

ORIGINAL RESEARCH

Prediction of vestibular schwannoma recurrence using artificial neural network

Mehdi Abouzari MD, PhD^{1,2} | Khodayar Goshtasbi BS¹ | Brooke Sarna BS¹ |
 Pooya Khosravi BS^{1,3} | Trevor Reutershan MS^{1,3} | Navid Mostaghni BS^{1,3} |
 Harrison W. Lin MD¹ | Hamid R. Djalilian MD^{1,3} 

¹Division of Neurotology and Skull Base Surgery, Department of Otolaryngology-Head and Neck Surgery, University of California, Irvine, California

²Division of Pediatric Otolaryngology, Children's Hospital of Orange County, Orange, California

³Department of Biomedical Engineering, University of California, Irvine, California

Correspondence

Mehdi Abouzari, Division of Neurotology and Skull Base Surgery, Department of Otolaryngology-Head and Neck Surgery, University of California Irvine, 200 S. Manchester Avenue, Suite 400, Orange, CA 92868.
 Email: mabouzar@hs.uci.edu

Funding information

National Center for Advancing Translational Sciences, Grant/Award Number: TL1TR001415-04; National Center for Research Resources, Grant/Award Number: TL1TR001415-04; NIH Medical Scientist Training Program, Grant/Award Number: T32GM008620

Abstract

Objectives: To compare two statistical models, namely logistic regression and artificial neural network (ANN), in prediction of vestibular schwannoma (VS) recurrence.

Methods: Seven hundred eighty-nine patients with VS diagnosis completed an online survey. Potential predictors for recurrence were derived from univariate analysis by reaching the cut off *P* value of .05. Those nine potential predictors were years since treatment, surgeon's specialty, resection amount, and having incomplete eye closure, dry eye, double vision, facial pain, seizure, and voice/swallowing problem as a complication following treatment. Multivariate binary logistic regression model was compared with a four-layer 9-5-10-1 feedforward backpropagation ANN for prediction of recurrence.

Results: The overall recurrence rate was 14.5%. Significant predictors of recurrence in the regression model were years since treatment and resection amount (both $P < .001$). The regression model did not show an acceptable performance (area under the curve [AUC] = 0.64; $P = .27$). The regression model's sensitivity and specificity were 44% and 69%, respectively and correctly classified 56% of cases. The ANN showed a superior performance compared to the regression model (AUC = 0.79; $P = .001$) with higher sensitivity (61%) and specificity (81%), and correctly classified 70% of cases.

Conclusion: The constructed ANN model was superior to logistic regression in predicting patient-answered VS recurrence in an anonymous survey with higher sensitivity and specificity. Since artificial intelligence tools such as neural networks can have higher predictive abilities compared to logistic regression models, continuous investigation into their utility as complementary clinical tools in predicting certain surgical outcomes is warranted.

Mehdi Abouzari and Khodayar Goshtasbi contributed equally to this manuscript.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2020 The Authors. *Laryngoscope Investigative Otolaryngology* published by Wiley Periodicals, Inc. on behalf of The Triological Society.

KEYWORDS

acoustic neuroma, artificial intelligence, artificial neural network, logistic regression, recurrence, vestibular schwannoma

1 | INTRODUCTION

Emerging evidence suggests the benefits of mathematical and statistical models on decision support in medicine.¹⁻⁵ Artificial neural networks (ANN) and logistic regression are among popular mathematical algorithms, besides other models such as support vector machines or gradient-boosted tree methods, which can potentially enhance diagnostic/management accuracy and thus result in better treatment decisions and more appropriate use of health care resources.⁶⁻¹⁰ Prediction of recurrence especially in cancer patients is one promising application of both regression and ANN models.¹¹⁻¹⁵ In otolaryngology, however, the utilization of these methods is scarce likely due to a lack of wide understanding or easily implementable application.

ANN is a collection of connected units or nodes called artificial neurons structured in two or more layers (ie, input, output, and one or more hidden layers), which loosely model the neurons in a biological brain. The ANN itself is not an algorithm, but rather a framework for many different machine learning algorithms to work together and process complex data inputs. Hence, ANN is a great tool for suggesting possible associations between different predictors and outcome, though the majority of current deep learning methods are not readily interpretable regarding specific influential nodes and connections. The variability and complicated number of interactions in otologic/neurotologic diseases allow the presumptuous occurrence of nonlinear relations. This problem could be at least theoretically approached by using ANN.

Vestibular schwannoma (VS), also known as acoustic neuroma, is the most common benign tumor originating in the cerebellopontine angle with its reported incidence increasing in recent years.^{16,17} Although the recurrence of VS is relatively small,¹⁸ several risk factors including tumor- and treatment-related factors may influence the recurrence rate.¹⁹⁻²² This makes the entity multivariable and suitable for implementation in an ANN model. To our knowledge, this is the first study that aims to compare the performance of logistic regression and ANN models in predicting recurrence in VS patients.

