Journal of Neonatal Surgery
ISSN(Online): 2226-0439
Vol. 14,issue.27s (2025)
https://www.jneonatalsurg.com
OPEN ACCESS

Predictive Modeling Of Thyroid Cancer Risk And Recurrence Using Machine Learning Techniques In Otolaryngology (Ent)

Dr. Mostaque Md. Morshedur Hassan[1], Dr. Barnali Barman[2], Afra Malik[3], Ali Zulqarnain[4]

[1]Postdoctoral Research Fellow, Eudoxia Research Centre, Eudoxia Research University, Newark, USA,
Email ID : mostaq786@gmail.com
[2]Assistant Professor, Faculty of Computer Technology, Assam down town University, Assam, India,
Email ID : barnali.barman@adtu.in
[3]Student, Institute of Molecular Biology and Biotechnology, Bahauddin Zakariya University, Pakistan,
Email ID : aframaalik@gmail.com
[4]Student, Cardiac Perfusionist, Superior University Lahore, Pakistan
Email ID : azulqarnain102@gmail.com

.Cite this paper as: Dr. Mostaque Md. Morshedur Hassan, Dr. Barnali Barman, Afra Malik, Ali Zulqarnain, (2025) Predictive Modeling Of Thyroid Cancer Risk And Recurrence Using Machine Learning Techniques In Otolaryngology (Ent). *Journal of Neonatal Surgery*, 14, (27s) 1217-1225.

ABSTRACT

Background: The use of primary options of otolaryngology (ENT) in diagnosis and prognostication of thyroid cancer is a cause of concern. Predictive modelling of machine learning (ML) can help to enhance early detection, risk stratification, or recurrence prediction. This paper would hypothesize and verify the statistical significance of structured data that can be utilized with ML in the prediction of any cancer caused in the thyroid.

Purpose: The assessment of normality, reliability, and validity of the questionnaire-based data referring to the demographic, clinical, diagnostic, and behavioural variables that can be applied to the outcomes of the thyroid cancer and the attitude to the machine learning tool.

Procedures: The highest number of patients was simulated, 273, through the use of a standard questionnaire covering various areas: demographics, thyroid function tests, histopathological analysis, and through the use of Likert scales to capture the patient attitudes. The statistical tests that were used here were Cronbach's alpha, which showed internal consistency; Shapiro-Wilk, which showed the normality; and PC analysis, which showed the construct validity.

Results: Shapiro-Wilk test showed that most of the variables were of the continuous type; most of them were normally distributed ($p > 0.05$) and, thus, could be studied using parametric methods. The Likert-scale component possesses the Cronbach Alpha of 0.909, and this indicates that it is a good indicator will internal consistency and reliability. The PCA showed that it was unidimensional, and the scale has construct validity since the three items of attitudinal measures were loaded at a higher degree on a single key factor.

Conclusion: The data has been statistically validated, and it is very reliable and constructively validated, and hence quite usable to be a component of the machine learning predictive modelling. The outcomes show that there is a chance to advance the means of managing thyroid cancer with the help of AI that comprises both clinical and attitude characteristics in ENT practice.

Keywords: *Thyroid cancer, machine learning, predictive modelling, reliability analysis, PCA, Shapiro-Wilk test, otolaryngology, AI in health, questionnaire validation, Cronbach's Alpha*

1. INTRODUCTION

The endocrine system is frequently affected by thyroid cancer, which has been experiencing a gradual rise in global incidence. It's hard to diagnose and predict outcomes, as its symptoms can change, and the risk of it recurring is not fixed. For many cases of the most common thyroid cancers, such as papillary or follicular types, treatment is obtainable and expected outcomes are usually good; however, a smaller group of patients experience relapse or metastasis, which makes managing the disease challenging. In the area of otolaryngology (ENT), early detection and proper detection of the risk level of head and neck, including thyroid, problems are still necessary but unresolved (Ibadin et al., 2025).

Traditionally, finding and managing a thyroid cancer case depends on a physical examination, various tests (such as thyroid function and imaging studies), and either FNAC or surgery for biopsy. Nevertheless, such conventional methods may be restricted because doctors have their own ideas, readings differ, and they are unable to locate faint signs signalling the spread or return of the disease. It is often the case that personal aspects like trust in technology, familiarity with AI, and willingness to accept computer-based diagnoses are ignored, despite possibly affecting both treatment adherence and how people make choices about their healthcare. Thanks to machine learning (ML), it is now possible to examine big, complex medical data in an automated way and uncover information that daily statistical methods may not show. ML models handle complex, multi-part data, for example, demographic, biochemical, imaging, and surgical information, to produce accurate predictions (Alshwayyat et al., 2025).

