Research Article *Open Access*

Journal of Clinical Oncology & Advanced Therapy

ISSN: 3070-3786

DOI: doi.org/10.63721/26JCOAT0125

Quantifying Progression Along the Thyroid Cancer Severity Continuum Using Cumulative Logit Modelling

Francis Ayiah Mensah[1*], Emmanuel Harris[2] and Mohammed Frempong[1]

[1]Department of Mathematics, Statistics and Actuarial Science, Takoradi Technical University, Sekondi-Takoradi, Ghana
[2]Department of Statistics and Actuarial Science, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana

Citation: *Francis Ayiah Mensah, Emmanuel Harris, Mohammed Frempong (2026) Quantifying Progression Along the Thyroid Cancer Severity Continuum Using Cumulative Logit Modelling. J of Clin Onco & Adv Thpy 2(2), 01-12. WMJ/JCOAT-125*

Abstract

Background: *Most studies on thyroid cancer have been aimed at prediction of recurrence, survival analysis, and machine-learning classification models that are often binary or unordered, thus ignoring the ordering of disease severity. The purpose of this study was to explore factors associated with the severity of thyroid cancer in an ordinal cumulative logit model that maintains the natural order of disease progression. Also, to estimate stage-transition probabilities and test the ability of an ordinal severity model to predict disease severity.*

Methods: *This study was carried out by a retrospective analysis of data from 383 thyroid cancer patients that were retrieved from the UCI Machine Learning Repository. The likelihood-based model selection and ordinal cumulative logit regression methods were used. Models were compared for adequacy using threshold parameters, information criteria, likelihood ratio tests, predicted probabilities, and classification performance.*

Results: *The findings revealed that age significantly increased the odds of progression to a higher disease stage (OR = 1.132, 95% CI: 1.090-1.175, $p < 0.001$), while uni-focal tumours substantially reduced the odds of severe disease (OR = 0.106, 95% CI: 0.039-0.294, $p < 0.001$). The final ordinal model had a markedly better fit than the null model (AIC: 157.613, LogLik: −72.807) as compared to the null model (AIC: 265.048, LogLik: −130.524). The predicted probabilities indicated that the chances of having Stage I disease were 94.6% for low-risk patients and 53.7% for patients > 60 years of age, with advanced stages of disease having significantly higher probabilities as well.*

Conclusion*: This study proposes an ordinal modelling method based on severity index, which overcomes the information loss problem of binary and multinomial models, and succeeded in giving clinically interpretable estimates of the progression of thyroid cancer. To facilitate early intervention and risk-stratified management, enhanced age-centred surveillance and intensified monitoring of multifocal tumours are recommended.*

***Corresponding author**: Francis Ayiah Mensah, Department of Mathematics, Statistics and Actuarial Science, Takoradi Technical University, Sekondi-Takoradi, Ghana.

Submitted: 15.06.2026 **Accepted**: 20.06.2026 **Published**: 30.06.2026

Keywords: Thyroid Cancer, Disease Severity, Ordinal Cumulative Logit Model, Ordinal Regression, Cancer Staging, Risk Stratification

Introduction

Clinical, pathological, and staging information (such as age, sex, thyroid function, adenopathy, pathology, focality, Tumour Node Metastasis (TNM) classification, risk category, treatment response, and recurrence status) are commonly used in the management of thyroid cancer [1-4].While many thyroid cancers are low risk and have a good prognosis, the severity of the disease is widely variable between patients, and it is important to stage and risk stratify the disease accurately to ensure that treatment is appropriate, surveillance is undertaken and that recurrence is prevented [5-9]. Current guidelines focus on risk-adapted management and dynamic risk stratification, rather than only on static baseline classification.

The classification of thyroid cancer by stage and risk must be ordered, as these indicators reflect disease progression [10-13]. However, some studies classify these outcomes as dichotomous or use multinomial approaches without respecting the ordinal nature of the progression of cancer. These may decrease the statistical efficiency, lose clinically useful ordering and impair interpretation. Analysing the order of the stages of thyroid cancer, combined with demographics, clinical characteristics, and treatments, indicates that the ordinal cumulative logit model is appropriate for the understanding of the risk of advanced thyroid cancer [14-17]. Several models to predict thyroid cancer have been developed, including the estimation of the probability of recurrence, risk of death, and risk stratification based on regression analysis, TNM staging, or American Thyroid Association (ATA) risk classification [2,18,19].

The latter takes into account the fact that recurrence risk is associated with the size of the tumour, regional lymph nodes involvement, the development of the metastasis, pathologic and treatment features and, thus, recommends conducting regular risk stratification even after certain follow-up periods [3,20,21]. Literature reviews of differentiated thyroid cancer indicate that the conventional technique of risk stratification is still valid for the medical practice yet provides inaccurate results [22-24]. Although progress has been made in improving thyroid cancer prognosis modelling, there are still some major statistical gaps to be overcome [15,25-27].

The first is that many studies divide up order into stages and/or lose clinical information. Secondly, multinomial models might not consider the natural transition from Stage I to Stage IV. Third, before the next recurrence, there is limited knowledge regarding the drivers of disease severity, and binary models of disease recurrence are not able to capture these drivers [15,19,28-30]. Hence, this study is warranted because of the need for creating a severity-based ordinal scale for thyroid cancer which maintains the ordered nature of thyroid cancer stages. The goal is to employ an ordinal cumulative logit model to determine demographic, clinical, pathological, and treatment-response factors associated with increasing levels of severity of thyroid cancer disease to enhance the interpretability, statistical efficiency and clinical relevance of the model compared to traditional binary or unordered modelling methods.