2 | MATERIALS AND METHODS

2.1 | Dataset and variable selection

The dataset we used consisted of 789 VS patients who completed a comprehensive survey (Appendix S1) distributed to members of the acoustic neuroma association (ANA) from January to March 2017 using a secure and confidential RedCap interface (Nashville, Tennessee).²³ Of these, 698 (88.5%) patients who answered the question regarding the presence or lack of VS recurrence according to their

neurotologists' or neurosurgeons' recognition were included. Survey links were distributed via ANA website, Facebook, and email list. The study was approved by the Institutional Review Board Committee of the University of California, Irvine. The survey evaluated patients' demographics, tumor size, presenting symptoms, post-treatment complications, treatment centers, and years since initial treatment. The cohort was also reviewed for type of surgeon (ie, neurosurgeon, neurotologist, or both), resection amount (ie, gross-total or subtotal), and applied surgical approach (ie, translabyrinthine, retrosigmoid/suboccipital, and middle fossa).

2.2 | Neural network model

The dataset was randomly divided to three subsets: training set (350 cases), validation set (174 cases), and test set (174 cases). Subjects

TABLE 1 Factors which univariate analysis revealed were not statistically predictors of recurrence of vestibular schwannoma

Factor	Recurrence	Nonrecurrence	P value
Sex (M/F)	31/70	208/389	.31
Age at diagnosis	50.04 ± 12.60	51.58 ± 11.53	.22
Tumor size (cm)	2.25 ± 1.13	2.08 ± 1.25	.21
Neurofibromatosis			.67
Yes	2 (11.7%)	15 (88.3%)	
No	77 (15.4%)	424 (84.6%)	
Unknown	22 (12.2%)	158 (87.8%)	
Surgical approach			.12
Translabyrinthine	18 (15.6%)	97 (84.4%)	
Retrosigmoid/ suboccipital	29 (13.0%)	194 (87.0%)	
Middle fossa	7 (12.3%)	50 (87.7%)	
Unknown	3 (60.0%)	2 (40.0%)	
Medical center			.43
Academic hospital	65 (16.4%)	332 (83.6%)	
Private hospital	32 (12.9%)	217 (87.1%)	
VA hospital	4 (7.7%)	48 (92.3%)	
Treatment complication ^a			.75
Present	12 (13.0%)	80 (87.0%)	
Absent	89 (14.7%)	517 (85.3%)	

Abbreviations: M/F, male/female ratio; VA, veteran administration.

^aDetailed post-treatment complications other than noted in Table 2 (ie, tinnitus, imbalance, cognitive problems, cerebrospinal fluid leak, synkinesis, headache, hydrocephalus, meningitis, and stroke) were not statistically different between the recurrence and nonrecurrence groups.

were taken from the entire sample according to a conventional random number generation procedure and were then allocated to the three subsets with predetermined sizes. The 2:1:1 size ratio chosen for the training, validation, and test sets is a popular method in classification algorithms. A feedforward backpropagation ANN was developed and

TABLE 2 Factors derived from univariate analyses with significant relationship with recurrence of vestibular schwannoma for subsequent statistical modeling

Factor	Recurrence	Nonrecurrence	P value
Years since treatment	8.94 ± 7.18	7.44 ± 7.06	.04
Type of surgeon			.007
Neurosurgeon	22 (25.5%)	64 (74.5%)	
Neurotologist	2 (8.3%)	22 (91.7%)	
Both	77 (26.5%)	213 (73.5%)	
Resection amount			<.001
Gross-total	17 (5.9%)	270 (94.1%)	
Subtotal	35 (31.0%)	78 (69.0%)	
Incomplete eye closure ^a			.02
Present	30 (19.7%)	122 (80.3%)	
Absent	71 (13.0%)	475 (87.0%)	
Dry eye ^a			.03
Present	43 (18.7%)	187 (81.3%)	
Absent	58 (12.4%)	410 (87.6%)	
Double vision ^a			.01
Present	18 (25.0%)	54 (75.0%)	
Absent	83 (13.2%)	543 (86.8%)	
Facial pain ^a			.04
Present	18 (22.8%)	61 (77.2%)	
Absent	83 (13.4%)	536 (86.6%)	
Seizure ^a			.04
Present	3 (50.0%)	3 (50.0%)	
Absent	98 (14.2%)	594 (85.8%)	
Voice/swallowing problems ^a			.01
Present	18 (25.0%)	54 (75.0%)	
Absent	83 (13.2%)	543 (86.8%)	

^aPost-treatment complication.