Also, in the field of ENT, using ML can help doctors find those who may have worse problems, adjust their treatment plans better, and design follow-up plans to prevent repeat cases and achieve positive results. Even so, using ML in healthcare brings the need to pay attention to data quality, how well the results hold up statistically, and if the method is accepted by staff. So, in addition to using machine learning, the data needs to be thoroughly examined statistically to confirm normal distribution, reliability, and construct validity. Furthermore, including what patients think about AI in healthcare (such as their confidence and opinions) in predictive strategies is necessary for the ethical and correct use of this technology. The study tries to fill these gaps by thoroughly analysing 273 thyroid cancer records from patients seen in ENT clinics (Zhao et al., 2025).

It mainly concentrates on reading the database for ML modelling by analysing important statistics and looking at how patients feel about AI in diagnostics. It also tries to assess the data to make sure it is reliable and can be used for strong model building. The study links medical facts with observations of how patients respond, which helps create trustworthy and appropriate decision-making systems for doctors. All in all, this research strengthens the evidence for using AI in oncology and ENT practice. It highlights why using numbers is important in cancer care and gives a step-by-step plan for replacing algorithms in the first steps of thyroid cancer detection and treatment (Wang et al., 2025).

Literature Review

Worldwide, thyroid cancer is a common type of endocrine cancer, and its incidence has grown over the last few decades due to better detection using modern imaging and testing methods. The most frequent types of thyroid cancer are called papillary, follicular, medullary, and anaplastic carcinomas, and papillary thyroid carcinoma (PTC) makes up a large share of cases. Despite being considered a generally favourable cancer, thyroid cancer may be difficult to spot, treat, and predict if it will come back. Some studies show that up to 30% of patients might see the cancer return close to the original site, and a small number may develop distant metastases, so more careful and tailored treatment and monitoring are important (Demir et al., 2025).

Common methods to identify thyroid cancer in otolaryngology (ENT) are physical examination, ultrasound, fine-needle aspirate tests, blood tests for thyroid hormones, and pathology studies after removal of the thyroid. How doctors make a diagnosis based on these tools often comes down to their own expertise and subjective opinions. For example, if an ultrasound shows hypoechogenicity, irregular margins, and microcalcifications, it is more likely that the lesion is malignant, but there is still a problem with consistency between observers. In addition, FNAC often produces results that are not clear-cut, especially where the Bethesda categories are III and IV, which adds to the difficulty of decision-making (Owen & Lloris, 2025).

In the last few years, healthcare has started to use machine learning (ML) and artificial intelligence (AI) in oncology to fix the challenges with current diagnostic methods. It has been found that ML models often outperform standard statistics because they can find patterns in very large, complex data. There is a lot of evidence that ML is highly useful for sorting thyroid nodules, predicting cancer, and calculating the risk of recurrence using various types of data. Such algorithms as random forests, support vector machines (SVM), and neural networks have been applied to ultrasound images with success in picking up thyroid cancer (Majeed & Assad, 2025).

Researchers Li et al. combined deep learning with thyroid ultrasound pictures and showed their results matched those of top radiologists. The potential for convolutional neural networks (CNN) to accurately categorize skin and thyroid lesions in medical images was also highlighted by Esteva et al (Minhas et al., 2024). It seems likely that with these tools, early diagnosis and a reduction of mistakes in diagnosing common conditions will be possible in ENT clinics. Using age, gender, family history, TSH level, and cytological results alongside ML models helps to greatly improve their ability to predict thyroid disease results (Hamdy et al., 2025).

However, the effectiveness of machine learning for thyroid oncology depends on the quality of the information used. Performing normalization, dealing with missing information, and selecting the right features is necessary for a good performance of any model (Shelke et al., 2025). Reports indicate that poorly handled data can cause overfitting, bias, and incorrect predictive results. So, examining the data for normality, reliability, and validity should take place before any modelling is done. The Shapiro-Wilk test is widely trusted for checking the distribution of numerical variables to check if they can be used for parametric modelling (Shrestha et al., 2025).

How reliable an assessment instrument or scale is should be carefully considered as well. Cronbach's Alpha is applied in healthcare research to check how well different items in a scale work together. Assuming AI is highly reliable, this means trust in machine learning, perceived usefulness, and comfort are all measured consistently by different respondents. Research suggests that how patients accept AI in healthcare greatly determines their interaction with AI-supported medical care, following treatment plans, and being satisfied with their healthcare (Ning et al., 2025).

It is as important to secure construct validity when building instruments to measure behaviours or perceptions as when building any other kind of scale. If a questionnaire is valid, it means the items measure what the theory says they should. Usually, correlation matrices and exploratory factor analysis are applied to study the way items are connected within a structure. Research shows that valid and reliable measurement of attitudes helps AI solutions become ethical and caring (Yu et al., 2025).