Materials and Methods

Study Design and Data Source

The design of this study was retrospective, observational, and analytical, and the clinical dataset used in the study was a publicly available dataset of thyroid cancer taken from the UCI Machine Learning Repository. The dataset includes information for 383 thyroid cancer patients and was designed to aid in prognosis and thyroid cancer recurrence studies. In the present study, the authors re-analysed the data to study the factors determining the severity of thyroid cancer in an ordinal modelling framework, maintaining the natural order

of the disease stages. The data include demographic, clinical, pathological, therapeutic and outcome. Data are anonymised, and therefore, no ethical approval or informed consent was required.

Dataset Repository:

https://archive.ics.uci.edu/dataset/915/thyroid+cancer+recurrence

Ordinal Cumulative Logit Model

Model Specification

The principal analytical framework was the Ordinal Cumulative Logit Model, also known as the Proportional Odds Model. This approach was selected because thyroid cancer stage is an ordered categorical outcome and the model preserves clinically meaningful information regarding disease pr(gression.

Let Y_i denote the observed stage of patient i.

The cumulative probability of being at or below stage k is:

$$P(Y_i \leq k)$$

for

$$k = 1,2$$

The cumulative logit model is expressed as:

$$\log\left[\frac{P(Y_i \leq k)}{P(Y_i > k)}\right] = \alpha_k - \beta_1 X_{1i} - \beta_2 X_{2i} - \cdots - \beta_p X_{pi}$$

where:

- α_k represents the threshold parameter separating adjacent disease stages,
- β_j represents the regression coefficient associated with predictor j,
- X_ji denotes the value of predictor j for subject i,
- p denotes the number of predictors.

The negative sign follows the parameterisation adopted by the cumulative link model implementation in R.

Interpretation of Model Parameters

Exponentiation of regression coefficients yields cumulative odds ratios:

$$OR_j = \exp(\beta_j)$$

where:

- $OR_j >$ indicates increased odds of belonging to a more advanced disease stage,
- $OR_j < 1$ indicates decreased odds of progression toward severe disease.

For example, an odds ratio of 1.132 for age indicates that each additional year increases the cumulative odds of belonging to a higher disease stage by approximately 13.2%.

Maximum Likelihood Estimation

Model parameters were estimated using Maximum Likelihood Estimation (MLE).

The likelihood function is:

$$L(\beta) = \prod_{i=1}^{n} P(Y_i = y_i)$$

where:

- y_i denotes the observed disease stage for patient i.

The corresponding log-likelihood function is:

$$\ell(\beta) = \sum_{i=1}^{n} \log[P(Y_i = y_i)]$$

Parameter estimates were obtained through iterative numerical optimisation until convergence was achieved.

Threshold Parameters

One of the unique characteristics of ordinal regression is the estimation of threshold parameters that define boundaries between disease severity categories.

The threshold Parameters are:

Stage I | Stage II

and

Stage II | Stage III

These thresholds determine the latent severity levels required for transition from one disease stage to another. Statistically significant thresholds indicate that the disease stages represent distinct and meaningful categories.

Model Selection and Evaluation

Competing models were compared using information criteria and likelihood-based measures.

Akaike Information Criterion

$$AIC = -2\ell + 2p$$

where:

- ℓ is the maximised log-likelihood,

• p is the number of model parameters.

Bayesian Information Criterion

$$BIC = -2\ell + p\ln(n)$$

where:
• n denotes sample size.
Lower AIC and BIC values indicate superior model fit.

Likelihood Ratio Test

The overall significance of the final model was assessed using the likelihood ratio statistic:

$$LR = -2(\ell_0 - \ell_1)$$

where:
• ℓ_0 represents the log-likelihood of the null model,
• ℓ_1 represents the log-likelihood of the fitted model.
Under the null hypothesis, the statistic follows a Chi-square distribution.

Pseudo Coefficient of Determination

Model explanatory power was quantified using McFadden's pseudo-R^2:

$$R^2_{\text{McFadden}} = 1 - \frac{\ell_1}{\ell_0}$$

and Nagelkerke's pseudo-R^2:

$$R^2_{\text{Nagelkerke}} = \frac{1 - \exp\left[\frac{2}{n}(\ell_0 - \ell_1)\right]}{1 - \exp\left(\frac{2\ell_0}{n}\right)}$$

These measures quantify the extent to which the predictors explain variation in thyroid cancer severity.

Assessment of the Proportional Odds Assumption

The proportional odds assumption was evaluated using nominal effects tests and scale effects tests.
The nominal effects test examines whether predictor effects differ across cumulative logits, while the scale effects test evaluates whether predictor-specific variability changes across disease stages. A non-significant result indicates that the proportional odds assumption is satisfied and that a single cumulative odds ratio can adequately describe the effect of each predictor across all stage transitions.

Predicted Probabilities

Predicted probabilities for each disease stage were computed from the fitted ordinal model.
The probability of belonging to stage k i

$$P(Y_i = k) = P(Y_i \leq k) - P(Y_i \leq k - 1)$$

These probabilities provide individualised estimates of disease severity and facilitate clinical interpretation of risk.

Classification Accuracy

The predicted stage for each patient was determined using the maximum probability rule.

