Predicting the outcomes of internet-based cognitive behavioral therapy for tinnitus: Applications of Artificial Neural Network and Support Vector Machine

Hansapani Rodrigo,[1,3] Eldré W. Beukes,[2,3,4] Gerhard Andersson,[5,6] & Vinaya Manchaiah,[2,3,6,7]

1. School of Mathematical and Statistical Sciences, University of Texas Rio Grande Valley, Edinburg, Texas, USA
2. Department of Speech and Hearing Sciences, Lamar University, Beaumont, Texas, USA
3. Virtual Hearing Lab, Collaborative initiative between Lamar University and University of Pretoria
4. Vision and Hearing Sciences Research Centre, School of Psychology and Sport Science, Anglia Ruskin University, Cambridge, United Kingdom
5. Department of Behavioral Sciences and Learning, Department of Biomedical and Clinical Sciences, Linköping University, Linköping, Sweden
6. Department of Clinical Neuroscience, Division of Psychiatry, Karolinska Institute, Stockholm, Sweden
7. Department of Speech and Hearing, School of Allied Health Sciences, Manipal, Karnataka, India

Corresponding Author

Author:	Hansapani Rodrigo, PhD
Communication address:	School of Mathematical and Statistical Sciences, University of Texas Rio Grande Valley, 1201 W University Drive, Edinburg, TX, 78539, USA
Email:	hansapani.rodrigo@utrgv.edu
Tel:	+1 (956) 665 2349

Conflicts of Interest

None declared.

Abstract

Purpose: Internet-based cognitive behavioral therapy (ICBT) has been found to be effective for tinnitus management, although there is limited understanding about who will benefit the most from ICBT. Traditional statistical models have largely failed to identify the non-linear associations and hence find strong predictors of success with ICBT. The current study aimed at examining the use of an artificial neural network (ANN) and support vector machine (SVM) to identify variables associated with treatment success in ICBT for tinnitus.

Method: The study involved a secondary analysis of data from 228 individuals who had completed ICBT in previous intervention studies. A 13-point reduction in Tinnitus Functional Index (TFI) was defined as a successful outcome. There were 33 predictor variables, including demographic, tinnitus, hearing-related, and treatment-related variables and clinical factors (anxiety, depression, insomnia, hyperacusis, hearing disability, cognitive function, and life satisfaction). Predictive models using ANN and SVM were developed and evaluated for classification accuracy. SHapley Additive exPlanations (SHAP) analysis was used to identify the relative predictor variable importance using the best predictive model for a successful treatment outcome.

Results: The best predictive model was achieved with the ANN with an average area under the receiver operating characteristic (AUC) value of 0.73 ± 0.03. The SHAP analysis revealed that having a higher education level and a greater baseline tinnitus severity were the most critical factors that influence treatment outcome positively.

Conclusions: Predictive models such as ANN and SVM help predict ICBT treatment outcomes and identify predictors of outcome. However, further work is needed to examine predictors that were not considered in this study as well as to improve the predictive power of these models.

Registration: Clinicaltrials.gov NCT02370810; Clinicaltrials.gov NCT02665975.

Key Words

Tinnitus, Internet interventions, Digital therapeutics, Cognitive behavioral therapy, Artificial intelligence, Machine learning, Artificial neural networks (ANN), Support vector machines (SVM)

Introduction

Tinnitus is a prevalent hearing-related condition that affects 10-15% of the adult population (McCormack et al., 2016). It is a highly heterogeneous condition in the way it is perceived and how individuals react to it. While tinnitus does not have an effect on the daily life of most people afflicted, a significant proportion report that tinnitus severely affects the performance of their essential day-to-day tasks (Beukes et al., 2021). The common problems and life effects reported by individuals with tinnitus may include difficulties in emotional function, sleep, hearing, reduced energy levels, problems in handling stress, and social problems, including work and family (Beukes et al., 2018a; Manchaiah et al., 2018). Tinnitus is also associated with various clinical conditions such as hearing loss, anxiety, and depression. While there is no cure for tinnitus, there are various management strategies with varied empirical support (Tunkel et al., 2014).

According to most reviews and treatment guidelines, Cognitive behavioral therapy (CBT) based on psychological principles has the most robust research support for tinnitus management (Fuller et al., 2020; Landry et al., 2020). CBT is based on the basic principle that what we think, how we feel, and how we behave are all closely connected, and each of these factors well-being. Due to this interconnectedness, addressing either unhelpful thoughts, emotions, reactions, or behaviors can lead to improvements associated with numerous disorders. CBT principles have been tailored as effective interventions for many different disorders including anxiety (Axelesson et al., 2019), depression (López-López et al., 2019) and insomnia (van der Zweerde et al., 2019), all which are often associated with tinnitus.

However, CBT is not easily accessible due to the limited number of trained professionals who can offer CBT for tinnitus (Bhatt et al., 2016). To increase the accessibility and affordability of CBT for tinnitus, an internet-based CBT (ICBT) was developed as a self-help format of CBT with minimal guidance from therapists. Numerous advantages of this approach have expanded its applicability, including its accessibility and flexibility and being convenient and informative (Beukes et al., 2018b). It furthermore requires less resources, having less associated stigma, partly due to maintaining anonymity and privacy and providing a standardized-evidence based treatment (Barak et al., 2008). As it is easy transferable, translatable and cultural adaptably, it has potential to reach global communities. (Warmerdam et al., 2010)

As its efficacy and effectiveness have been demonstrated of other conditions associated with tinnitus e.g. insomnia (Yu et al. (2019), it was later also developed for tinnitus (Andersson et al., 2002). Several controlled trials across the globe have shown positive outcomes of ICBT for tinnitus (for review, see Beukes et al., 2019)., but there is limited understanding of who will benefit from ICBT.

To identify for who ICBT is more suitable, studies have examined predictors of ICBT outcomes. Lower engagement has been found to be one of the main barriers of ICBT, estimated in about a quarter of all ICBT patients. Predictors of non-compliance are related to greater severity and higher levels of associated anxiety.