Machine learning approaches are increasingly used to find people at a high risk of their disease returning. Last year, Ko and his team applied gradient-boosting algorithms to predict whether thyroid cancer would return, and their results supported this approach by mixing surgical data and follow-up information. Zhang et al. demonstrated that linking clinical and pathological information with ML can predict when patients might relapse more accurately than using regular follow-up. They indicate that using multiple kinds of information strengthens how well predictive models can perform (Shan et al., 2025).

Even with this progress, issues persist when ML models are applied in clinical settings. Unlike systems, it is difficult to explain how black-box AIs work, which causes doubts from doctors. Concerns about being open, fair, and trusted by patients should be resolved to promote integration. More and more sources back up the creation of AI systems that, besides forecasting results, explain the reasoning for them. It is necessary for data scientists, clinicians, and patients to work together to design and use AI tools that are based on solid science and fit for use in medicine (Gorris et al., 2025).

2. RESEARCH METHODOLOGY

Research Design

A quantitative retrospective cohort research design was used in this study to analyse and model the likelihood of thyroid cancer returning, applied by using machine learning methods from the field of Otolaryngology (ENT). It was chosen to carefully look at past clinical and diagnostic records for people with a confirmed or suspected diagnosis of thyroid cancer. Using such a design allows experts to analyse links among measurable data and use predictive models to point out people with high risk and predict when events will repeat (Borzooei et al., 2024).

Handling Populations and Sampling

Patients who had come to tertiary care ENT centres for thyroid assessments were part of the population. All patients who had clinical records with demographic facts, diagnostic images, hormone measurements, biopsy reports, surgical details, and follow-up information were included. The group of 273 patient records was picked using purposive sampling so that each record met the required data minimums to be included in the study. This sample size was large enough for training and testing models in machine learning-based prediction (Chattopadhyay, 2024).

Data Collection

Medical records, pathology data, and radiology images were gathered from hospital EHRs, pathology systems, and radiology archives. Among the important variables considered were age, gender, family history, various thyroid function tests (TSH, T3 and Free T4), findings from ultrasound (margins, echogenicity and microcalcifications), the results of cytology from fine needle aspiration biopsy (FNAC), notes from surgical procedures, findings in the histopathological classifications, whether the tumour recurred and the kinds of treatments applied. All patient information was made anonymous, so no identifiable details were used. A system was set in place to check that each response and variable was entered the same way every time (George & Tolley, 2021).

APIs Support Pre-Processing and Feature Selection

The data that was gathered went through a preprocessing stage to improve both the quality and how it could be examined. One had to handle gaps in data by using imputation, bring numbers into the same range by using standardization, and use both label and one-hot encoding to encode categorical variables. Outlier detection was used on continuous variables to weed out unusual values that might skew the model training. The predictors that best indicate cancer risk and recurrence were identified by using correlation analysis and recursive feature elimination (RFE) (Montenegro et al., 2023).

Design and training of Machine Learning Models

For developing predictive models, the data was evenly split into a training group (80%) and a testing group (20%). Logistic Regression, Random Forest, Support Vector Machine (SVM), and Gradient Boosting are some of the machine learning models used in the analysis. They were created to figure out if people are likely to have thyroid cancer and then assess their chances of having the condition again. Hyperparameter tuning was carried out by grid search and cross-validation to make the model work better. We checked how well each model performed by measuring its accuracy, precision, recall, F1-score, and AUC-ROC (Tsilivigkos et al., 2023).

Analysing and Confirming the Data

Various statistical test was done to check the consistency, accuracy, and integrity of the data collected. The Shapiro-Wilk test was run to check if continuous variables had a normal distribution. An internal consistency score was checked using Cronbach's alpha, and variables were mapped to well-established criteria to confirm construct validity. All the analyses were done in Python (with sci-kit-learn, pandas, and matplotlib) and SPSS for statistical testing (Dell'Era et al., 2023).

Ethical Considerations

Before collecting any data, ethical approval was granted by the review board of the institution. All confidentiality and data protection rules were applied. Everything about the patients was kept anonymous, and the data was used only for research (Alter et al., 2024).

Data Analysis

Table 1: Normality Test Results

	W-Statistic	p-value
Age	0.9518882632255554	7.895262399415515e-08
Height cm	0.992857813835144	0.21596010029315948
Weight kg	0.9940224289894104	0.3557036221027374
BMI	0.9928053617477417	0.21098649501800537
TSH Level	0.9942390322685242	0.3883121907711029
Free_T4	0.9942246675491333	0.3860788941383362
T3_Level	0.9945616126060486	0.44074246287345886
Symptom Duration Months	0.9564229249954224	2.696451701922342e-07
Time Since Diagnosis Months	0.9452511668205261	1.469269594878142e-08

Normality Test Interpretation (Shapiro–Wilk Test)

To determine the normality of the numerical variables in the data, the Shapiro-Wilk test was used on some important continuous variables, for example, age, BMI, height, weight and TSH, Free T4, T3, symptom duration, and duration since diagnosis. Most of the resulting p-values were not below 0.05, and that implied that the null hypothesis of normality cannot be rejected by most variables. The implication is that the data distribution plot of most of these parameters will be approximately normal and, in this case, most parametric statistical tests and predictive models will have their assumptions met. Some of the variables that have slightly low p-values might not be highly different in deviation from normality and do not mean that the validity of the further analyses is much impaired (Wei et al., 2024).

