjopen-2021-049384](https://doi.org/10.1136/bmjopen-2021-049384)] [Medline: [34417217](https://pubmed.ncbi.nlm.nih.gov/34417217/)]
20. Juhola M, Viikki K, Laurikkala J, Auramo Y, Kentala E, Pyykkö I. Application of artificial intelligence in audiology. *Scand Audiol Suppl* 2001(52):97-99. [doi: [10.1080/010503901300007209](https://doi.org/10.1080/010503901300007209)] [Medline: [11318498](https://pubmed.ncbi.nlm.nih.gov/11318498/)]
21. Niemann U, Boecking B, Brueggemann P, Mebus W, Mazurek B, Spiliopoulou M. Tinnitus-related distress after multimodal treatment can be characterized using a key subset of baseline variables. *PLoS One* 2020 Jan 30;15(1):e0228037. [doi: [10.1371/journal.pone.0228037](https://doi.org/10.1371/journal.pone.0228037)] [Medline: [31999776](https://pubmed.ncbi.nlm.nih.gov/31999776/)]
22. Rasku J, Pyykkö I, Levo H, Kentala E, Manchaiah V. Disease profiling for computerized peer support of Ménière's disease. *JMIR Rehabil Assist Technol* 2015 Sep 03;2(2):e9 [FREE Full text] [doi: [10.2196/rehab.4109](https://doi.org/10.2196/rehab.4109)] [Medline: [28582248](https://pubmed.ncbi.nlm.nih.gov/28582248/)]