Factors including higher engagement and an increase in knowledge have led to lager reductions in severity for ICBT for insomnia. (Kraepelien et al. 2021)   Only a few previous studies have examined predictors of ICBT outcomes for tinnitus. Kaldo-Sandström et al. (2004) examined predictors of ICBT outcomes in a clinical population in Sweden. They reported that intervention compliance, external referral to the treatment, and the number of previous tinnitus treatments were associated with positive outcomes. In contrast, the number of messages sent to their therapist concerning the treatment problem was associated with worse outcomes. Further identified trends were that patients referred from external routes and those undertaking previous treatments had better outcomes.

When evaluating the long-term (1-year) outcomes of ICBT in the UK from a self-selected research population, Beukes, et al., 2018c) found that baseline tinnitus severity, engagement with the ICBT program (more modules read), and higher self-reported satisfaction with the intervention were critical predictors of ICBT success. Rodrigo et al. (2021a) found that baseline tinnitus severity and education levels were predictors of ICBT outcome when applying univariate and multivariable models to a self-recruited research population undertaking ICBT in the UK. In a follow-up study, we used various exploratory techniques (Rodrigo et al., 2021b), namely classification and regression trees (CART) (Breiman et al., 1984), C5.0 decision trees (Quinlan, 1993), Gradient Boosting (Friedman, 2001), AdaBoost algorithm (Gandhi, 2008), eXtreme gradient Boosting (Chen, et al. 2016) and Random Forest (Breiman, 2001) to examine the presence of any non-linear associations with the response. The CART decision tree model was identified as the optimal decision tree model. Nevertheless, its predictive power was still considered low (accuracy of 74%, sensitivity of 74%, specificity of 64%, and area under the receiver operating characteristic curve (AUC) of 0.69. This brought the necessity of exploring other non-linear techniques to determine the most valuable predictors of outcome.

Artificial Intelligence and Machine Learning (AI/ML) algorithms have become popular methods to predict outcomes in various rehabilitative practices. AI/MLs have been applied to audiovestibular data since the 1980s (Juhola et al., 2001). Several studies have reported that machine learning techniques such as artificial neural network (ANN) and support vector machine (SVM) have yielded more favorable results than other techniques for audiovestibular data (Haro et al., 2020; Niemann et al., 2020; Wang, 2017; Zhao et al., 2019).

ANN is an information processing archetype inspired by biological neural networks systems, such as the brain (McCulloch et al., 1943). ANNs are typically organized in layers where each layer is made up of a number of interconnected 'neurons' (also known as nodes) which contain an 'activation function'. The purpose of the activation function is to introduce non-linearity into the output of a neuron. This makes ANNs capable of capturing complex relationships intrinsic to the data. SVM (Cortes et al., 1995) is another well-known classifier that falls under supervised learning algorithms that uses the concept of “margin” to classify between two classes. SVM uses different kernels like linear, polynomial, and radial basis to transform the inputs to higher dimensional feature space to assist in this task.

Few studies have applied ML to identify pre-intervention factors that may predict treatment outcomes for tinnitus. Niemann et al. (2020) applied AI techniques to determine treatment predictors for 1,416 patients completing a 7-day multi-modal treatment for tinnitus. The authors used 205 predictor variables, including sociodemographic and clinical factors. As per their findings, gradient boosted trees were identified to be the optimal classifier at predicting tinnitus-related distress post-intervention. The variables, perceived tinnitus-related impairment, depressive symptoms, sleep problems, physical health-related impairments in quality of life, time spent to complete questionnaires, and educational level, exhibited a high contribution towards the model prediction. However, no study has used AI/ML to predict outcomes for ICBT tinnitus treatment other than our recent study (Rodrigo et al., 2021b).

The current study aimed to examine the applications of AI/ML, specifically using ANN and SVM to identify variables associated with positive treatment outcomes in ICBT for tinnitus. First of all, we intend to identify the most appropriate AI/ML technique that can be used to predict treatment distress. This will be followed by identifying the most useful pre-intervention variables that can best predict post-intervention outcomes.

Method

Study Design

The study was a secondary analysis of three different clinical trials that were used to investigate the impact of ICBT within the period from 2016 to 2018, namely a single-group pre-test post-test design (n=42) (Beukes et al., 2017), an efficacy randomized control (RCT) design (n=142) (Clinical Trials.gov: NCT02370810; Beukes, Baugley, et al., 2018d), and an effectiveness RCT design (n=46) (Clinical Trials.gov: NCT02665975; Beukes, Andersson, et al., 2018e). Participants had completed the baseline and post-intervention outcome measures, resulting in a pooled single data set (n = 228).

Ethical clearance was obtained from the Faculty of Science and Technology Research Ethics Panel of Anglia Ruskin University (ARU reference: FST/FREP/14/478 and FST/FREP/14/478) and the East of England–Cambridge South Research Ethics Committee (REC reference: 16/EE/0148) and Health Research Authority (IRAS project ID: 195565). All participants had provided informed consent before participating in the study.

Participant Characteristics

The inclusion criteria were age at least 18 years, ability to read and type in English, access to a computer and the internet and having bothersome tinnitus. The average age was 55.14 years (SD: 12.92), and 43% were female (see S1). Participants had varied educational backgrounds, with 26% having completed high school education, 26% having an undergraduate degree, and 13% having a postgraduate degree. A quarter of the participants had sought tinnitus treatment previously, and the average tinnitus duration was 18 years (SD: 19 years)

Intervention

All participants had completed an 8-week ICBT intervention (Beukes et al., 2016, 2020). The intervention was presented in a self-help format with minimal therapist guidance and was administered using a secured ePlatform (Manchaiah et al., 2020). During this 8-week period, participants were represented with 2-3 learning modules that contained various elements of CBT specifically adapted for tinnitus, including applied relaxation, cognitive restructuring, and imagery. The digital assistive materials were presented using text, images, and videos. In addition, various exercises were presented to improve engagement.