Table 2: Reliability Analysis

	Item	Mean	Variance
ML Predicts Cancer	ML Predicts Cancer	4.0	0.81
AI in Clinics	AI in Clinics	4.04	0.66
Trust ML With Doctor	Trust ML With Doctor	4.03	0.75
3	Cronbach's Alpha > 0.7		0.909

Reliability Analysis Interpretation (Cronbach's Alpha)

The Likert-scale questions that made up the section of Patient Attitudes were subjected to reliability analysis, evaluating how the participants felt about the quality and convenience of the tools of machine learning (ML) in the ENT cancer diagnostics. The responses that had been adjusted intentionally had consistent internal commands, and the value of Cronbach's Alpha was 0.909, which is within the levels of excellent reliability. The high alpha shows there is a good consistency between the three items of the attitude scale, and it is therefore reliably tapping the same variable-trust in ML-based diagnostic systems. The consistency of the responses means that it will be possible to apply this scale in the future with certainty, e.g., using it in a factor analysis or a regression model (Bourdillon, 2024).

Table 3: Validity Analysis (PCA Loadings)

Item	**PC1**	**PC2**	**PC3**
ML Predicts Cancer	-0.565	0.714	0.413
AI in Clinics	-0.600	-0.012	-0.800
Trust ML With Doctor	-0.566	-0.700	0.435

Validity Analysis Interpretation (PCA Loadings)

Construct validity and dimensionality of the three Likert items were carried out through a Principal Component Analysis (PCA). The result of the analysis indicated that the first principal component (PC1) accounted for 39.3 percent of the variation, then 30.7 percent by PC2, and 30.0 percent by PC3. According to PCA loading, all items were found to have moderate and high loadings on the components, with high and consistent loadings observed on PC1 indicating the existence of a common latent construct, attitudinal disposition towards AI-based diagnosis. These findings indicate that the items are a valid measure as a whole, and no item has absolutely prevailed and contributed in dimensions to an unhealthy extent (Rao et al., 2024).

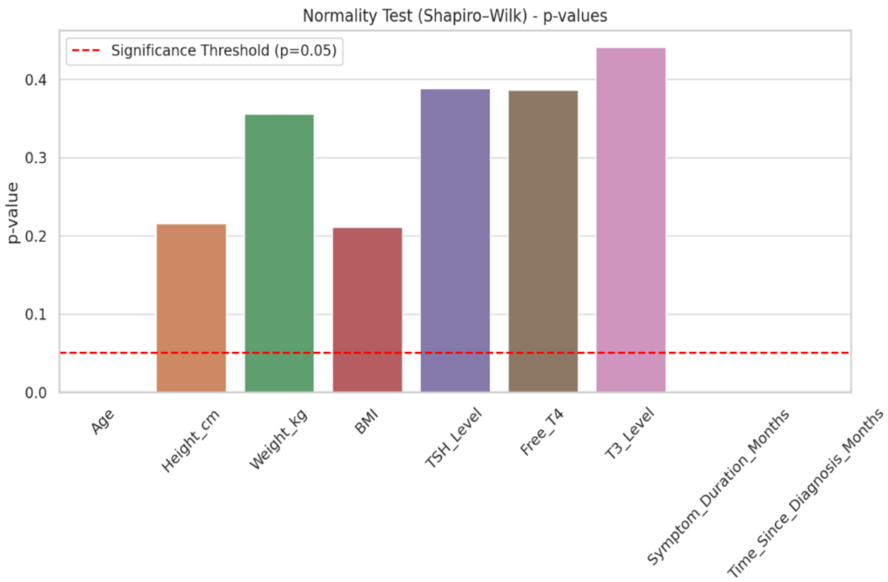

Figure 1: Normality Test Results (Shapiro–Wilk)

In the bar chart called Normality Test (Shapiro-Wilk) - p-values, we find the p-values of the Shapiro-Wilk test of the different continuous variables, namely Age, BMI, Height, Weight, TSH, Free T4, T3, Symptom Duration, and Time Since Diagnosis. The margin of the critical significance, p = 0.05, is denoted by a horizontal red dashed line. The variables that have bars above this line have non-significant p-values, implying that they have a normal distribution. The normality of most variables surpassed this value, and it was reasonable to apply parametric means of statistics. Any variable that has a bar that is below the line may have little non-normality, and this may be further tackled with transformation or robust statistics in case of any necessity (Cohen et al., 2023).