23. Shoushtarian M, Alizadehsani R, Khosravi A, Acevedo N, McKay CM, Nahavandi S, et al. Objective measurement of tinnitus using functional near-infrared spectroscopy and machine learning. *PLoS One* 2020 Nov 18;15(11):e0241695. [doi: [10.1371/journal.pone.0241695](https://doi.org/10.1371/journal.pone.0241695)] [Medline: [33206675](https://pubmed.ncbi.nlm.nih.gov/33206675/)]
24. Weininger O, Warnecke A, Lesinski-Schiedat A, Lenarz T, Stolle S. Computational analysis based on audioprofiles: A new possibility for patient stratification in office-based otology. *Audiol Res* 2019 Nov 05;9(2):230 [FREE Full text] [doi: [10.4081/audiores.2019.230](https://doi.org/10.4081/audiores.2019.230)] [Medline: [31728177](https://pubmed.ncbi.nlm.nih.gov/31728177/)]
25. Raj S. Effects of multi-collinearity in logistic regression, SVM, random forest(RF). Medium. 2019 Aug 09. URL: <https://medium.com/@raj5287/effects-of-multi-collinearity-in-logistic-regression-svm-rf-af6766d91f1b> [accessed 2021-09-22]
26. Beukes EW, Allen PM, Manchaiah V, Baguley DM, Andersson G. Internet-based intervention for tinnitus: Outcome of a single-group open trial. *J Am Acad Audiol* 2017 Apr;28(4):340-351. [doi: [10.3766/jaaa.16055](https://doi.org/10.3766/jaaa.16055)] [Medline: [28418328](https://pubmed.ncbi.nlm.nih.gov/28418328/)]
27. Beukes EW, Andersson G, Manchaiah V, Kaldo V. *Cognitive Behavioral Therapy for Tinnitus*. USA: Plural Publishing; Oct 28, 2020.
28. Manchaiah V, Vlaescu G, Varadaraj S, Aronson EP, Fagelson MA, Munoz MF, et al. Features, functionality, and acceptability of internet-based cognitive behavioral therapy for tinnitus in the United States. *Am J Audiol* 2020 Sep 03;29(3):476-490 [FREE Full text] [doi: [10.1044/2020\\_AJA-20-00002](https://doi.org/10.1044/2020_AJA-20-00002)] [Medline: [32880499](https://pubmed.ncbi.nlm.nih.gov/32880499/)]
29. Vlaescu G, Alasjö A, Miloff A, Carlbring P, Andersson G. Features and functionality of the Iterapi platform for internet-based psychological treatment. *Internet Interv* 2016 Oct 03;6:107-114 [FREE Full text] [doi: [10.1016/j.invent.2016.09.006](https://doi.org/10.1016/j.invent.2016.09.006)] [Medline: [30135819](https://pubmed.ncbi.nlm.nih.gov/30135819/)]
30. Meikle M, Henry J, Griest S, Stewart B, Abrams H, McArdle R, et al. The tinnitus functional index: development of a new clinical measure for chronic, intrusive tinnitus. *Ear Hear* 2012;33(2):153-176. [doi: [10.1097/AUD.0b013e31822f67c0](https://doi.org/10.1097/AUD.0b013e31822f67c0)] [Medline: [22156949](https://pubmed.ncbi.nlm.nih.gov/22156949/)]
31. Bastien CH, Vallières A, Morin CM. Validation of the Insomnia Severity Index as an outcome measure for insomnia research. *Sleep Med* 2001 Jul;2(4):297-307. [doi: [10.1016/s1389-9457\(00\)00065-4](https://doi.org/10.1016/s1389-9457(00)00065-4)] [Medline: [11438246](https://pubmed.ncbi.nlm.nih.gov/11438246/)]
32. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med* 2006 May 22;166(10):1092-1097. [doi: [10.1001/archinte.166.10.1092](https://doi.org/10.1001/archinte.166.10.1092)] [Medline: [16717171](https://pubmed.ncbi.nlm.nih.gov/16717171/)]
33. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. *Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire*. *JAMA* 1999 Nov 10;282(18):1737-1744. [doi: [10.1001/jama.282.18.1737](https://doi.org/10.1001/jama.282.18.1737)] [Medline: [10568646](https://pubmed.ncbi.nlm.nih.gov/10568646/)]
34. Newman CW, Weinstein BE, Jacobson GP, Hug GA. Test-retest reliability of the hearing handicap inventory for adults. *Ear Hear* 1991 Oct;12(5):355-357. [doi: [10.1097/00003446-199110000-00009](https://doi.org/10.1097/00003446-199110000-00009)] [Medline: [1783240](https://pubmed.ncbi.nlm.nih.gov/1783240/)]
35. Khalfa S, Dubal S, Veuillet E, Perez-Diaz F, Jouvent R, Collet L. Psychometric normalization of a hyperacusis questionnaire. *ORL J Otorhinolaryngol Relat Spec* 2002;64(6):436-442. [doi: [10.1159/000067570](https://doi.org/10.1159/000067570)] [Medline: [12499770](https://pubmed.ncbi.nlm.nih.gov/12499770/)]
36. Broadbent DE, Cooper PF, FitzGerald P, Parkes KR. The Cognitive Failures Questionnaire (CFQ) and its correlates. *Br J Clin Psychol* 1982 Feb;21(1):1-16. [doi: [10.1111/j.2044-8260.1982.tb01421.x](https://doi.org/10.1111/j.2044-8260.1982.tb01421.x)] [Medline: [7126941](https://pubmed.ncbi.nlm.nih.gov/7126941/)]
37. Diener E, Emmons RA, Larsen RJ, Griffin S. The Satisfaction With Life Scale. *J Pers Assess* 1985 Feb;49(1):71-75. [doi: [10.1207/s15327752jpa4901\\_13](https://doi.org/10.1207/s15327752jpa4901_13)] [Medline: [16367493](https://pubmed.ncbi.nlm.nih.gov/16367493/)]
38. James G, Witten D, Hastie T, Tibshirani R. *An Introduction to Statistical Learning with Applications in R*. USA: Springer; 2013.
39. Breiman L. Random forests. *Machine learning* 2001 Oct;45:5-32. [doi: [10.1023/A:1010933404324](https://doi.org/10.1023/A:1010933404324)]
40. Friedman JH. Greedy function approximation: A gradient boosting machine. *Ann Stat* 2001 Oct;29(5):1189-1232. [doi: [10.1214/aos/1013203451](https://doi.org/10.1214/aos/1013203451)]
41. Chen T, Guestrin C. XGBoost: A scalable tree boosting system. 2016 Aug Presented at: KDD '16: The 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; August 13-17, 2016; San Francisco, California, USA p. 785-794. [doi: [10.1145/2939672.2939785](https://doi.org/10.1145/2939672.2939785)]
42. Breiman L, Friedman JH, Olshen RA, Stone CJ. *Classification and Regression Trees*. Boca Raton, Florida, USA: Chapman and Hall/CRC; 1984.
43. Quinlan J. *C4.5: Programs for Machine Learning*. San Francisco, California, USA: Morgan Kaufmann Publishers; Mar 1993.
44. Gandhi R. Gradient boosting and XGBoost. Hacker Noon. 2018 May 06. URL: <https://hackernoon.com/gradient-boosting-and-xgboost-90862daa6c77> [accessed 2021-09-22]
45. Gini C. On the Measure of Concentration with Special Reference to Income and Statistics. *Colorado College Publication* 1936;208:73-79.
46. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. 2017 Dec Presented at: The 31st International Conference on Neural Information Processing Systems; December 4-9, 2017; Long Beach, California, USA p. 4768-4777.
47. Lundberg SM, Erion G, Chen H, DeGrave A, Prutkin JM, Nair B, et al. From local explanations to global understanding with explainable AI for trees. *Nat Mach Intell* 2020 Jan;2(1):56-67 [FREE Full text] [doi: [10.1038/s42256-019-0138-9](https://doi.org/10.1038/s42256-019-0138-9)] [Medline: [32607472](https://pubmed.ncbi.nlm.nih.gov/32607472/)]
48. Hansapani / AI-Tinnitus. GitHub. URL: <https://github.com/Hansapani/AI-Tinnitus> [accessed 2021-09-23]