trained using MATLAB R2010b (The MathWorks, Inc). Nine neurons, corresponding to nine independent variables as in the regression model, were created for the input layer. Therefore, the output of each neuron in the input layer represented the value of one potential predictor. The network was fully connected; that is, each neuron in a given layer was connected to all neurons in the next layer. One neuron was used for the output layer. This neuron classified the recurrence as 0 (nonrecurrence) or 1 (recurrence). The task performed by hidden layers was to detect features of the input space (mapped on the input layer) and feed them as input to the output layer for final classification decision. The neurons in hidden layers, therefore, did not represent actual variables. The optimal number of hidden layers and the number of neurons in each were determined by trial and error. Continuous tan-sigmoid functions were used as transfer functions in the network. The Levenberg-Marquardt algorithm was used as the training method. Training was based on the batch training method and stopped when the error (measured as mean squared error; MSE) obtained from the validation set started to increase. To account for overfitting in models that used a loss function, L2 regularization, namely adding squared magnitude of coefficients to the loss function as penalty, was incorporated. The network was trained for a maximum of 100 epochs or until the optimal performance goal (MSE < 10⁻¹⁰) was achieved. The network was applied to the test set, the receiver operating characteristic (ROC) curve was plotted, and the AUC was calculated. For each neural network architecture design, the top models trained using random initializations and predicting the highest AUC scores were chosen. These models were further ensemble into a voting classifier to characterize the final prediction according to majority vote for accuracy.

2.3 | Statistical analysis

The training and validation sets were combined for regression analysis. Independent sample *t* test and chi-square test were used for numerical and categorical variables, respectively. Variables with statistically significant *P* values (<.05) were selected for regression analysis. There are a number of approaches for prediction of an entity's outcome. One of the simplest methods is univariate analysis through which we are able to compare two groups of subjects (with and without a specific outcome [ie, recurrence]) for the presence of a potential risk factor. Following univariate analysis, multivariate binary logistic

Model	Specificity	Sensitivity	Accuracy	AUC score
Logistic regression	0.72	0.46	0.59	0.64
Random forest	0.91^a	0.25	0.58	0.73
Ada boost	0.65	0.50	0.58	0.66
Artificial neural network (9-5-10-1)	0.83	0.55^a	0.69^a	0.83^a
Artificial neural network (9-10-1)	0.81	0.49	0.65	0.74
Artificial neural network (9-5-10-10-1)	0.78	0.49	0.63	0.70

TABLE 3 Comparison of different predictive models trained on the dataset

Abbreviation: AUC, area under the curve.

^aThe best performance of each vertical category is bolded.

regression was performed with recurrence as the dependent variable and years since treatment, type of surgeon, resection amount, and incomplete eye closure, dry eye, double vision, facial pain, seizure, and voice/swallowing problems among post-treatment complications as potential predictors. The regression equation obtained was then applied to the test set, the ROC curve was plotted and the AUC was calculated. SPSS 18.0 software (SPSS, Inc, Chicago, Illinois) was used for statistical analysis.

3 | RESULTS

The overall recurrence rate was 14.5% (101 out of 698 subjects). Factors not predictive of recurrence on univariate analysis are demonstrated in

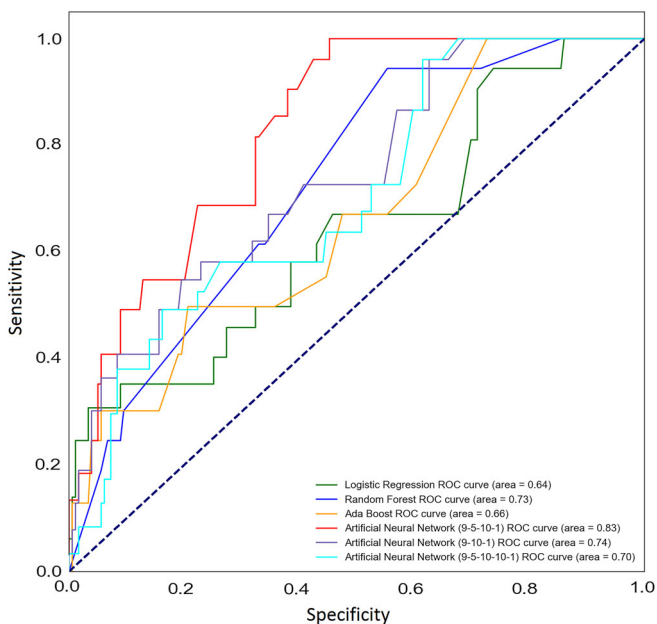


FIGURE 1 The receiver operating characteristic (ROC) curves for different predictive models trained with their respective area under the curves. Dotted line denotes to the reference line