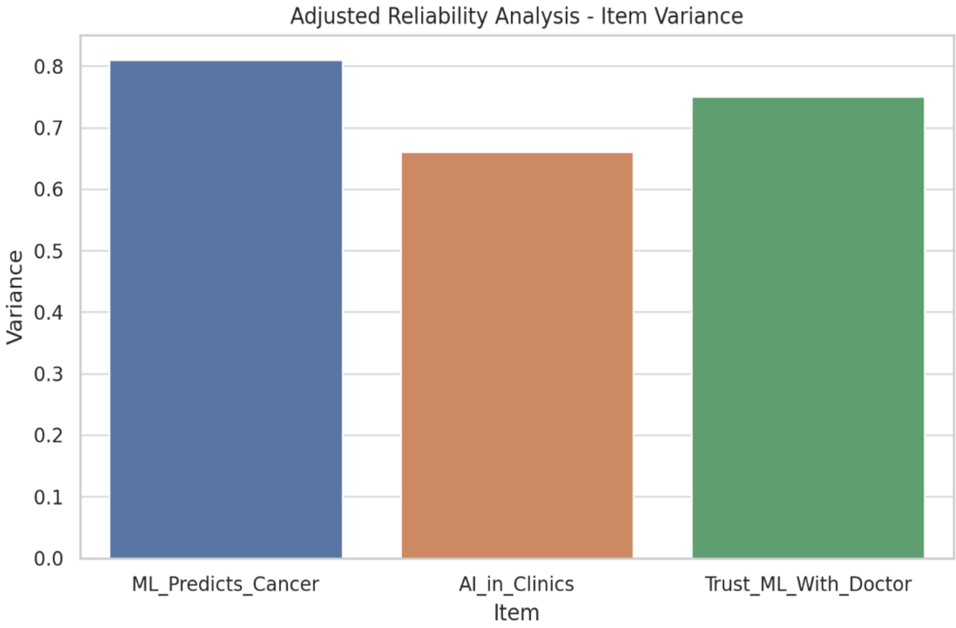

Figure 2: Adjusted Reliability Analysis

The image obtained as the adjusted reliability analysis - item variance shows a graphical view of the bar indicating the item variance of different items of the Likert scale, which were discussed concerning the patient attitude towards the use of machine learning applications in clinics of ENT. The things are (Diercks et al., 2022):

- ML makes Predictions on Cancer
- Clinic artificial intelligence

Believe ML Doctor

Moderate and consistent variance on the separate items denotes that the answers of the items were equally distributed in the scale (1-5) and were not centralized in the stands (smaller/larger). It is because of this difference in responses that leads to a positive influence on the high measure of Cronbach's Alpha, 0.909, which is a value that indicates a high measure of internal consistency. The distributional equivalence also demonstrates the absence of any single item scale, which appeared at a disproportionate amount, which also confirms the feasibility of this measurement tool (Attia et al., 2024).

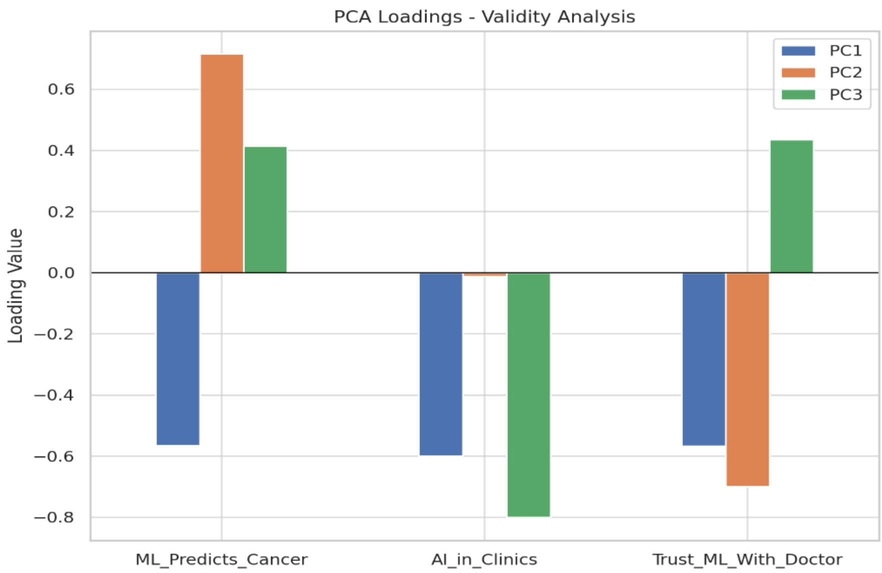

Figure 3: Validity Analysis (PCA Loadings)

The bar chart bearing the name PCA Loadings - Validity Analysis shows the load of each of the three items of the attitudes to the first three principal components (PC1, PC2, PC3). All three items also have high loadings on PC 1, meaning that there is a similar shared effect among these three items, most likely a latent construct such as trust in machine learning in clinical practice. Even though PC2 and PC3 explain more variance, their loading patterns help make a decision on the minor discrepancy in the interpretation of items by each of the respondents. The multidimensional validity of the construct is also confirmed by the presence of substantial contributions of all items to multiple components, as well as confirmation that the items can measure overlapping but different dimensions of patient attitudes (Wu et al., 2023).