Classification accuracy was calculated as:

$$\text{Accuracy} = \left(\frac{\text{Number of Correct Predictions}}{\text{Total Number of Observations}}\right) \times 100$$

This measure evaluates the predictive performance of the ordinal model.

Results

Table 1: Ordinal Cumulative Logit Model Results

Predictor	Estimate	SE	Cumulative OR [95% CI]	p-value
age	0.124	0.019	1.132 [1.09, 1.175]	< .001
genderM	0.125	0.68	1.133 [0.299, 4.294]	0.855
smokingYes	1.176	0.731	3.241 [0.774, 13.572]	0.108
focalityUni-Focal	-2.24	0.517	0.106 [0.039, 0.294]	< .001

Table 1 presents the results of the ordinal cumulative logit model, which offers new insights into the factors associated with the severity of thyroid cancer, as the outcomes are ordered disease stages instead of just collapsing to the two categories: recurrence or no recurrence. This methodological advancement fills a significant gap in past research on thyroid cancer, which often has been limited to logistic regression, machine-learning classifiers, or prediction models based on recurrence [15,25,26]. A great strength of the present study is the maintenance of the intrinsic ordering of the progression of cancer that results in clinically meaningful cumulative odds ratios that directly represent the probability that a cancer will be at a more advanced stage at the time of diagnosis.

Each year of age was a highly significant factor for the severity of the thyroid cancer (OR = 1.132, 95% CI: 1.090 – 1.175, $p < 0.001$), with an increase in the cumulative odds of more advanced disease stage of approximately 13.2% per year of age. This conclusion confirms earlier studies showing that age is a key factor in predicting how differentiated thyroid cancer will progress. It is also one of the main criteria in the traditional staging systems of both the ATA and TNM systems [3,6,18]. The present study, however, differs from previous studies in that it directly analysed the effects of age on disease severity, and not just on mortality or recurrence, thus extending the knowledge in the field. This result aligns with current recommendations for surveillance strategies according to age and is a contribution to SDG 3 Good Health and Well-being by better management of risks.

Gender was not found to be statistically associated with disease severity (OR = 1.133, $p = 0.855$), indicating that if other clinical and pathological factors are taken into account, there was no significant difference between disease severity in patients of both sexes. Interestingly, other studies showed that the role of gender was not consistent after accounting for tumour characteristics [2,19]. In light of the current ordinal scale, the implication from the findings could be that progression based on one clinical factor, such as gender, cannot adequately help move through the stages of the disease. The cumulative OR for smoking was marginally higher (OR = 3.241, $p = 0.108$). Despite being broad, the confidence intervals do show some clinical significance of the association as the effect size was quite large. Research on the impact of smoking on the growth of thyroid cancer has yielded conflicting results, with some studies indicating that smoking indirectly impacts the growth of thyroid cancer [7].

One of the most unique discoveries is about the issue of focality. The odds of progression to advanced disease were significantly lower for patients with uni-focal disease than those with multi-focal disease (OR = 0.106, 95% CI: 0.039–0.294, $p < 0.001$). This suggests a reduction in cumulative odds of severe disease of ~89.4% when treating patients with uni-focal tumours. The multifocality has been repeatedly reported in previous literature as a risk factor for increased recurrence risk and aggressive disease behaviour [3,21]. The present study, however, supports this evidence by showing a direct link between the ordinal disease severity and the same evidence. This discovery further underscores the significance of multifocality evaluation in the planning of surgery and in follow-up of the post-surgical period.

Table 2: Estimated Threshold Parameters

Threshold	Estimate	SE	CI_Lower	CI_Upper	p-value
I|II	7.54	1.056	5.471	9.61	< .001
II|III	11.094	1.421	8.31	13.879	< .001

The estimated threshold parameters are discussed in Table 2. One of the unique features of ordinal regression models and one parameter rarely reported in thyroid cancer studies is the threshold parameter reported in Table 2. Estimates of 7.540 and 11.094 are well above the level of $p < 0.001$, suggesting good separation between the data points for the three severity categories. The disease stages are meaningful and statistically distinguishable stages of clinical disease and not arbitrary, as these threshold estimates confirm that the disease stages represent meaningful and statistically separable levels of clinical progression rather than arbitrary classifications. Threshold values represent the spectrum of severity that is not seen in the development of cancer and is asso-

ciated with thyroid cancer. Unlike the common binary logistic regression models employed in recurrence studies, this is not the same model. The progressive trends of increasing threshold values indicate that the more advanced the disease, the more contributions from risk factors are needed for a progression to the next stage of disease, highlighting the biological complexity of thyroid cancer progression [25,26]. This result further confirms the validity of the classification of clinical stages in current clinical practice and the suitability of the modelling framework of order.

Table 3: Model Fit Comparison

Model	AIC	BIC	LogLik
Null ordinal model	265.048	272.869	-130.524
Final stable ordinal cumulative logit model	157.613	181.078	-72.807

The results of the model comparison in Table 3 show that the ordinal cumulative logit approach has significantly improved. The final model reduced the AIC from 265.048 to 157.613 and improved the log-likelihood from −130.524 to −72.807. The improvements have a more pronounced disease severity pattern than the null model. The main limitation of previous studies on thyroid cancer was the lack of measurement of the model's performance in terms of prediction accuracy or classification power. They didn't consider the process order throughout the disease, something in the works of Borzooei et al [15,19].

The better model fit here indicates that there are significant statistical advantages in this model, which contains ordinal information. The study is also methodologically innovative as it shows the possibility of the ordinal severity models to better account for disease heterogeneity than traditional methods. Improved disease stratification has implications in precision medicine and for SDG 3 in the diagnostic and prognostic decision-making process.