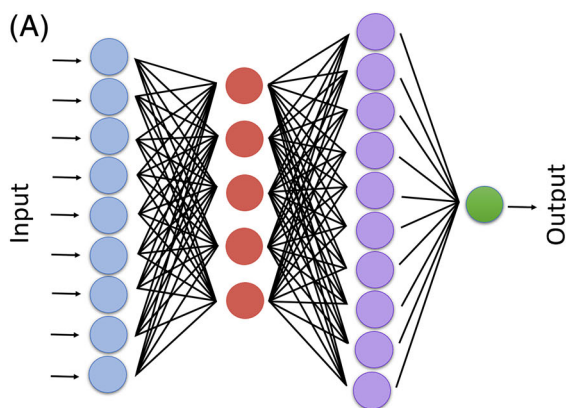


Table 1. None of the presenting symptoms (ie, tinnitus, imbalance, vertigo, headache, hearing loss, aural fullness, eye problems, facial numbness/paralysis, and voice/swallowing problems) were significantly different between the recurrence and nonrecurrence groups.

As shown in Table 2, univariate analyses extracted nine variables with significant relationship with recurrence including years since treatment ($P = .04$), type of surgeon ($P = .007$), and resection amount ($P < .001$). Incomplete eye closure ($P = .02$), dry eye ($P = .03$), double vision ($P = .01$), facial pain ($P = .04$), seizure ($P = .04$), and voice/swallowing problems ($P = .01$) were the only post-treatment complications which were significantly different between the recurrence and nonrecurrence groups. Among these variables, significant predictors of recurrence in the regression model were years since treatment (95% CI, 1.03-1.11; $P < .001$) and resection amount (95% CI, 5.16-22.93; $P < .001$). The regression model did not show an acceptable performance on the test set (AUC = 0.64; 95% CI, 0.45-0.84, $P = .27$). It had a sensitivity of 44% and a specificity of 69%, and correctly classified 56% of cases.

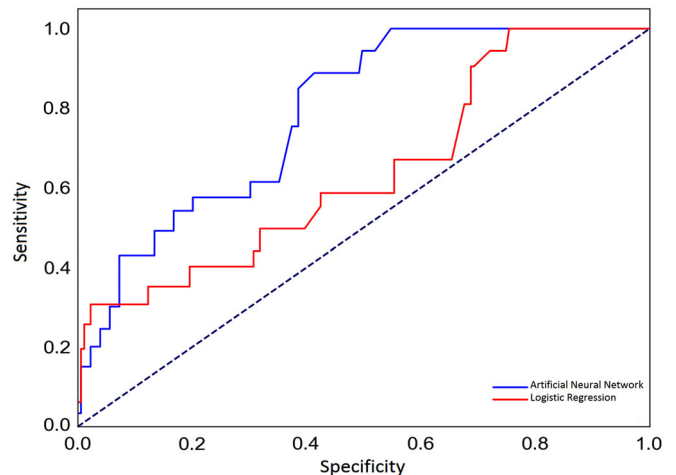


FIGURE 3 The receiver operating characteristic (ROC) curves for prediction of vestibular schwannoma recurrence using the logistic regression (solid red line) and artificial neural network (solid blue line) models. Dotted line denotes to the reference line

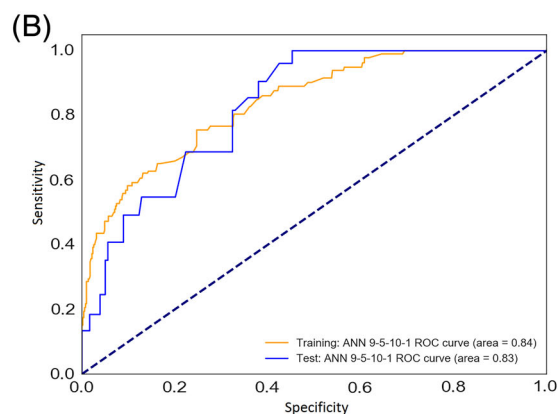


FIGURE 2 A, Schematic architecture of a four layer 9-5-10-1 neural network similar to the one designed in this study. B, Training and test sets' results for the 9-5-10-1 model with their associated area under the curves. Dotted line denotes to the reference line

Two hidden layers with 5 and 10 neurons provided the best structure for the neural network compared to other architectures (Table 3), where these models' AUC comparisons are further depicted in Figure 1. Therefore, a four-layer 9-5-10-1 feedforward backpropagation ANN was developed and trained (Figure 2). The ANN showed a superior performance compared to the regression model (AUC = 0.79; 95% CI, 0.67-0.82, $P = .001$). It had a sensitivity of 61% and a specificity of 81%, and correctly classified 70% of cases (Figure 3).