Data Collection

All data were collected using online questionnaires. The study participants completed an extensive pre-intervention questionnaire, including demographics, tinnitus-related and treatment-related information, and other clinically relevant factors. Study participants also completed standardized patient-reported outcome measures before and after the intervention. The primary outcome was a change in tinnitus distress, as measured by the Tinnitus Functional Index (TFI; Meikle et al., 2012). The secondary outcome measures included the Insomnia Severity Index (ISI; Bastien et al., 2001) as a measure of insomnia, the Generalized Anxiety Disorder (GAD-7; Spitzer et al., 2006) as a measure of anxiety, the Patient Health Questionnaire (PHQ-9; Spitzer et al., 1999) as a measure of depressive symptoms, the Hearing Handicap Inventory for Adults Screening version (HHIA-S; Newman et al., 1991) as a measure of self-reported hearing disability, the Hyperacusis Questionnaire (HQ; Khalfa et al., 2002) to assess the presence of hyperacusis (reduced tolerance of everyday sounds), the Cognitive Failures Questionnaire (CFQ; Broadbent et al., 1982) to evaluate cognitive functions, and the Satisfaction with Life Scales (SWLS; Diener et al., 1985) to assess global life satisfaction.

Data Analysis

Variables

The dependent variable in the study was a change in tinnitus distress. A 13-point reduction in TFI scores was defined as a clinically significant change (Meikle et al., 2012), indicating successful treatment outcomes. In addition, there were 33 predictor variables, including seven demographic variables (age, gender, education level, employment type, noise exposure, the presence of psychological conditions, tinnitus affecting the ability to work), 15 tinnitus and hearing-related variables (baseline tinnitus severity, tinnitus duration, how often tinnitus was heard, tinnitus location, nine different types of tinnitus, multiple tones heard, and the presence of hearing loss), four treatment-related variables (past treatment sought, tinnitus maskability [defined as sound enrichment being effective in making tinnitus less noticeable], hearing aid use, and medication use), and seven clinical factors (anxiety, depression, insomnia, hyperacusis, hearing disability, cognitive functions, and life satisfaction). Supplementary Tables S2-S5 provide details about these variables, including the specific questions and response options.

Preliminary analysis was performed to investigate differences in baseline characteristics between those who had shown clinically significant reduction in the TFI score vs. those who had not, using either *t*-test (for quantitative variables) or Chi Square/ Fishers' Exact (for categorical variables) tests along with the odds ratio calculation. All analyses were two-tailed and performed with a 0.05 significance level. R statistical software (V.3.6.3) was used for all analyses, including predictive model development. The code and data are available upon a reasonable request.

Predictive Models

We focused on two types of machine learning models, namely ANN (McCulloch & Pitts, 1943) and SVM (Cortes & Vapnik, 1995). The ANN model used in this study consisted of 3 layers; input, hidden and output. As there were 33 predictor variables (taken as inputs to the ANN) and one outcome (taken as the output of the ANN, indicating the success or failure of the ICBT treatment), 33 input nodes and one output node were used in training the ANN model. A varying number of hidden nodes (from 1 to 5) and different weight decay values were examined during the ANN model training process. The optimal ANN model had five hidden nodes with a weight decay of 0.1. The sigmoid activation function was used at the output layer to obtain the predictions.

SVM employs kernel tricks and maximal margin concepts to perform better in non-linear and high-dimensional tasks. Most of the time, even a powerful SVM model benefits from the proper feature selection and feature extraction/transformation techniques. Two SVM models were used in our study, one with linear kernel and one with Radial Basis Kernel (Chudzian, 2011).

Unlike traditional statistical methods are usually driven by certain distributional assumptions. If the model assumptions are violated, the conclusions made through these will not be valid. Moreover, when evaluating the effect of interactions within these models, those need to be added specifically and tested. Unlike these models, ML models like ANN and SVM do not depend on the distributions of the data and can handle complex interactions within the data (Niu, et al. 2019; Chong et al., 2021) even they not being specifically explored. More importantly, ANN and SVM are data driven models which have shown impressive predictive accuracies over other models (https://towardsdatascience.com/). However, ANN models are often criticized for their black-box nature as their predictions are non-transparent and not traceable by humans (Benitez at al., 1997). To overcome this a model-agnostic post-hoc framework, SHapley Additive exPlanations (SHAP) was applied to facilitate model interpretation and assess the feature importance when identifying the most influential factors leading to TFI score reduction after the treatment (Lundberg et al., 2017, 2020). Moreover, please note that, unlike traditional statistical methods, ANN models do not rely on model assumptions like multicollinearity (De Veaux and

Ungar, 1994). Nevertheless, multicollinearity could affect the performance of SVM in a similar way to that of multicollinearity in logistic regression. Therefore, we have identified the variables which demonstrated severe multicollinearity using the generalized variation inflation factor (GVIF), with a multiple logistic regression model. Variables with VGIF >10 indicate severe multicollinearity. Three variables; employment type (GVIF: 525.10), tinnitus location (GVIF: 15.73) and tinnitus affecting the ability to work (GVIF: 13.78) had shown severe multicollinearity. We ran two additional SVM models (one with linear and other with radial basis kernels) without these three variables, to investigate the impact of their removal on the model performance. Despite these changes, the discrimination power (as assessed by the mean AUC) of SVM models remain lower than to the ANN model. Here, we summarize the predictive model development process step-by-step order to facilitate the readers' understanding.

Model Training

The data set with 228 participants was divided into two parts used for model training (80% of data, n = 183) and testing. During the model training process, a 3-fold cross-validation process was applied. The training data set was split into three folds, and each fold was given a chance to act as its validation set to minimize model overfitting. Ten different replicate models were created with different random initializations for each AI method.

Model Performance Evaluation

The discriminative power of the trained models was assessed using the mean predictive accuracy, sensitivity (true positive rate), specificity (true negative rate), and AUC using the testing data set. These are given as mean $\pm$ SD based on the ten replicated models for each AI technique. The best predictive model was selected for further analysis based on the highest AUC value.

Shapley Additive exPlanations for Predictor Importance

To better understand the contribution of each predictor to the final model predictions, we used SHAP analysis (Lundberg & Su-ln, 2017). SHAP measures the impact of variables taking into account the interaction with other variables. SHAP estimates the importance of a predictor variable by comparing what a model predicts with and without that predictor variable. As the order in which a model sees variables can affect its predictions, every possible order was applied to ensure that the predictors were fairly compared.

For an overall comparison of predictor variables, a matrix of SHAP values was obtained for each participant. This matrix has one row per participant and one column per predictor, where we summed the absolute SHAP values for each predictor variable across the data. Predictors with large absolute SHAP values are considered globally significant. A visualization of the importance of these predictor variables in descending order is provided for the best predictive model.