49. Beukes EW, Fagelson M, Aronson EP, Munoz MF, Andersson G, Manchaiah V. Readability following cultural and linguistic adaptations of an internet-based intervention for tinnitus for use in the United States. *Am J Audiol* 2020 Jun 08;29(2):97-109 [FREE Full text] [doi: [10.1044/2019\\_AJA-19-00014](https://doi.org/10.1044/2019_AJA-19-00014)] [Medline: [32105512](https://pubmed.ncbi.nlm.nih.gov/32105512/)]
50. Beukes EW, Baguley DM, Jacquemin L, Lourenco MPCG, Allen PM, Onozuka J, et al. Changes in tinnitus experiences during the COVID-19 pandemic. *Front Public Health* 2020 Nov 05;8:592878 [FREE Full text] [doi: [10.3389/fpubh.2020.592878](https://doi.org/10.3389/fpubh.2020.592878)] [Medline: [33251179](https://pubmed.ncbi.nlm.nih.gov/33251179/)]
51. Beukes EW, Onozuka J, Brazell TP, Manchaiah V. Coping with tinnitus during the COVID-19 pandemic. *Am J Audiol* 2021 Jun 14;30(2):385-393. [doi: [10.1044/2021\\_AJA-20-00188](https://doi.org/10.1044/2021_AJA-20-00188)] [Medline: [33979227](https://pubmed.ncbi.nlm.nih.gov/33979227/)]
52. Wind TR, Rijkeboer M, Andersson G, Riper H. The COVID-19 pandemic: The 'black swan' for mental health care and a turning point for e-health. *Internet Interv* 2020 Apr;20:100317 [FREE Full text] [doi: [10.1016/j.invent.2020.100317](https://doi.org/10.1016/j.invent.2020.100317)] [Medline: [32289019](https://pubmed.ncbi.nlm.nih.gov/32289019/)]
53. Beukes EW, Allen PM, Baguley DM, Manchaiah V, Andersson G. Long-term efficacy of audiologist-guided internet-based cognitive behavior therapy for tinnitus. *Am J Audiol* 2018 Nov 19;27(3S):431-447 [FREE Full text] [doi: [10.1044/2018\\_AJA-IMIA3-18-0004](https://doi.org/10.1044/2018_AJA-IMIA3-18-0004)] [Medline: [30452747](https://pubmed.ncbi.nlm.nih.gov/30452747/)]
54. Beukes EW, Andersson G, Fagelson MA, Manchaiah V. Dismantling internet-based cognitive behavioral therapy for tinnitus. The contribution of applied relaxation: A randomized controlled trial. *Internet Interv* 2021 May 12;25:100402 [FREE Full text] [doi: [10.1016/j.invent.2021.100402](https://doi.org/10.1016/j.invent.2021.100402)] [Medline: [34040997](https://pubmed.ncbi.nlm.nih.gov/34040997/)]

## Abbreviations

**AUC:** area under the receiver operating characteristic curve  
**CART:** classification and regression tree  
**CBT:** cognitive behavioral therapy  
**GB:** gradient boosting  
**ICBT:** internet-based cognitive behavioral therapy  
**RF:** random forest  
**SHAP:** Shapley additive explanations  
**TFI:** Tinnitus Functional Index  
**XGBoost:** extreme gradient boosting

*Edited by R Kukařka; submitted 21.03.21; peer-reviewed by G Liu, JA Benítez-Andrades; comments to author 28.06.21; revised version received 29.08.21; accepted 12.09.21; published 02.11.21*

*Please cite as:*

Rodrigo H, Beukes EW, Andersson G, Manchaiah V

*Exploratory Data Mining Techniques (Decision Tree Models) for Examining the Impact of Internet-Based Cognitive Behavioral Therapy for Tinnitus: Machine Learning Approach*

*J Med Internet Res* 2021;23(11):e28999

URL: <https://www.jmir.org/2021/11/e28999>

doi: [10.2196/28999](https://doi.org/10.2196/28999)

PMID:

©Hansapani Rodrigo, Eldré W Beukes, Gerhard Andersson, Vinaya Manchaiah. Originally published in the Journal of Medical Internet Research (<https://www.jmir.org/>), 02.11.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://www.jmir.org/>, as well as this copyright and license information must be included.