4 | DISCUSSION

We found that the VS recurrence rate in a large sample of ANA respondents was 14.5% and years since treatment and resection amount were the two independent predictors of recurrence. Various studies have reported a recurrence rate of 5% to 45% associated with different treatment modalities.^{21,22,24-29} The observed recurrence rate in this study is within the wide range of reported rates in the literature, which is likely due to its comprehensive patient-centered and survey-based nature. It is plausible to consider that patients with either poor or excellent outcomes may be more inclined to participate in these types of nonprofit organization. In other words, one could conceive that an organization focused on patient advocacy and support might disproportionately attract patients with outcomes or treatment courses that deviate from the average VS patient experience.³⁰

Our statistical models revealed that the amount of resection was one of the two independent predictors of recurrence. This finding is in agreement with the previous reports showing that incomplete resection of VS is associated with a significant risk of tumor recurrence requiring subsequent intervention.^{21,22,31} Although the observed association between incomplete resection and recurrence is not novel, the incorporation of an ANN model to a patient-reported survey of VS recurrence factors and reaching this logical association has not yet been reported in the literature. Since patient-reported answers are only a proxy for primary clinical data, future studies are encouraged to utilize more objective or reliable measures. This work presents as a proof-of-concept utilization of ANN in VS patients and offers potentially promising future uses of machine learning methods if proper mathematical algorithms are utilized in the hands of an experienced user with sufficient clinical and statistical knowledge. Complete VS resection is case-dependent and possibly challenging in some operations to preserve nerve function; accordingly, the literature suggest longer term follow-up with imaging in partial or subtotal resection is necessary.^{29,32-36}

Although our positive findings of years since treatment and degree of tumor resection's association with VS recurrence are reflected in the literature, there is a lack of agreement on the prevalence of recurrence with our observed 14.5% rate falling within a wide range (5%-45%).^{21,22,24-29} It is important to note that most of these literature-reported values concern specific patient groups undergoing selective inclusion criteria or particular treatment modalities and

surgical techniques. This may affect the outcome depending on patients' baseline and degree of intervention and follow-up. Some studies may struggle with following patients for a long time which is essential for documenting a credible rate of tumor recurrence in patients with VS.²⁹ On the other hand, this may be one of the strengths and advantages of such a survey-based study whose nature allows for a wider range of years of follow-up. Although our survey-based study is not without its limitations as we will later discuss, it allows for the opportunity to gather comprehensive information from a diverse and nation-based patient population which would have been a nearly impossible task if done in a prospective single-institutional setting. To our knowledge, this is one of the first reports of a VS recurrence rate based on a large, general, and heterogeneous patient population with a wide range of years from treatment to study participation.

Although the utilized univariate analysis is a common approach for prediction of an entity's outcome, the confounding effects of other possible risk factors are ignored in this approach. To solve this problem, more sophisticated techniques such as multivariate logistic regression have been developed to evaluate the independent predictive roles of more than one variable. Regression models have two main shortcomings including the assumption of normality for residuals beside their inability to identify nonlinear relationships.³⁷ Artificial intelligence tools such as machine learning methods are evolving to avoid limitations of traditional outcome prediction methods. For instance, decision trees, Bayesian networks, and ANNs are gaining increased applications in the field of otolaryngology.^{4,38,39} These methods are capable of detecting complex association and potentially nonlinear relationships between potential predictors and outcome. In this study, among nine potential predictors of VS recurrence, only more years since initial treatment and less resection amount at the time of surgery were independent predictors of recurrence in the regression model. The ANN model, on the other hand, used nonlinear, potentially subtle effects of linearly nonsignificant variables to lead to a better accuracy outcome compared to the regression model.

We achieved a reasonable classification accuracy of 70% in a large sample of VS patients by utilizing nine variables in a four-layered neural network. Using an input layer of nine variables and a four-layer rather than a three-layer ANN model probably contributed to the relatively high sensitivity and specificity of the model as well as the accuracy for its prediction of recurrence. In other words, the conventional regression model can be viewed as a simple ANN with one input layer, one output layer, and no hidden layer resulting in a poor performance. This contrast was not unexpected given the wide range of factors that may potentially influence the recurrence in VS, and the complexity of the corresponding relationships. This finding is in line with the result of a comparative review of the performance of logistic regression and ANNs in 72 articles.⁴⁰ Dreiseitl and Ohno-Machado showed that in 61 out of 72 studies which used statistical methods for comparisons, regression was superior in only one case, ANN was superior in 18 cases, and they were equally good in the remaining 42 studies.⁴⁰

To our knowledge, the application of an ANN for prediction of patient-reported recurrence in VS patients is a novel idea. The most

plausible interpretation from our ANN model is that VS patients undergoing subtotal tumor removal are more likely to experience tumor recurrence and thus should be evaluated more frequently via serial imaging. By also demonstrating years since treatment as another independent variable, we suggest that though stable patients may incrementally increase the durations between each serial magnetic resonance imaging (MRI), monitoring should not be discontinued. It is important to emphasize that we do not suggest that ANN can substitute MRI as a VS monitoring modality. Instead, we propose that ANN has the potential of predicting certain treatment outcomes with good accuracy. Namely, we demonstrated that partial (subtotal) tumor removal and years since initial treatment deem more frequent monitoring compared to the other analyzed variables. Likewise, "wait and watch" VS patients can be monitored as needed based on a frequency that is both optimal for the patient and one that reduces extra health care costs associated with unnecessary imaging orders. Future ANN analyses of large VS cohorts including additional clinical variables can potentially result in a mathematical algorithm that generates a frequency value (eg, 6 months vs 2 years) based on any individual VS's likeliness of progression and recurrence. This can be expanded to other areas of medicine, where such ANN models can recommend an optimal serial imaging or medical lab frequency based on a patient's clinical characteristics and disease progression risk factors.