Results

ICBT Outcome

Tinnitus distress was reduced from a mean of 58/100 points (SD: 19) to 34/100 points (SD: 23) after completing the ICBT intervention as measured by the TFI. This reduction was significant [$t(227) = 16.37$, p<.001]. A clinically meaningful score change (13-point reduction in TFI) was achieved by 150/228 (66%) of the participants. We have performed a series of chi-square tests to evaluate this potential issue having higher TFI reduction mainly among the participants who had shown higher anxiety, higher depression, and higher insomnia (tinnitus related distress factors at baseline). However, no association was found between TFI reduction and anxiety $(OR\ 1.09, 95\ \%\ CI\colon 0.60, 1.98, p = .774)$, between TFI reduction and depression $(OR\ 1.19, 95\ \%\ CI\colon 0.49, 2.88, p = .694)$,, and between TFI reduction and insomnia $(OR\ 0.86, 95\ \%\ CI\colon 0.49, 1.50, p = .587)$. This indicates the fact that the effect of TFI reduction do not just associate with the participants with higher tinnitus related distress measures. Supplementary Table S1 presents the *t*-test or Chi-Square/ Fishers' Exact test results along with the corresponding odds ratios between the groups who had either shown clinically successful reduction or not.

Participants with a Master's degree or above showed the highest odds ratio$(OR\ 4.50, 95\ \%\ CI\colon 1.59, 18.47, p = .003)$ of success with the ICBT treatment, indicating the importance of education. Baseline tinnitus severity was significantly different between the two groups $(\chi_1 = 53.27\ (SD\colon 20.87), \chi_2 = 60.35\ (SD\colon 17.79), t(136.06) = -2.52$, p = .012). Moreover, those for which tinnitus was not possible to mask was significantly different $(OR\ 3.31, 95\%CI\colon 1.02, 10.72, p = .042)$ between the groups. No other variables were identified as significantly different among the two groups. The findings align with the Rodrigo et al., 2021a and Rodrigo et al., 2021b studies which had performed analysis on similar data using additional methods including linear and logistic regressions and decision tree models.

AI/ML Model Performance Evaluations

Results for the three predictive models examined (i.e., ANN, SVM with linear kernels, SVM with radial basis kernels) are provided in Table 1. The ANN model led to the highest mean accuracy of 72.89% (SD: 5.22) with ten random initializations. The mean sensitivity and specificity values were 82.67% (SD: 9.66) and 53.33 (SD: 9.94), respectively. The mean AUC value was 0.73 (SD: 0.03). The AUC is a measure of the discrimination capacity of the model, i.e., the ability of the model to correctly identify those who might have a successful treatment outcome vs. those who do not. A general rule of thumb is that an AUC between 0.90-1.00 indicates excellent discriminative power, an AUC between 0.8-0.9 indicates good discriminatory power and an AUC between $0.70 - 0.80$ indicates an adequate discriminative power. Hence, all three predictive models presented in this paper show sufficient discriminative capacity. In contrast, ANN shows the highest average AUC out of all models considered here and in Rodrigo et al. (2021b).

Table 1: Predictive model evaluations

Classification Model	Accuracy (%; SD)	Sensitivity (true positive rate%; SD)	Specificity (true negative rate%; SD)	Area under the ROC curve (AUC%; SD)
ANN	72.89 ± 5.22	82.67 ± 9.66	53.33 ± 9.94	0.73 ± 0.03
SVM with Linear Kernel	69.12 ± 1.62	72.65 ± 4.40	61.99 ± 8.93	0.72 ± 0.01
SVM with Linear Kernel (without variables VIF>10)	65.78 ± 3.95	84.00 ± 7.67	29.33 ± 25.77	0.68 ± 0.00
SVM with Radial Basis Kernel	67.11 ± 4.54	74.33 ± 7.22	52.66 ± 2.10	0.70 ± 0.00
SVM with Radial Basis Kernel (without variables VIF>10)	66.66 ± 4.45	82.00 ± 5.71	35.99 ± 23.38	0.67 ± 0.01

Note. ROC = receiver operating characteristic; AUC = area under the receiver operating characteristic; ANN = artificial neural network; SVM = support vector machine; VIF = variation inflation factor.