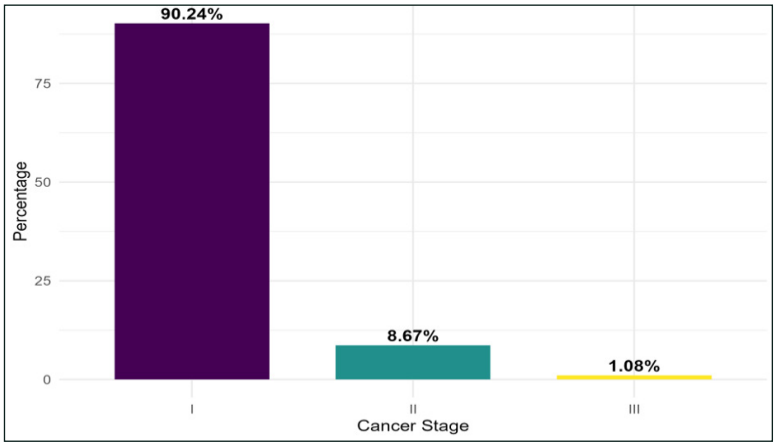

Figure 1: Distribution of Thyroid Cancer Stage

Figure 1 shows that around 90.24% were Stage I, only 8.67% were Stage II and 1.08% were classified as Stage III. The distribution is similar to the overall favourable prognosis and early detection rate of differentiated thyroid cancer [2,5]. The majority of the cases in Stage I are in line with previous epidemiological studies reporting a rise in the incidence of localised thyroid tumours with the improvement of imaging and surveillance. The present study, however, goes beyond the previous descriptive studies in that it uses this stage information as an ordered outcome variable, rather than as frequencies. It is a key methodological step as it allows us to identify factors associated with moving up or down the severity continuum. The results are supportive of ongoing investment in early detection and risk-stratified management programmes.

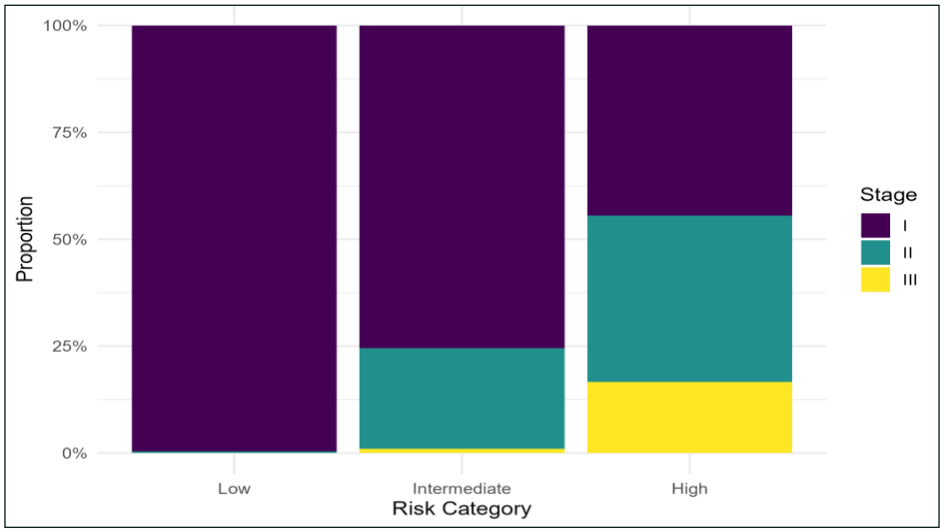

Figure 2: Cancer Stage Composition Across Risk Categories

As can be seen in Figure 2, there is a clear gradient between ATA risk category and severity of the disease. Stage I (Low) had significantly more patients who were at low risk, whereas Stage II (Intermediate) and Stage III (High) had significantly more patients at high risk. This finding aligns with the ATA's proposed risk classification scheme and corroborates other studies that have found a relationship between risk groups and poorer outcomes [3,20]. So, it fits with what other studies have found too. The current study builds on previous evidence, however, by quantifying this relationship in an ordinal model instead of binary models of recurrence. The results showed the effectiveness of the risk categories to capture progressive increases in the severity of the disease, which will be continued to use in clinical decision-making. The observed gradient also indicates potential areas for intervention for high-risk groups to prevent progression to more advanced stages of disease.