Despite the promising results achieved in this study, the proposed ANN model should be tested on a general VS population rather than a survey-based ANA respondent population. It has been shown that inherent limitations and biases (ie, participation bias) are associated with survey studies from national patient support organizations such as ANA.^{30,41} Prummer et al showed that the population profile of ANA survey respondents likely differs significantly from the greater population of VS patients that may be encountered at a tertiary referral center.³⁰ Meanwhile, some of the participants completed the survey long time after their initial diagnosis and treatment, which can lead to recall bias. The study is also limited by the lack of details on the tumors' characteristics or the experience of the surgeons or the physicians who performed the treatment planning. Also, the degree of adherence to other nerves or brainstem, vascularity of the tumor, and internal auditory canal penetration are potential factors that could affect recurrence rate.

There were some potential limitations in the neural network model as well. There is limited generalizability of the role of this ANN as a VS decision-making clinical tool beyond its better predictive ability than logistic regression in our database. The regression model's low predictivity (56%) was also relatively low, warranting future investigations comparing ANN with higher predictive regression models. For instance, a regression model utilizing interaction terms might have a better performance than our developed model in this study. Since the ANN model was limited to remaining variables following statistical dimension reduction via univariate analysis, the potential influence of all possible relationships even those insignificant on univariate analysis were lost. This exploratory study builds the foundation for a future study to develop more sophisticated ANN models utilizing larger

databases and including all available variables. Lastly, the current ANN model is unable to compare or contrast which survey variables were most influential. Even though the constructed ANN model was superior to logistic regression, it still underperformed compared to ANN models applied to complex datasets achieving >80% classification performance.⁴²⁻⁴⁵ This can be due to the chosen neural network architecture and/or the small sample size.

In our experience with this patient-surveyed database, there were four main limitations that could be addressed in future investigations. First, the VS sample cohort may have not been large enough to appropriately train and validate various ANN models. Future studies may benefit from large national databases which can provide significantly larger cohorts for developing machine learning algorithms.^{46,47} Second, various advanced techniques to potentially control for overfitting (eg, cross-validation, Bayesian regularization, and ensembling) were not implemented, and the involved hyperparameters were not modulated in prospective investigations. Third, though beyond the scope of this proof-of-concept report, future studies can improve the machine learning architecture by considering the transfer of model learning from pre-existing ANN models, developing and analyzing a range of ANN models, as well as employing other machine learning algorithms (eg, support vector machines, K nearest neighbors, gradient boosted trees, etc). Lastly, and more specific to the utilized patient-answered survey, the constructed model may have missed important predictors that were not available in the database or answered by experienced clinicians. Despite these limitations, we hope that the presented manuscript will encourage future utilization of machine learning and ANN in predictive modeling regarding VS and other head and neck conditions.

5 | CONCLUSION

Our statistical models showed that the degree of tumor resection and years since initial treatment were the two independent predictors of VS recurrence. The constructed ANN model was superior to logistic regression in predicting patient-answered VS recurrence in an anonymous survey with higher sensitivity and specificity. The predictions in this model have to be compared against real clinical data to validate this proof-of-concept report. Since artificial intelligence tools such as neural networks can have higher predictive abilities compared to logistic regression models, continuous investigation into their utility as complementary clinical tools in predicting certain surgical outcomes is warranted.

ACKNOWLEDGMENTS

The authors sincerely thank the Acoustic Neuroma Association directors and members for their participation in facilitating this study. The authors wish to acknowledge the reviewers for their insightful comments on the original manuscript.

CONFLICT OF INTEREST

The authors declare no potential conflict of interest.