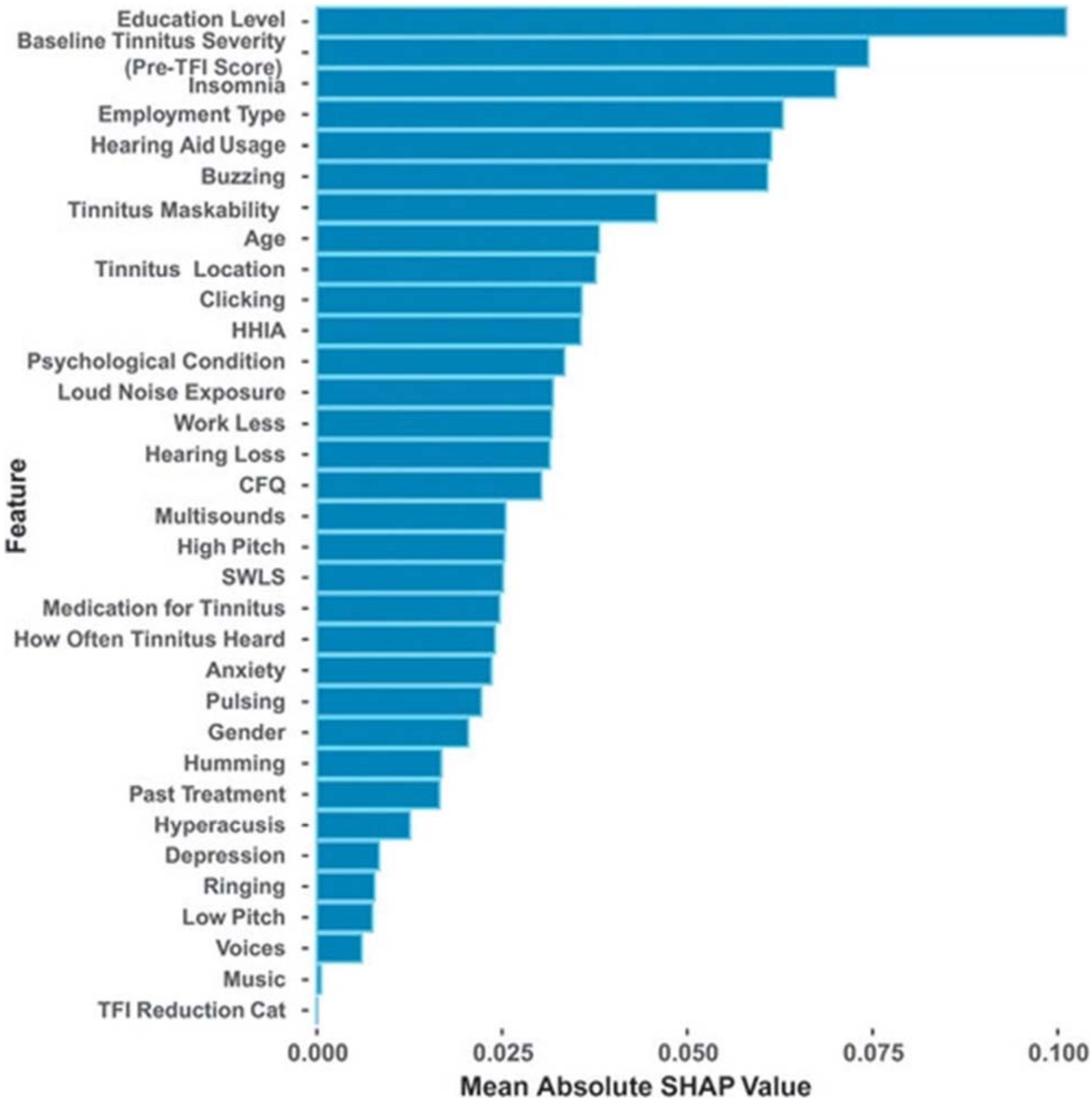

Figure 1: Feature importance based on mean absolute SHapley Additive exPlanations (SHAP) values from the optimal artificial neural network (ANN) model.

Predictor Variable Importance with the Best ANN Predictive Model

Figure 1 shows the relative importance of each predictor variable towards the outcome variable based on the mean absolute SHAP values as identified by the optimal ANN model. These mean

SHAP values represent the absolute change in log odds associated with a particular predictor variable. Hence, predictors with higher magnitude SHAP values indicate greater predictive power. As per the mean absolute SHAP values, Education level, baseline tinnitus severity, presence of insomnia, employment type, and hearing aid usage were the top five predictive variables that are most influential towards the outcome. Following is a summary of individual predictor variable category effects for those top five predictor variables.

Demographic variables

- Higher education level (positive effect Bachelor's degree and Master's degree or above).
- Age (positive effect for age >57 years).
- Employment type (positive effect professional, administrative, skilled tradesman, service occupation, medical, sales, home maker, student).

Tinnitus-related variables

- Greater baseline TFI score (positive effect with Pre TFI > 55.2/100).
- Tinnitus location (positive effect for tinnitus in only one ear, unsure and other categories).
- Positive effect for participants who described their tinnitus as a buzzing sound.
- Negative effect for participants who described their tinnitus as a clicking sound.

Treatment-related variables

- Tinnitus maskability defined as sound enrichment being effective in making tinnitus less noticable (positive effect for participants who never experienced tinnitus maskability at all or just partially).
- Hearing aid use (positive effect for those who did not use a hearing aid or used an aid only in one ear).

Clinical variables

- Insomnia (positive effect with lower insomnia < 14/28)

Figure 2 shows the impact of each predictor variable category (feature effect) on the outcome. A positive SHAP value on a given feature, for the plot named "1" (where 1 indicates the group with successful treatment outcome) reflects a positive impact from that category on the treatment outcome, while a negative SHAP value indicates a negative effect on the outcome. Positive treatment effects were found for the groups who had: higher education level (Bachelor's degree and Master's degree or above), and greater baseline tinnitus severity (Pre TFI > 55.2/100), lower insomnia (ISA < 14/28), employment types (professionals, administrative, skilled tradesmen, service occupational, medical, sales, homemaker, and students), hearing aid usage (using only one hearing aid or none), presence of buzzing type of tinnitus, maskability of tinnitus [defined as sound enrichment being effective in making tinnitus less noticeable], (participants who never had tinnitus maskability at all or just partially), participants with age>57 years, and tinnitus location (who had their tinnitus in only one ear, unsure and other categories).

Figure 2: Effect of individual features on the outcomes according to the ANN model. Each graph represents a feature vs. corresponding Shapley Additive explanation (SHAP) value.

Discussion

AI/ML approaches have not previously been applied to predict treatment outcomes. The current study examined the applications of a few AI/ML based models (ANN and SVM) in predicting the ICBT outcomes for tinnitus. Of the three models reviewed, the ANN model had the best predictive accuracy of post-intervention improvement with a fair AUC value (0.73). The SVM models with linear kernel and also radial basis kernel had AUC values of 0.72 and 0.70, respectively. The ANN was also more accurate than other data mining models like CART decision tree models (Rodrigo et al., 2021b). These findings align with several previous studies on audiovestibular data, identifying that ANN models are superior when examining audiovestibular data (Haro et al., 2020; Zhao et al., 2019). In contrast, a study examining the outcomes of a multi-modal 7-day program for tinnitus found the gradient boosting model has the highest predictive accuracy (Niemann et al., 2020). The difference could be due to different predictor variables, differences in sample size and different outcome measures.

Education level and baseline tinnitus severity have been consistently significant predictors throughout all our analyses (Rodrigo et al., 2021a, Rodrigo et al., 2021b). Those with a master's education or higher and greater baseline tinnitus severity had shown high potential for ICBT success. This was expected as having good literacy skills is essential when understanding the intervention materials. The intervention materials used in these studies were written at an average of ninth-grade reading level, suggesting that they were not easily accessible for participants with only a high school education (Rodrigo et al., 2021a). The association of greater tinnitus severity for a successful ICBT treatment has been demonstrated earlier by Beukes et al., 2018c. More details on how other factors that have been identified in this study are related to ICBT treatment success can be found in Rodrigo et al., 2021b. For other ICBT studies, greater severity and higher levels of associated anxiety have been found to be predictors. Higher intervention engagement and an increase in knowledge been further predictors to reduce the severity of symptoms(Rozental et al., 2019) , from which parallels can also be drawn. Similar results have also been obtained from ICBT for insomnia(Yeung et al., 2015) indicating greater insomnia, anxiety and depression predicted non-completion.