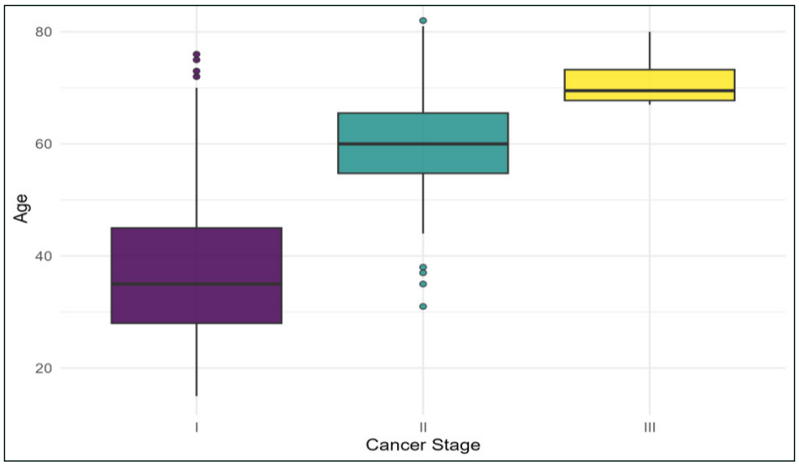

Figure 3: Age Distribution Across Thyroid Cancer by Stage

The age distribution's shown in Figure 3 graphically confirms the presence of age effect, as described in Table 1. Median age for each disease stage increased in a consistent manner, and was much higher for Stage III patients compared to Stage I patients. This profile is similar to the modern staging systems taking age into account as an important prognostic factor [3,18]. Previous studies have mainly correlated age to mortality and/or recurrence, but the current data show an association of age with the severity of the disease. This is very significant as it means that not only is there an effect of age on outcomes following diagnosis, but there is also an effect on when disease develops. The results of this study further emphasise the importance of more intense monitoring in the elderly and the potential value of age-specific approaches to management.

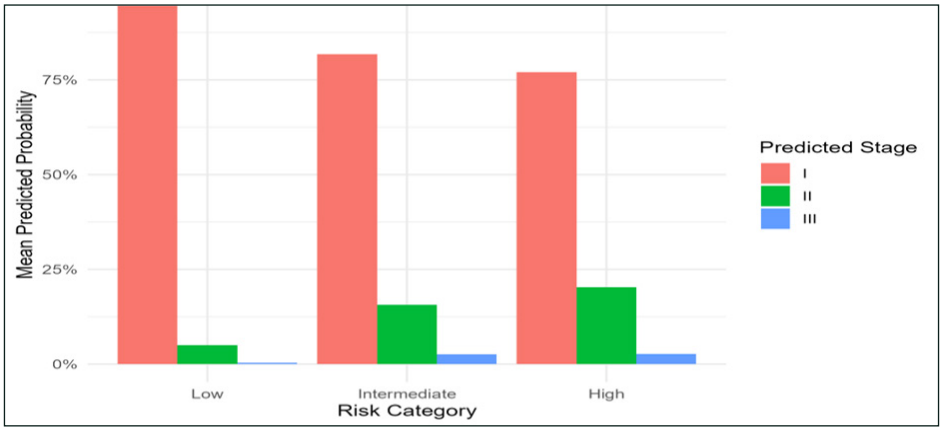

Figure 4: Predicted Probability of Disease Stage by Risk Category

As shown in Figure 4, the probabilities forecasted give new insights into the severity gradient of the risk categories. The percentage of patients expected to have Stage 1 disease was 94.6% for low-risk patients, 81.7% for intermediate-risk patients, and 77.0% for high-risk patients. The probabilities, on the other hand, of the higher grades rose progressively with the risk groups. The ordinal model produces clinically meaningful probabilities, which can be directly used for patient counselling and treatment planning, rather than acting as a "black-box" predictive model as in common machine-learning studies [15,19]. This probabilistic interpretation is an important step in precision oncology as it enables the physicians to estimate the probability of progression in a rank-ordered set of severity grades instead of just predicting for recurrence. While discussing Figure 5 address the students' expectations and describe the results. When discussing Figure 5, discuss students' expectations and what they saw.

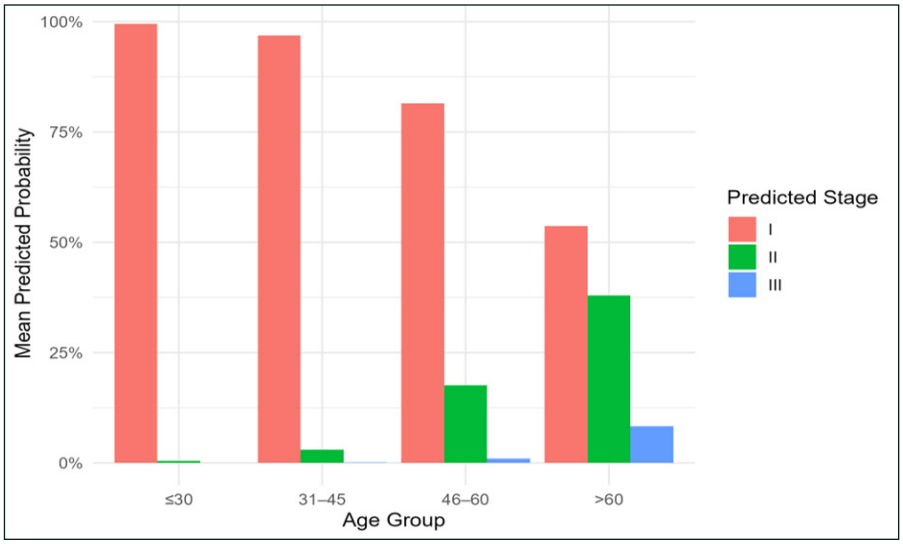

Figure 5: Predicted Probability of Disease Stage by Age Group

There is a clear age course-going pattern as revealed by Figure 5 and Table 12. Patients with age ≤30 years had a predicted probability of 99.5% of Stage I disease, while patients >60 years of age had a predicted probability of just 53.7% for Stage I disease and significantly higher probabilities for Stage II (38.0%) and Stage III (8.4%) disease. The results strongly support the importance of age for progression of thyroid cancer in a quantitative manner. While earlier research suggested that age was a prognostic indicator, this study is among only a few to provide estimates of age-based probabilities of disease staging through ordinal regression modelling.