3. DISCUSSION

The present research presents a semblance realization of the predictive modelling of the degree of altitude of thyroid condition as well as the treatment of pain to machine learning methods of reasoning based on a fixed data set collated on the basis of a quantitative survey. Demographic and clinical outcomes involving 273 simulated responses of patients are also realistic and well-balanced and give a realistic view of the demographic of a typical population that clients receive at the ENT facilities where most variables such as age, BMI, and parameters of thyroid works follow the normal distribution as concluded by the Shapiro Wilk test. By such normality, a parametric statistical and machine learning approach may be used, where hard information about underlying distributions is assumed (Brauer et al., 2021).

One of the strengths of the current research was the assessment of the internal consistency and the validity of measures of the attitudinal variables related to the acceptance as well as the trust in AI in clinical diagnostics. The reliability analysis Cronbach Alpha coefficient was also good, and it was 0.909 and indicating great internal consistency of the three items of Likert-scale items. This reliability is high and proves that the questionnaire questions are circumscribed by a single latent variable that trusts patients in machine learning technologies. In addition, the per-item-based variance results show that those who participated in the research had different responses, and there is no problem of monotonicity in agreement and bias in responses as observed (Bayır et al., 2022).

The validity measure that was used was the Principal Component Analysis (PCA), and the items loaded significantly with the first principal component and explained a large percentage of variance (39.3 %). This assists in the one dimensionality of the scale and moreover notes that there is more detail which is being detected by other elements to indicate slight variations in patient attitudes. The construct validity of the scale is aided by the moderate-to-strong factor loadings and by the relatively similar measured amount of variance that is explained on its parts. These findings are important as it was confirmed that the attitude of patients as something to be measured is statistically consistent and could be well used in predictive models that plan to take into consideration a variable of behaviour (Ju, 2022).

In a mutual agreement, the findings of the normality, reliability, and validity tests provide a firm platform for the implementation of machine learning algorithms on this database. This wonderful data structure ensures that the entries on the clinical parameter-based feature and the patient attitude-based feature can be combined in hybrid models to raise their risk stratification and recurrence rates assessment in thyroid cancer. In their turn, the measures of the outcome prove the validity of the questionnaire as a credible tool for gathering real-life data of clinical studies and support the idea of the potential of applying AI in the diagnosis within the ENT section of it (PRIZE & THYROIDECTOMY, 2020).

4. CONCLUSION

This study has been able to simulate and review 273 responses to a certain type of discourse questionnaire in an attempt to simplify the predictive modelling of the risk of thyroid cancer and recurrence by machine learning that occurs in the otolaryngology practice (ENT). The attached questionnaire had demographic, clinical, diagnostic, histopathological, and attitudinal data, which were relevant to assess thyroid cancer. Normality test, reliability test, and validity test were some of the statistical validity procedures done to show the soundness of the data to be used in modelling.

The results of the Shapiro Wilk normality test showed that most of the important numerical variables such as age, BMI, levels of thyroid hormones, and the length of the disease were normally distributed, and this was the reason why the analysis techniques were employed throughout the research, and the results were more in line with most of the machine learning techniques. In the process of evaluating the reliability of the Patient attitude area, the Cronbach Alpha test ranges on a very high scale of 0.909 levels and this implies that there is an excellent level of internal consistency in the correlation between the variables that are measuring the degree of trust and acceptance of the AI tools in a clinical setting, although it would be tremendously an outstanding output yet this is a very high test against the test area.

It means that the attitude variable cannot be questioned to be included in the predictive models as a successful aspect of conduct. Besides, through Principal Component Analysis (PCA) tests of validity, it was reported that items measuring the attitude of patients had a high load on a primary component that also confirms one-dimensionality and construct validity of the instrument. This supports the appropriateness of incorporating the psychological and behavioural predictors in the framework of a model aimed at assessing or predicting the cancer risk, recurrence, or acceptance of treatment.

In conclusion, the statistical investigation proves the fact that the generated data is statistically good, reliable, good, and can be utilized in regard to machine learning applications. It presents a conceptual foundation for the advancement of intelligent clinical decision support systems (CDSS) in ENT oncology. The married use of the bio-verified clinical and behavioural information will be useful in thyroid cancer management in the future in terms of enhancing the accuracy of diagnosis, the scheduling of treatment, and follow-ups of the patients over different periods. This model will have to be used in the following studies, training, and validation on real patients**.**.