Although the ANN had the best predictive accuracy for the present study, its discriminative capacity was only moderate (Swets, 1988). In general, a model giving an AUC value of 0.90 or above has high discriminative power. Using a larger sample size and incorporating more relevant predictive variables might improve the predictive accuracy (Figueroa et al., 2012). Moreover, as the performance of models evaluated in this study and previous studies was similar (Niemann et al., 2020; Rodrigo et al., 2021b), it may be helpful to consider multiple models in future studies.

In addition to evaluating the predictive accuracy of data mining models, it is essential to examine which predictive variables contributed most to these models. In the current study, education type and the baseline tinnitus severity were the most significant predictors of ICBT success. Most of these findings are consistent with previous studies that have used univariate, multivariable and other data mining techniques (Rodrigo et al., 2021a, 2021b). As ICBT is a self-help intervention, being able to read and follow written instructions is very important. For this reason, it is not surprising that those with a master's level of education or above had higher success. However, efforts should be made to improve the accessibility of ICBT so that even those with limited education will receive good benefits. This can be done by revising the intervention for readability as we have done in the recent US version of the program (Beukes et al., 2020), as well as using a video-based intervention. Moreover, individuals with higher tinnitus distress (those with TFI scores of more than 50) are the ones who require more aggressive interventions such as the 8-week CBT program. For this reason, it is also not surprising that baseline tinnitus severity was the best predictor of ICBT success. However, efforts are needed to create different versions of the program to cater to individuals with varying levels of severity. In addition, the study was limited in terms of the number of predictive variables included. For this reason, future studies should aim to have other possible variables (e.g., health literacy, motivation, engagement, adherence) that may influence the ICBT success.

In conclusion, this study has shown decision tree models such as ANN and SVM help predict ICBT treatment outcomes as well as in identifying the predictors of outcome. Further work is recommended using more variables and larger sample sizes to improve the predictive power of these models.

Acknowledgement

This work was partially funded by the National Institute on Deafness and Communication Disorders (NIDCD) of the National Institute of Health (NIH) under the award number R21DC017214 awarded to Vinaya Manchaiah.

References

Andersson, G., Strömgren, T., Ström, L., & Lyttkens, L. (2002). Randomized controlled trial of internet-based cognitive behavior therapy for distress associated with tinnitus. *Psychosomatic medicine*, **64(5),** 810-816. https://doi.org/10.1097/01.psy.0000031577.42041.f8.

Axelsson, E., & Hedman-Lagerlöf, E. (2019). Cognitive behavior therapy for health anxiety: systematic review and meta-analysis of clinical efficacy and health economic outcomes. *Expert Review of Pharmacoeconomics & Outcomes Research,* **19(6),** 663-676. https://doi.org/10.1080/14737167.2019.1703182.

Barak A.Hen L. Boniel-Nissim M. Shapira Na'ama (2008) A comprehensive review and a meta-analysis of the effectiveness of Internet-based psychotherapeutic interventions. Journal of Technology in Human Service 2(4), 109-160. https://doi.org/10.1080/15228830802094429

Benitez J.M., Castro J.L., Requena I. (1997) Are artificial neural networks black boxes? *IEEE Trans Neural Netw.,* **8(5),**1156-64. https://doi.org/10.1109/72.623216. PMID: 18255717.

Beukes, E. W., Vlaescu, G., Manchaiah, V., Baguley, D. M., Allen, P. M., Kaldo, V., & Andersson, G. (2016). Development and technical functionality of an Internet-based intervention for tinnitus in the UK. *Internet Interventions*, **6**, 6-15. https://doi.org/10.1016/j.invent.2016.08.002

Beukes, E. W., Allen, P. M., Manchaiah, V., Baguley, D. M., & Andersson, G. (2017). Internet-based intervention for tinnitus: Outcome of a single-group open trial. *American Academy of Audiology, **12*****(4),** 340-351. https://doi.org/10.3766/jaaa.16055

Beukes, E. W., Manchaiah, V., Andersson, G., Allen, P. M., Terlizzi, P. M., & Baguley, D. M. (2018a). Situationally influenced tinnitus coping strategies: a mixed methods approach. *Disability and rehabilitation*, ***40*(24),** 2884–2894. https://doi.org/10.1080/09638288.2017.1362708

Beukes, E. W., Manchaiah, V., Davies, A. S., Allen, P. M., Baguley, D. M., & Andersson, G. (2018b). Participants' experiences of an Internet-based cognitive behavioural therapy intervention for tinnitus. *International journal of audiology*, **57(12),** 947-954. https://doi.org/10.1080/14992027.2018.1514538

Beukes, E.W., Manchaiah, V., Baguley, D.M., Allen, P.M. & Andersson, G. (2018c). Long-term efficacy of Audiologist-guided Internet-based Cognitive Behavior Therapy for Tinnitus. *American Journal of Audiology*, **27(3S),** 431-447. https://doi.org/10.1044/2018_AJA-IMIA3-18-0004

Beukes, E. W., Baguley, D. M., Allen, P. M., Manchaiah, V., & Andersson, G. (2018d). Audiologist-guided Internet-based cognitive behavior therapy for adults with tinnitus in the United Kingdom: A randomized controlled trial. *Ear and Hearing*, **39(3),** 423–433. https://doi.org/10.1097/AUD.0000000000000505

Beukes, E. W., Andersson, G., Allen, P. M., Manchaiah, V., & Baguley, D. M. (2018e). Effectiveness of guided Internet-based Cognitive Behavioral Therapy vs face-to-face clinical care for treatment of tinnitus: A randomized clinical trial. *JAMA Otolaryngology—Head & Neck Surgery*, **144(12),** 1126–1133. https://doi.org/10.1001/jamaoto.2018.2238

Beukes, E. W., Manchaiah, V., Allen, P. M., Baguley, D. M., & Andersson, G. (2019). Internet-Based Interventions for Adults With Hearing Loss, Tinnitus, and Vestibular Disorders: A Systematic Review and Meta-Analysis. *Trends in hearing*, **23**, https://doi.org/10.1177/2331216519851749