ORCID

Hamid R. Djalilian  <https://orcid.org/0000-0003-2270-5207>

REFERENCES

- Kapetanovic IM, Rosenfeld S, Izmirlan G. Overview of commonly used bioinformatics methods and their applications. *Ann N Y Acad Sci*. 2004;1020:10-21.
- Ramesh AN, Kambhampati C, Monson JR, Drew PJ. Artificial intelligence in medicine. *Ann R Coll Surg Engl*. 2004;86:334-338.
- Friedland DR, Tarima S, Erbe C, Miles A. Development of a statistical model for the prediction of common vestibular diagnoses. *JAMA Otolaryngol Head Neck Surg*. 2016;142:351-356.
- Morse JC, Shilts MH, Ely KA, et al. Patterns of olfactory dysfunction in chronic rhinosinusitis identified by hierarchical cluster analysis and machine learning algorithms. *Int Forum Allergy Rhinol*. 2019;9:255-264.
- Johnson A, Tarima S, Wong S, Friedland DR, Runge CL. Statistical model for prediction of hearing loss in patients receiving cisplatin chemotherapy. *JAMA Otolaryngol Head Neck Surg*. 2013;139:256-264.
- Obermeyer Z, Emanuel EJ. Predicting the future—big data, machine learning, and clinical medicine. *N Engl J Med*. 2016;375:1216-1219.
- Chen JH, Asch SM. Machine learning and prediction in medicine—Beyond the peak of inflated expectations. *N Engl J Med*. 2017;376:2507-2509.
- Shipp MA, Ross KN, Tamayo P, et al. Diffuse large B-cell lymphoma outcome prediction by gene-expression profiling and supervised machine learning. *Nat Med*. 2002;8:68-74.
- Ghahramani Z. Probabilistic machine learning and artificial intelligence. *Nature*. 2015;521:452-459.
- Das A, Ben-Menachem T, Cooper GS, et al. Prediction of outcome in acute lower-gastrointestinal haemorrhage based on an artificial neural network: internal and external validation of a predictive model. *Lancet*. 2003;362:1261-1266.
- Faradmal J, Soltanian AR, Roshanaei G, Khodabakhshi R, Kasabian A. Comparison of the performance of log-logistic regression and artificial neural networks for predicting breast cancer relapse. *Asian Pac J Cancer Prev*. 2014;15:5883-5888.
- Kourou K, Exarchos TP, Exarchos KP, Karamouzis MV, Fotiadis DI. Machine learning applications in cancer prognosis and prediction. *Comput Struct Biotechnol J*. 2014;13:8-17.
- Hu X, Cammann H, Meyer HA, Miller K, Jung K, Stephan C. Artificial neural networks and prostate cancer—tools for diagnosis and management. *Nat Rev Urol*. 2013;10:174-182.
- Jerez-Aragónés JM, Gómez-Ruiz JA, Ramos-Jiménez G, Muñoz-Pérez J, Alba-Conejo E. A combined neural network and decision trees model for prognosis of breast cancer relapse. *Artif Intell Med*. 2003;27:45-63.
- Planz B, Deix T, Caspers HP. Prediction of tumor recurrence and progression of superficial bladder carcinoma using an artificial neural network. *Urologe A*. 2007;46:1138-1139.
- Stangerup SE, Tos M, Thomsen J, Caye-Thomasen P. True incidence of vestibular schwannoma? *Neurosurgery*. 2010;67:1335-1340.
- Propp JM, McCarthy BJ, Davis FG, Preston-Martin S. Descriptive epidemiology of vestibular schwannomas. *Neuro-Oncol*. 2006;8:1-11.
- Shelton C. Unilateral acoustic tumors: how often do they recur after translabyrinthine removal? *Laryngoscope*. 1995;105:958-966.
- Ahmad RA, Sivalingam S, Topsakal V, et al. Rate of recurrence vestibular schwannoma after total removal via different surgical approaches. *Ann Otol Rhinol Laryngol*. 2012;121:156-161.
- Cerullo L, Gutsch J, Osterdock R. Recurrence of vestibular (acoustic) schwannomas in surgical patients where preservation of facial and cochlear nerve is the priority. *Br J Neurosurg*. 1998;12:547-552.
- El-Kashlan HK, Zeitoun H, Arts HA, Hoff JT, Telian SA. Recurrence of acoustic neuroma after incomplete resection. *Am J Otol*. 2000;21:389-392.
- Fukuda M, Oishi M, Hiraishi T, Natsumeda M, Fujii Y. Clinicopathological factors related to regrowth of vestibular schwannoma after incomplete resection. *J Neurosurg*. 2011;114:1224-1231.
- Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*. 2009;42:377-381.
- Kameyama S, Tanaka R, Kawaguchi T, Honda Y, Yamazaki H, Hasegawa A. Long-term follow-up of the residual intracanalicular tumours after subtotal removal of acoustic neurinomas. *Acta Neurochir*. 1996;138:206-209.
- Ohta S, Yokoyama T, Nishizawa S, Uemura K. Regrowth of the residual tumour after acoustic neurinoma surgery. *Br J Neurosurg*. 1998;12:419-422.
- Samii M, Gerganov VM, Samii A. Functional outcome after complete surgical removal of giant vestibular schwannomas. *J Neurosurg*. 2010;112:860-867.
- Seol HJ, Kim CH, Park CK, et al. Optimal extent of resection in vestibular schwannoma surgery: relationship to recurrence and facial nerve preservation. *Neurol Med Chir (Tokyo)*. 2006;46:176-180.
- Bloch DC, Oghalai JS, Jackler RK, Osofsky M, Pitts LH. The fate of the tumor remnant after less-than-complete acoustic neuroma resection. *Otolaryngol Head Neck Surg*. 2004;130:104-112.
- Roche PH, Ribeiro T, Khalil M, et al. Recurrence of vestibular schwannomas after surgery. *Prog Neurol Surg*. 2008;21:89-92.
- Prummer CM, Kerezoudis P, Tombers NM, Peris-Celda M, Link MJ, Carlson ML. Influence of selection bias in survey studies derived from a patient-focused organization: a comparison of response data from a single tertiary care center and the Acoustic Neuroma Association. *Otol Neurotol*. 2019;40:504-510.
- Lownie SP, Drake CG. Radical intracapsular removal of acoustic neurinomas Long-term follow-up review of 11 patients. *J Neurosurg*. 1991;74:422-425.
- Nakatomi H, Jacob JT, Carlson ML, et al. Long-term risk of recurrence and regrowth after gross-total and subtotal resection of sporadic vestibular schwannoma. *J Neurosurg*. 2017;1-7. <https://doi.org/10.3171/2016.11.JNS16498>.
- Carlson ML, Van Abel KM, Driscoll CL, et al. Magnetic resonance imaging surveillance following vestibular schwannoma resection. *Laryngoscope*. 2012;122:378-388.
- Tang S, Griffin AS, Waksal JA, et al. Surveillance after resection of vestibular schwannoma: measurement techniques and predictors of growth. *Otol Neurotol*. 2014;35:1271-1276.
- Chen Z, Prasad SC, Di Lella F, et al. The behavior of residual tumors and facial nerve outcomes after incomplete excision of vestibular schwannomas. *J Neurosurg*. 2014;120:1278-1287.
- Schwartz MS, Kari E, Strickland BM, et al. Evaluation of the increased use of partial resection of large vestibular schwannomas: facial nerve outcomes and recurrence/regrowth rates. *Otol Neurotol*. 2013;34:1456-1464.
- Schroeder LD, Sjoquist DL, Stephan PE. *Understanding Regression Analysis*. Beveley Hills, CA: Sage; 1986.
- Chowdhury NI, Smith TL, Chandra RK, Turner JH. Automated classification of osteomeatal complex inflammation on computed tomography using convolutional neural networks. *Int Forum Allergy Rhinol*. 2019;9:46-52.
- Goggin LS, Eikelboom RH, Atlas MD. Clinical decision support systems and computer-aided diagnosis in otology. *Otolaryngol Head Neck Surg*. 2007;136:S21-S26.
- Dreiseitl S, Ohno-Machado L. Logistic regression and artificial neural networks classification models: a methodology review. *J Biomed Inform*. 2002;35:352-359.
- Goshtasbi K, Abouzari M, Djalilian HR. Are acoustic neuroma association patients characteristically different than the general population