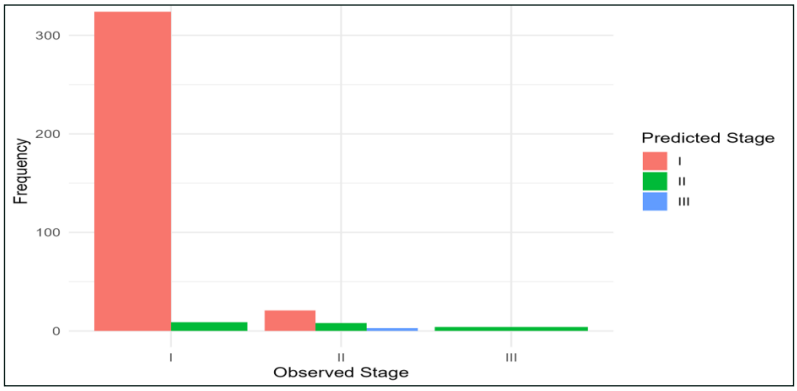

Figure 6: Actual Versus Predicted Thyroid Cancer Stage

The ordinal cumulative logit model provides good classification, as shown by the good agreement between the observed and predicted disease stages in Figure 6. The distributions of the observed and predicted HABs data are very close, indicating that the model is able to capture the main factors governing HABs severity. This is relevant to a weakness of many existing studies predicting thyroid cancer, that they only report on predictive accuracy measures, and neglect to report on the ability of models to reproduce observed stage distributions. The ordinal model shows statistical validity and clinical relevance as it accurately reflects the severity structure. The results confirm that cumulative logit models are a useful way of assessing the severity of the disease and stratifying disease and risk.

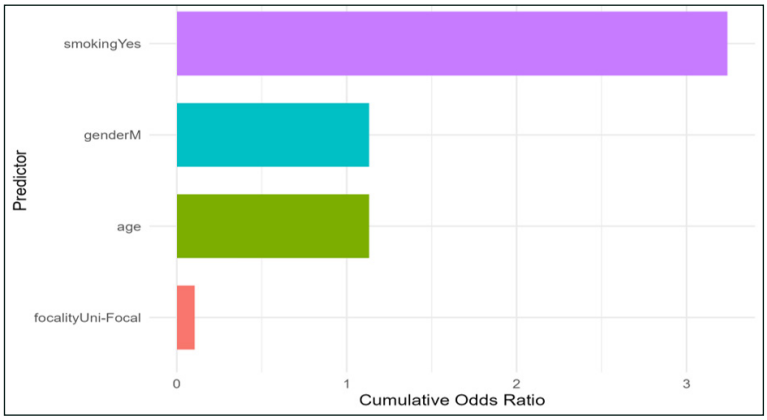

Figure 7: Cumulative Odds Ratios from the Ordinal Logit Model

Figure 7 represents a summary of the size of the predictor effects on the severity of the disease, in a visual format. Smoking had the highest positive cumulative odds ratio, while single focal disease had a high protective factor, and age had a moderate positive factor. The graphical representation shows the relative contribution of predictors to the process of moving to more advanced disease stages. As in previous studies, age and multifocality were the major factors influencing the severity [3,21]. The novelty of the present study is, however, that these effects are quantified in an ordered disease framework, instead of a recurrence framework. The results, therefore, contribute to a more complete picture of the progression of thyroid cancer and can assist in management for an individual patient, communication to patients, and policy decisions to enhance cancer prognosis.

Discussion

This study aimed to examine the factors that influence the severity of thyroid cancer in an ordinal cumulative logit modelling framework, which maintained the natural ordering of the disease stages. Based on this understanding, most previous studies on thyroid cancer have been binary recurrence studies or involved machine-learning-based classification methods, but have not directly measured disease severity; this study directly looked at progression

across staged cancer. The results showed that older age was a significant risk factor for more advanced thyroid cancer stages (13.2% increased odds of being in a higher stage of thyroid cancer for each additional year of age; OR = 1.132, 95% CI: 1.090–1.175, $p < 0.001$). This finding validates the key role age plays in the pathogenesis of thyroid cancer and provides quantitative evidence for a systematic increase in disease severity categories with ageing.

The study also found that one of the best indicators of the severity of thyroid cancer was focality. The cumulative odds of advanced disease were significantly lower for the patients with uni-focal tumours (OR = 0.106, 95% CI: 0.039–0.294, $p < 0.001$) than for the patients with multi-focal tumours. The implication from this discovery is that a significant level of protection against progression to higher stage thyroid cancer can be provided by uni-focal disease and this underlines the clinical significance of tumour focality in disease evaluation and management. Smoking had an increased cumulative odds ratio (OR = 3.241), but was not statistically significant, and gender did not have an independent association with disease severity after controlling for other factors.

The calculated thresholds parameter estimates showed that there was a clear separation between each disease stage and were statistically significant, supporting the use of the ordered outcome of thyroid cancer. The ordinal cumulative logit model was compared to the null model, and good model fit was achieved as evidenced by the improved AIC (265.048 to 157.613) and log-likelihood (−130.524 to −72.807). The results here support significant merits of the ordinal framework over the non-ordered disease models for the estimation of clinically meaningful heterogeneity in progression of thyroid cancer.