Beukes, E. W., Fagelson, M., Aronson, E. P., Munoz, M. F., Andersson, G., & Manchaiah, V. (2020). Readability following cultural and linguistic adaptations of an Internet-based intervention for tinnitus for use in the United States. *American Journal of Audiology*, **29(2)**, 97–109. https://doi.org/10.1044/2019_AJA-19-00014

Beukes, E. W., Manchaiah, V., Allen, P. M., Andersson, G., & Baguley, D. M. (2021). Exploring tinnitus heterogeneity. *Progress in brain research*, ***260***, 79–99. https://doi.org/10.1016/bs.pbr.2020.05.022

Bhatt, J.M., Lin, H.W., & Bhattacharyya, N. (2016). Prevalence, severity, exposures, and treatment patterns of tinnitus in the United States. *JAMA Otolaryngology–Head & Neck Surgery*, 142 (10), 959-965. https://doi.org/10.1001/jamaoto.2016.1700

Breiman L., Friedman J., Sone C.J., Olshen, R.A. (1984). Classification and Regression Trees. Chapman and Hall, USA.

Breiman, L. (2001) Random Forests. *Machine Learning,* **45,** 5–32. https://doi.org/10.1023/A:1010933404324

Broadbent, D. E., Cooper, P. F., FitzGerald, P., et al. (1982). The Cognitive Failures Questionnaire (CFQ) and its correlates. *Br J Clin Psychol*, **21(1)**, 1–16. https://doi.org/10.1111/j.2044-8260.1982.tb01421.x

Chen, T., and Guestrin, C. (2016) XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM; 2016 Aug 13;785–794. New York, NY, USA. https://doi.org/10.1145/2939672.2939785

Chong, W.S., Zaki, N. H. M., Hossain, M. S., Muslim, A. M., and Pour A. B. (2021) Introducing Theil-Sen estimator for sun glint correction of UAV data for coral mapping. *Geocarto International,* 1-30. https://doi.org/10.1080/10106049.2021.1892206

Chudzian, P. (2011) Radial Basis Function Kernel Optimization for Pattern Classification. In: Burduk, R., Kurzyński, M., Woźniak, M., Żołnierek, A. (eds) Computer Recognition Systems 4. Advances in Intelligent and Soft Computing, vol 95. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-20320-6_11

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273–297.

De Veaux, R.D. and Ungar, L.H. (1994), “Multicollinearity: A tale of Two Nonparametric Regressions,” in: P. Cheeseman, P. & Oldford, R.W . (eds.) Selecting models from data: AI and Statistics IV, pp. 293–302 (New York, NY: Springer-Verlag).

Diener, E., Emmons, R. A., Larsen, R. J., and Griffin S. (1985). The Satisfaction With Life Scale. *J Pers Assess*, **49(1)**, 71–75. https://doi.org/10.1207/s15327752jpa4901_13

Figueroa, R. L., Zeng-Treitler, Q., Kandula, S., & Ngo, L. H. (2012). Predicting sample size required for classification performance. *BMC Medical Informatics and Decision Making*, **12**. https://doi.org/10.1186/1472-6947-12-8

Friedman J.H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics,* **29(5)**,1189–1232. https://doi.org/10.1214/aos/1013203451

Fuller, T., Cima, R., Langguth, B., Mazurek, B., Vlaeyen, J.W.S., & Hoare, D.J. (2020). Cognitive behavioural therapy for tinnitus. *Cochrane Database of Systematic Reviews*, 1, CD012614. https://doi.org/10.1002/14651858.CD012614.pub2

Gandhi R. (2018) Boosting Algorithms: AdaBoost, Gradient Boosting and XGBoost. Retrieved from https://hackernoon.com/gradient-boosting-and-xgboost-90862daa6c77.

Haro, S., Smalt, C. J., Ciccarelli, G. A., & Quatieri, T. F. (2020). Deep Neural Network Model of Hearing-Impaired Speech-in-Noise Perception. *Frontiers in Neuroscience*, **14**, 588448. https://doi.org/10.3389/fnins.2020.588448

Juhola, M., Viikki, K., Laurikkala, J., Auramo, Y., Kentala, E., & Pyykkö, I. (2001). Application of artificial intelligence in audiology. *Scandinavian Audiology*,**30(1)**, 97–99. https://doi.org/10.1080/010503901300007209

Kaldo-Sandström, V., Larsen, H. C., & Andersson, G. (2004). Internet-based cognitive-behavioral self-help treatment of tinnitus: clinical effectiveness and predictors of outcome. *American Journal of Audiology*, ***13*(2)**, 185–192. https://doi.org/10.1044/1059-0889(2004/023)

Khalfa, S., Dubal, S., Veuillet, E., et al. (2002). Psychometric normalization of a hyperacusis questionnaire. *ORL J Otorhinolaryngol Relat Spec.*, **64**, 436–442. https://doi.org/10.1159/000067570.

Kraepelien, M., Blom, K., Jernelöv, S. et al. (2021) Weekly Self-Ratings of Treatment Involvement and Their Relation to Symptom Reduction in Internet Cognitive Behavioral Therapy for Insomnia. *Cognitive Therapy and Research*, **45**, 262–271 https://doi.org/10.1007/s10608-020-10151-y

Landry, E. C., Sandoval, X., Simeone, C. N., Tidball, G., Lea, J., & Westerberg, B. D. (2020). Systematic Review and Network Meta-analysis of Cognitive and/or Behavioral Therapies (CBT) for Tinnitus. *Otology & Neurotology,* **41(2),** 153–166. https://doi.org/10.1097/MAO.0000000000002472

López-López, J. A., Davies, S. R., Caldwell, D. M., et al. (2019). The process and delivery of CBT for depression in adults: a systematic review and network meta-analysis. *Psychological medicine*, **49(12),** 1937-1947 https://doi.org/10.1017/S003329171900120X

Lundberg, S. M. & Lee, S.-I. (2017), A Unified Approach to Interpreting Model Predictions, in I. Guyon; U. V. Luxburg; S. Bengio; H. Wallach; R. Fergus; S. Vishwanathan & R. Garnett, ed., 'Advances in Neural Information Processing Systems 30', Curran Associates, Inc., 4765-4774.