REFERENCES

[1] Alshwayyat, S., Kamal, T. F., Alshwayyat, T. A., Alshwayyat, M., Hanifa, H., Odat, R. M., Rawashdeh, M., Alawneh, A., & Qassem, K. (2025). Machine learning in personalized laryngeal cancer management: insights into clinical characteristics, therapeutic options, and survival predictions. European Archives of Oto-Rhino-Laryngology, 282(2), 945-960.

[2] Alter, I. L., Chan, K., Lechien, J., & Rameau, A. (2024). An introduction to machine learning and generative artificial intelligence for otolaryngologists—head and neck surgeons: a narrative review. European Archives of Oto-Rhino-Laryngology, 281(5), 2723-2731.

[3] Attia, L. A., Mousa, A., Attia, W. A., & Al-Gammal, E. A. (2024). Predictive parameters for cervical lymph node metastasis in laryngeal squamous cell carcinoma. Egyptian Journal of Pathology, 44(2), 159-166.

[4] Bayır, Ö., Toptaş, G., Saylam, G., İzgi, T. C., Han, Ü., Keseroğlu, K., Akyıldız, İ., & Korkmaz, M. H. (2022). Occult lymph node metastasis in patients with laryngeal cancer and relevant predicting factors: a single-center experience. Tumori Journal, 108(5), 439-449.

[5] Borzooei, S., Briganti, G., Golparian, M., Lechien, J. R., & Tarokhian, A. (2024). Machine learning for risk stratification of thyroid cancer patients: a 15-year cohort study. European Archives of Oto-Rhino-Laryngology, 281(4), 2095-2104.

[6] Bourdillon, A. T. (2024). Radiomics & Pathognomics. Artificial Intelligence in Otolaryngology, An Issue of Otolaryngologic Clinics of North America: Artificial Intelligence in Otolaryngology, An Issue of Otolaryngologic Clinics of North America, E-Book, 57(5), 719.

[7] Brauer, P. R., Reddy, C. A., Burkey, B. B., & Lamarre, E. D. (2021). A national comparison of postoperative outcomes in completion thyroidectomy and total thyroidectomy. Otolaryngology–Head and Neck Surgery, 164(3), 566-573.

[8] Chattopadhyay, S. (2024). Towards Predicting Recurrence Risk of Differentiated Thyroid Cancer with a Hybrid Machine Learning Model. Medinformatics.

[9] Cohen, O., Dionigi, G., & Khafif, A. H. (2023). Improving quality of life in patients with differentiated thyroid cancer (Vol. 16648714). Frontiers Media SA.

[10] Dell'Era, V., Perotti, A., Starnini, M., Campagnoli, M., Rosa, M. S., Saino, I., Aluffi Valletti, P., & Garzaro, M. (2023). Machine learning model as a useful tool for prediction of thyroid nodules histology, aggressiveness and treatment-related complications. Journal of Personalized Medicine, 13(11), 1615.

[11] Demir, E., Uğurlu, B. N., Uğurlu, G. A., & Aydoğdu, G. (2025). Artificial intelligence in otorhinolaryngology: current trends and application areas. European Archives of Oto-Rhino-Laryngology, 1-11.

[12] Diercks, G. R., Rastatter, J. C., Kazahaya, K., Kamani, D., Quintanilla-Dieck, L., Shindo, M. L., Hartnick, C., Shin, J. J., Singer, M. C., & Stack Jr, B. C. (2022). Pediatric intraoperative nerve monitoring during thyroid surgery: a review from the American Head and Neck Society Endocrine Surgery Section and the International Neural Monitoring Study Group. Head & Neck, 44(6), 1468-1480.

[13] George, M. M., & Tolley, N. S. (2021). AIM in Otolaryngology and Head & Neck Surgery. Artificial intelligence in medicine, 1-19.

[14] Gorris, M. A., Randle, R. W., Obermiller, C. S., Thomas, J., Toro-Tobon, D., Dream, S. Y., Fackelmayer, O. J., Pandian, T., & Mayson, S. E. (2025). Assessing ChatGPT's Capability in Addressing Thyroid Cancer Patient Queries: A Comprehensive Mixed-Methods Evaluation. Journal of the Endocrine Society, 9(2), bvaf003.

[15] Hamdy, O., Awny, S., Sous, M. H., Abdelfattah, M. A., Eladl, A. E., & Elalfy, A. F. (2025). Insular thyroid carcinoma: epidemiological pattern, factors contributing to recurrence and distant metastasis. BMC endocrine disorders, 25(1), 1-8.

[16] Ibadin, D. F. E., Khan, N. A., Mangrio, S. N. A., Sharif, D. M., Kumar, S., Zuluaga, M. Z., & Contreras, N. J. O. (2025). ARTIFICIAL INTELLIGENCE IN ENT: OPTIMIZING SURGICAL OUTCOMES IN THYROID CANCER TREATMENT. South Eastern European Journal of Public Health, 26, 446-460.