Important gradients of disease severity were found between the different disease risk categories and between different age categories with the help of predicted probability analyses. Patients that were determined to be low risk had 94.6% probability of Stage I and high-risk patients had significantly higher probabilities of Stage II and Stage III. Likewise, the younger patients (≤30 years) had a predicted probability of 99.5% for Stage I disease and markedly higher predicted probabilities of advanced disease were seen in those older than 60 years. The results show that there are strong associations between both age and risk stratification with progression along the thyroid cancer severity continuum. The stage distributions of the actual data and the predicted data were well matched, suggesting that the fitted model was predicting the data well. Together, these results show that the ordinal cumulative logit model can give clinically meaningful estimates of disease progression while capturing the information in the ordered categories. The study thus presents a novel statistical contribution to the study of thyroid cancer that focuses, rather than on recurrence prediction, on direct modelling of the severity and progression of the disease.

Conclusion

This research was able to use an ordinal cumulative logit model to analyse factors related to severity of the thyroid cancer without losing the natural order of the stages. The results showed age and tumour focality were the most important factors determining the progression to more advanced stages of thyroid cancer. A significant positive association was seen with increasing age and cumulative odds of severe disease, while a significant negative association was found between uni-focal tumours and the odds of progression compared with multi-focal disease. These findings highlight the need for clinical surveillance and extensive pathological assessment in a patient-specific manner to manage thyroid cancer.

One of the important limitations of the literature that this research aims to overcome is the lack of a comprehensive analysis of the costs of the policy. Most existing works have concentrated on binary recurrence outcomes, survival prediction or machine-learning classification models which have failed to take into account the ordered thyroid cancer progression. The present study preserved clinically meaningful information by directly modelling the severity of the disease as an ordinal outcome, boosted the statistical efficiency, and provided interpretable cumulative odds ratios and stage-specific probabilities that are directly applicable to clinical decision-making.

Ordinal severity modelling provides a more complete model of a cancer's progression than traditional binary modelling of progression, as seen by the superior model fit and meaningful predicted probability patterns across age groups and risk categories. The results are encouraging for incorporating ordinal modelling

frameworks into future research on predicting thyroid cancer prognosis, and indicate the importance of risk stratification by severity for guiding patient care.

The findings indicate that older patients and patients with multifocal disease should be monitored more closely and treated to mitigate progression to advanced disease from a clinical and public health point of view. Larger multicenter cohorts, molecular biomarkers, genomic characteristics and longitudinal follow-up data should be included in future studies to further develop the models of severity prediction and streamline individualised care of thyroid cancer. In conclusion, the use of ordinal severity-based analytical frameworks could help to facilitate earlier interventions, better risk stratification, treatment planning and improved outcomes for patients, ultimately contributing to efforts to meet Sustainable Development Goal 3 – good health and well-being.

Data Availability

This study used secondary longitudinal data on HIV prevalence among females aged 15–24 years obtained from the World Bank World Development Indicators (WDI) database. The dataset is publicly accessible via the World Bank Open Data platform: https://databank.worldbank.org/source/world-development-indicators.

Ethical Approval

No ethical approval was needed since the study is based on publicly available aggregated data and will not involve the use of human subjects.

The authors confirm that they do not have any recognised competing financial interests or personal relations that might have manifested themselves to affect the work presented in this paper.

Consent for Publication

The publication of this manuscript was fully agreed upon by all the authors.

Authors Contributions

F.A.M.: Conceptualisation, Validation, Methodology, Formal Analysis, Writing - Original Draft, Software, Writing - Review & Editing.

E.H.: Visualisation, Writing - Review & Editing, Investigation, Methodology. Data Collection.

M.F.: Visualisation, Writing - Review & Editing, Investigation, Methodology. Data Collection.

V.N.: Validation, Writing - Review & Editing, Visualisation, Investigation

J.K.A.J.: Visualisation, Validation, Writing - Review & Editing, Investigation.

References

1. Chakrabarty N, Mahajan A, Basu S, D'Cruz AK (2024) Comprehensive review of the imaging recommendations for diagnosis, staging, and management of thyroid carcinoma. Journal of clinical medicine 13: 2904.
2. Forma A, Kłodnicka K, Pająk W, Flieger J, Teresińska B, et al. (2025) Thyroid cancer: epidemiology, classification, risk factors, diagnostic and prognostic markers, and current treatment strategies. International Journal of Molecular Sciences 26: 5173.
3. Ringel MD, Sosa JA, Baloch Z, Bischoff L, Bloom G, et al. (2025) 2025 American Thyroid Association management guidelines for adult patients with differentiated thyroid cancer. Thyroid® 35: 841-985.
4. Nguyen M, He G, Lam AKY (2022) Clinicopathological and molecular features of secondary cancer (metastasis) to the thyroid and advances in management. International Journal of Molecular Sciences 23: 3242.
5. Grani G, Sponziello M, Filetti S, Durante C (2024) Thyroid nodules: diagnosis and management. Nature Reviews Endocrinology 20: 715-728.
6. Haddad RI, Bischoff L, Ball D, Bernet V, Blomain E, et al. (2022) Thyroid carcinoma, version 2.2022, NCCN clinical practice guidelines in oncology. Journal of the National Comprehensive Cancer Network 20: 925-951.
7. Li LR., Song JL, Liu HQ, Chen C (2023) Metabolic syndrome and thyroid Cancer: risk, prognosis, and mechanism. Discover Oncology, 14: 23.
8. Chen W, Li J, Peng S, Hong S, Xu H, et al. (2022) Association of total thyroidectomy or thyroid lobectomy with the quality of life in patients with differentiated thyroid cancer with low to intermediate risk of recurrence. JAMA surgery 157: 200-209.
9. Palella M, Giustolisi FM, Modica Fiascaro A, Fichera M, Palmieri A, et al. (2023) Risk and prognosis of thyroid cancer in patients with Graves' disease: an umbrella review. Cancers 15: 2724.
10. Bekker J, Repanos C, Nixon I, Wreesmann V (2026) Subcentimeter Thyroid Nodules: Diagnostic Challenges, Risk Stratification and Management in Contemporary Surgical Oncology Practice. European Journal of Surgical Oncology 111848.
11. Baloch ZW, Asa SL, Barletta JA, Ghossein RA, Juhlin CC, et al. (2022) Overview of the 2022 WHO classification of thyroid neoplasms. Endocrine pathology 33: 27-63.