Lundberg, S.M., Erion, G., Chen, H., DeGrave, A., Prutkin, J.M., Nair, B., et al. (2019). Explainable AI for Trees: From Local Explanations to Global Understanding. arXiv preprint arXiv:190504610.

Manchaiah, V., Beukes, E. W., Granberg, S., Durisala, N., Baguley, D. M., Allen, P. M., & Andersson, G. (2018). Problems and Life Effects Experienced by Tinnitus Research Study Volunteers: An Exploratory Study Using the ICF Classification. *Journal of the American Academy of Audiology*, **29(10),** 936–947. https://doi.org/10.3766/jaaa.17094

Manchaiah, V., Vlaescu, G., Varadaraj, S., Aronson, E. P., Fagelson, M. A., Munoz, M. F., Andersson, G., & Beukes, E. W. (2020). Features, Functionality, and Acceptability of Internet-Based Cognitive Behavioral Therapy for Tinnitus in the United States. *American Journal of Audiology*, **29(3)**, 476–490. https://doi.org/10.1044/2020_AJA-20-00002

McCormack, A., Edmondson-Jones, M., Somerset, S., & Hall, D. (2016). A systematic review of the reporting of tinnitus prevalence and severity. *Hearing Research*, **337**, 70–79. https://doi.org/10.1016/j.heares.2016.05.009

McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, ***5*(4)**, 115–133.

Meikle, M. B., Henry, J. A., Griest, S. E., et al. (2012). The tinnitus functional index: development of a new clinical measure for chronic, intrusive tinnitus. *Ear and Hearing*, **33(2)**, 153–176.

Newman C.W., Weinstein B.E., Jacobson G.P., Hug G.A. (1991) Test-retest reliability of the hearing handicap inventory for adults. *Ear and Hearing*. **12(5)**, 355-357. https://doi.org/10.1097/00003446-199110000-00009.

Niemann, U., Boecking, B., Brueggemann, P., Mebus, W., Mazurek, B., & Spiliopoulou, M. (2020). Tinnitus-related distress after multi-modal treatment can be characterized using a key subset of baseline variables. *PloS One*, ***15*(1)**, e0228037. https://doi.org/10.1371/journal.pone.0228037

Niu, W.-J., Feng, Z.-K., Feng, B.-F., Min Y.-W. et al. (2019). Comparison of Multiple Linear Regression, Artificial Neural Network, Extreme Learning Machine, and Support Vector Machine in Deriving Operation Rule of Hydropower Reservoir. *Water,* **11(1),** 88. https://doi.org/10.3390/w11010088

Quinlan J.R. (1993) C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers, USA

Rodrigo, H.S., Beukes, E.W., Andersson, G., & Manchaiah, V. (2021a). Predictors of outcomes of internet-based cognitive behavioral therapy intervention for individuals with tinnitus. *BMJ Open,* ***11*(8)**, https://doi.org/10.1136/bmjopen-2021-049384.

Rodrigo, H.S., Beukes, E.W., Andersson, G., & Manchaiah, V. (2021b). Exploratory data mining to examine the outcome of internet-based cognitive behavioral therapy for tinnitus: Applications of decision tree and random forest models. *Journal of Medical Internet Research,* **23(11):**e28999, https://doi.org/10.2196/28999

Rozental, A., Andersson, G., & Carlbring, P. (2019). In the absence of effects: An individual patient data meta-analysis of non-response and its predictors in internet-based cognitive behavior therapy. *Frontiers in Psychology*, **10**, 589. https://doi.org/10.3389/fpsyg.2019.00589

Spitzer, R. L., Kroenke, K., Williams, J. B., et al. (1999). 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*, **282**, 1737–1744. https://doi.org/10.1001/jama.282.18.1737

Spitzer, R. L., Kroenke, K., Williams, J. B., and Lowe B., (2006). A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med*, **166**, 1092–1097. https://doi.org/10.1001/archinte.166.10.1092

Swets J. A. (1988). Measuring the accuracy of diagnostic systems. *Science (New York, N.Y.)*, **240(4857)**, 1285–1293. https://doi.org/10.1126/science.3287615

Tunkel, D. E., Bauer, C. A., Sun, G. H., et al., (2014). Clinical practice guideline: tinnitus. *Otolaryngology-Head and Neck surgery,* **151(2 Suppl),** S1–S40. https://doi.org/10.1177/0194599814545325

van der Zweerde, T., Bisdounis, L., Kyle, S. D., Lancee, J., & van Straten, A. (2019). Cognitive behavioral therapy for insomnia: a meta-analysis of long-term effects in controlled studies. *Sleep Medicine Reviews*, **48**, 101208 https://doi.org/10.1016/j.smrv.2019.08.002

Wang, D. (2017). Deep Learning Reinvents the Hearing Aid: Finally, wearers of hearing aids can pick out a voice in a crowded room. *IEEE spectrum*, ***54*(3)**, 32–37. https://doi.org/10.1109/MSPEC.2017.7864754

Warmerdam, L., Smit, F., van Straten A., Riper H., Cuijpers P. (2010) Cost-utility and cost-effectiveness of internet-based treatment for adults with depressive symptoms: randomized trial. *J Med Internet Res*. **12(5)** e53. https://doi.org/10.2196/jmir.1436.

Yeung, W. F., Chung, K. F., Ho, F. Y., & Ho, L. M. (2015). Predictors of dropout from internet-based self-help cognitive behavioral therapy for insomnia. *Behaviour research and therapy*, **73**, 19–24. https://doi.org/10.1016/j.brat.2015.07.008

Yu, H., Zhang, Y., Liu, Q., & Yan, R. (2021). Efficacy of online and face-to-face cognitive behavioral therapy in the treatment of neurological insomnia: a systematic review and meta-analysis. *Annals of Palliative Medicine*, **10(10),** 10684-10696. https://doi.org/10.21037/apm-21-2387.

Zhao, Y., Li, J., Zhang, M., Lu, Y., Xie, H., Tian, Y., & Qiu, W. (2019). Machine Learning Models for the Hearing Impairment Prediction in Workers Exposed to Complex Industrial Noise: A Pilot Study. *Ear and Hearing*, ***40*(3),** 690–699. https://doi.org/10.1097/AUD.0000000000000649