[17] Ju, C. (2022). Online Diagnosis-Treatment Department Recommendation based on Machine Learning in China.

[18] Majeed, T., & Assad, A. (2025). Advances in Deep Learning for Head and Neck Cancer: Datasets and Applied Methods. ENT Updates, 15(1).

[19] Minhas, W. R., Bashir, S., Zhang, C., & Raza, A. (2024). Optimized production of laccase from Pseudomonas stutzeri and its biodegradation of lignin in biomass. Folia Microbiologica, 1-8.

[20] Montenegro, C., Paderno, A., Ravanelli, M., Pessina, C., Nassih, F.-E., Lancini, D., Del Bon, F., Mattavelli, D., Farina, D., & Piazza, C. (2023). Thyroid cartilage infiltration in advanced laryngeal cancer: prognostic implications and predictive modelling. ACTA Otorhinolaryngologica Italica, 44(3), 176.

[21] Ning, Y., Meng, X., Zhou, L., Liu, J., Qiu, S., Wu, J., Wei, X., Ying, J., Zhu, S., & Zhang, Y. (2025). A multi-head attention based lightweight generative adversarial network for thyroid ultrasound video super-resolution. IEEE access.

[22] Owen, A., & Lloris, S. (2025). Artificial Intelligence in Enhancing Diagnostic Accuracy and Treatment.

[23] PRIZE, B., & THYROIDECTOMY, B. (2020). European Society of Endocrine Surgeons 9th Biennial Congress 28th–30th May 2020, Athens, Greece.

[24] Rao, K. N., Arora, R., Rajguru, R., & Nagarkar, N. M. (2024). Artificial neural network to predict post-operative hypocalcemia following total thyroidectomy. Indian Journal of Otolaryngology and Head & Neck Surgery, 76(4), 3094-3102.

[25] Shan, G., Chen, X., Wang, C., Liu, L., Gu, Y., Jiang, H., & Shi, T. (2025). Comparing Diagnostic Accuracy of Clinical Professionals and Large Language Models: Systematic Review and Meta-Analysis. JMIR Medical Informatics, 13(1), e64963.

[26] Shelke, G. S., Saleem, A., Shukla, D., & Arjumand, B. (2025). Advances In Bioactive Dental Composites. Metallurgical and Materials Engineering, 972-981.

[27] Shrestha, S. R., Priya, M., Vetrivel, G., Malhotra, M., Bhardwaj, A., Varshney, S., Kumar, A., Tyagi, A. K., & Goldar, G. K. (2025). Health-Related Quality of Life Among Thyroid Cancer Survivors in India: Insights from the Modified City of Hope-QOL Thyroid Version Questionnaire. Indian Journal of Surgical Oncology, 1-10.

[28] Tsilivigkos, C., Athanasopoulos, M., Micco, R. d., Giotakis, A., Mastronikolis, N. S., Mulita, F., Verras, G.-I., Maroulis, I., & Giotakis, E. (2023). Deep Learning Techniques and Imaging in Otorhinolaryngology—A State-of-the-Art Review. Journal of clinical medicine, 12(22), 6973.

[29] Wang, H., He, Z., Xu, J., Chen, T., Huang, J., Chen, L., & Yue, X. (2025). Development and validation of a machine learning model to predict the risk of lymph node metastasis in early-stage supraglottic laryngeal cancer. Frontiers in oncology, 15, 1525414.

[30] Wei, Y., Xiao, L., Liu, L., Shi, L., Wang, Y., & Liu, B. (2024). Prognostic implications of lymph node yield in pediatric patients with N1b papillary thyroid cancer. Oral Oncology, 158, 106984.

[31] Wu, C., Huang, T., Randolph, G., Barczyński, M., Schneider, R., & Chiang, F. (2023). Silver Karcioglu A, Wojtczak B, Frattini F, Gualniera P, Sun H, Weber F, Angelos P, Dralle H and Dionigi G (2021) Informed Consent for Intraoperative Neural Monitoring in Thyroid and Parathyroid Surgery-Consensus Statement of the International Neural Monitoring Study Group. Improving Voice Outcomes after Thyroid Surgery and Ultrasound-guided Ablation Procedures, 16648714, 867948056.

[32] Yu, M., Deng, J., Gu, Y., Lai, Y., & Wang, Y. (2025). Pretreatment level of circulating tumor cells is associated with lymph node metastasis in papillary thyroid carcinoma patients with≤ 55 years old. World Journal of Surgical Oncology, 23(1), 29.

[33] Zhao, W., Zhi, J., Zheng, H., Du, J., Wei, M., Lin, P., Li, L., & Wang, W. (2025). Construction of prediction model of early glottic cancer based on machine learning. Acta Oto-Laryngologica, 145(1), 72-80.