- patients? The benefits and drawbacks of survey-based studies. *Otol Neurotol*. 2019;40:979-980.
42. Mala K, Sadasivam V, Alagappan S. Neural network based texture analysis of CT images for fatty and cirrhosis liver classification. *Appl Soft Comput*. 2015;32:80-86.
43. Chang P, Grinband J, Weinberg BD, et al. Deep-learning convolutional neural networks accurately classify genetic mutations in gliomas. *AJNR Am J Neuroradiol*. 2018;39:1201-1207.
44. Nasr-Esfahani E, Samavi S, Karimi N, et al. Melanoma detection by analysis of clinical images using convolutional neural network. *Conf Proc IEEE Eng Med Biol Soc*. 2016;2016:1373-1376.
45. Maki S, Furuya T, Horikoshi T, et al. A deep convolutional neural network with performance comparable to radiologists for differentiating between spinal schwannoma and meningioma. *Spine*. 2019;1. <https://doi.org/10.1097/BRS.0000000000003353>.
46. Harris AHS, Kuo AC, Weng Y, Trickey AW, Bowe T, Giori NJ. Can machine learning methods produce accurate and easy-to-use prediction models of 30-day complications and mortality after knee or hip arthroplasty? *Clin Orthop Relat Res*. 2019;477:452-460.
47. Hopkins BS, Yamaguchi JT, Garcia R, et al. Using machine learning to predict 30-day readmissions after posterior lumbar fusion: an NSQIP study involving 23,264 patients. *J Neurosurg Spine*. 2019;1-8. <https://doi.org/10.3171/2019.9.SPINE19860>.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Abouzari M, Goshtasbi K, Sarna B, et al. Prediction of vestibular schwannoma recurrence using artificial neural network. *Laryngoscope Investigative Otolaryngology*. 2020;5:278–285. <https://doi.org/10.1002/lio2.362>