12. Scerrino G, Marcianò M, Vicari B, Bullaro MA, Di Vuolo R, et al. (2026) Current Evidence, Selective Indications, and the Role of Lymph-Node Assessment in Intraoperative Frozen Section in Thyroid Cancer Surgery: A Literature Review. Journal of Clinical Medicine 15: 1611.
13. Xu Y, Huang P, Wang L, Ke N, Guo F, et al. (2025) Association of Radioactive Iodine Administration with Outcome Among Patients with Low-Risk Differentiated Thyroid Cancer: A Real-World Data Analysis. Clinical Endocrinology 102: 205-213.
14. Tsokos CP, Abeywardana D (2025) Analysis and Statistical Modeling of Thyroid Cancer. In International Encyclopedia of Statistical Science 55-58.
15. Thakur D, Gera T, Bhardwaj V, Mazen R, Lasisi A, et al. (2025) A comparative study on advanced predictive modeling of thyroid cancer recurrence using multi algorithmic machine learning frameworks 16: 3385.
16. Yin J, Guo Z, Yi W, He Y, Luo Y, et al. (2026) Development and validation of a multifactorial risk prediction model for breast cancer patients with co-occurring thyroid cancer: a retrospective matched case-control study. Frontiers in Oncology 16: 1772910.
17. Xue J, Xu Y, Li J, Ma S, Fang K (2025) Ordinal Sparse Neural Networks for Modeling Gene-and Imaging-Environment Interactions. Statistics in Medicine 44: 23-24.
18. Van Velsen EF, Peeters RP, Stegenga MT, van Kemenade FJ, van Ginhoven TM, et al. (2023) Evaluating disease-specific survival prediction of risk stratification and TNM systems in differentiated thyroid cancer. The Journal of Clinical Endocrinology & Metabolism 108: e267-e274.
19. Borzooei S, Briganti G, Golparian M, Lechien JR, Tarokhian A (2024) Machine learning for risk stratification of thyroid cancer patients: a 15-year cohort study. European Archives of Oto-Rhino-Laryngology 281: 2095-2104.
20. Wu J, Hu XY, Ghaznavi S, Kinnear S, Symonds CJ, et al. (2022) The prospective implementation of the 2015 ATA guidelines and modified ATA recurrence risk stratification system for treatment of differentiated thyroid cancer in a Canadian tertiary care referral setting. Thyroid 32: 1509-1518.
21. Chidambaranathan N, Thiagarajan S, Gurukeerthi B, Sathe P, Samel P, et al. (2024) The significance of the presence of extranodal extension in the metastatic node of differentiated thyroid cancer: a proposal for modification in the American Thyroid Association (ATA) risk stratification. European Archives of Oto-Rhino-Laryngology 281: 1923-1931.
22. Avram AM, Zukotynski K, Nadel HR, Giovanella L (2022) Management of differentiated thyroid cancer: the standard of care. Journal of nuclear medicine 63: 189-195.
23. Abelleira E, Jerkovich F (2024) Dynamic risk assessment in patients with differentiated thyroid cancer. Reviews in Endocrine and Metabolic Disorders 25: 79-93.
24. Van Houten P, Netea Maier RT, Smit JW (2023) Differentiated thyroid carcinoma: an update. Best Practice & Research Clinical Endocrinology & Metabolism 37: 101687.
25. Alam S, Hider MA, Al Mukaddim A, Anonna FR., Hossain MS et al. (2024) Machine learning models for predicting thyroid cancer recurrence: A comparative analysis. Journal of Medical and Health Studies 5: 113-129.
26. Wang H, Zhang C, Li Q, Tian T, Huang R et al. (2024) Development and validation of prediction models for papillary thyroid cancer structural recurrence using machine learning approaches. BMC cancer 24: 427.
27. Lixandru Petre IO, Dima A, Musat M, Dascalu M, Gradisteanu Pircalabioru G, et al. (2025) Machine learning for thyroid cancer detection, presence of metastasis, and recurrence predictions—A scoping review. Cancers 17: 1308.
28. Canberk S, Isaza A, Bojarsky M, Simplício M, Barroca H, et al. (2026) The influence of age-independent somatic driver alterations on clinical outcomes in paediatric and young adult thyroid cancer. European Thyroid Journal 15.
29. Ji G, Sun H, Chen S, Sun X, Chang L, et al. (2025) Single-cell RNA sequencing and multi-omics analysis of prognosis-related staging in papillary thyroid cancer. Cancer Immunology, Immunotherapy 74: 267.
30. Al Tashi Q, Saad MB, Muneer A, Qureshi R, Mirjalili S, et al. (2023) Machine learning models for the identification of prognostic and predictive cancer biomarkers: a systematic review. International journal of molecular sciences 24: 7781.