Histopathology

Histopathology 2024 DOI: 10.1111/his.15141

Novel prognostic nomogram for predicting recurrence-free survival in medullary thyroid carcinoma

Yagiz A Aksoy,[1,2,3] Bin Xu,[4] Kartik Viswanathan,[5] Mahsa S Ahadi,[1,2,3] Abir Al Ghuzlan,[6] Bayan Alzumaili,[7] Mohamed-Amine Bani,[6] Justine A Barletta,[8] Nicole Chau,[8] Angela Chou,[1,2,3] Adele Clarkson,[1,2] Roderick J Clifton-Bligh,[3,9] Antonio De Leo,[10] Snjezana Dogan,[4] Ian Ganly,[11] Ronald Ghossein,[4] Matti L Gild,[3,12] Anthony R Glover,[3,9] Julien Hadoux,[13] Livia Lamartina,[13] Daniel J Lubin,[5] Kelly Magliocca,[5] Fedaa Najdawi,[8] Aradhya Nigam,[11] Alex Papachristos,[3,9,12] Andrea Repaci,[14] Bruce G Robinson,[3,9,12] Amy Sheen,[1,2] Qiuying Shi,[5] Stan B Sidhu,[3,9,12] Loretta Sioson,[1,2] Erica Solaroli,[15] Mark S Sywak,[3,9] Giovanni Tallini,[10] Venessa Tsang,[3,12] John Turchini,[1,3,16] Brian R Untch,[11] Anthony J Gill[1,2,3] & Talia L Fuchs[1,3,16]

[1]Cancer Diagnosis and Pathology Group, Kolling Institute of Medical Research, [2]NSW Health Pathology, Department of Anatomical Pathology, Royal North Shore Hospital, St Leonards, [3]Sydney Medical School, University of Sydney, Sydney, NSW, Australia, [4]Department of Pathology and Laboratory Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, [5]Department of Pathology, Emory University Hospital Midtown, Atlanta, GA, USA, [6]Medical Pathology and Biology Department, Gustave Roussy Campus Cancer, Villejuif, France, [7]Department of Pathology, Massachusetts General Hospital, [8]Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA, [9]University of Sydney Endocrine Surgery Unit, Royal North Shore Hospital, St Leonards, NSW, Australia, [10]Department of Medical and Surgical Sciences (DIMEC), University of Bologna Medical Center, IRCCS Azienda Ospedaliero-Universitaria di Bologna, Bologna, Italy, [11]Department of Surgery, Memorial Sloan Kettering Cancer Center, New York, NY, USA, [12]Department of Endocrinology, Royal North Shore Hospital, St Leonards, NSW, Australia, [13]Endocrine Oncology, Gustave Roussy Campus Cancer, Villejuif, France, [14]Division of Endocrinology and Diabetes Prevention and Care, IRCCS Azienda Ospedaliero-Universitaria di Bologna, [15]Endocrinology Unit, Azienda USL di Bologna, Bologna, Italy and [16]Douglass Hanly Moir Pathology, Macquarie Park, New South Wales, Australia

Date of submission 6 November 2023
Accepted for publication 2 January 2024

Aksoy Y A, Xu B, Viswanathan K, Ahadi M S, Al Ghuzlan A, Alzumaili B, Bani M-A, Barletta J A, Chau N, Chou A, Clarkson A, Clifton-Bligh R J, De Leo A, Dogan S, Ganly I, Ghossein R, Gild M L, Glover A R, Hadoux J, Lamartina L, Lubin D J, Magliocca K, Najdawi F, Nigam A, Papachristos A, Repaci A, Robinson B G, Sheen A, Shi Q, Sidhu S B, Sioson L, Solaroli E, Sywak M S, Tallini G, Tsang V, Turchini J, Untch B R, Gill A J & Fuchs T L

(2024) *Histopathology*. https://doi.org/10.1111/his.15141

Novel prognostic nomogram for predicting recurrence-free survival in medullary thyroid carcinoma

Aims: Recently, there have been attempts to improve prognostication and therefore better guide treatment for patients with medullary thyroid carcinoma (MTC). In 2022, the International MTC Grading System (IMTCGS) was developed and validated using a multi-institutional cohort of 327 patients. The aim of the

Address for correspondence: AJ Gill, Department of Anatomical Pathology, Royal North Shore Hospital, Pacific Highway, St Leonards, NSW 2065, Australia. e-mail: affgill@med.usyd.edu.au
Anthony J. Gill and Talia L. Fuchs contributed equally to this work.
[Correction added on 02 february 2024, after first online publication: In the author byline, Aradhya Nigam's first name was misspelt and has been corrected in this version.]

current study was to build upon the findings of the IMTCGS to develop and validate a prognostic nomogram to predict recurrence-free survival (RFS) in MTC.
Methods and Results: Data from 300 patients with MTC from five centres across the USA, Europe, and Australia were used to develop a prognostic nomogram that included the following variables: age, sex, AJCC stage, tumour size, mitotic count, necrosis, Ki67 index, lymphovascular invasion, microscopic extrathyroidal extension, and margin status. A process of 10-fold cross-validation was used to optimize the model's performance. To assess discrimination and calibration, the area-under-the-curve (AUC) of a receiver operating characteristic (ROC) curve, concordance-index (*C*-index), and dissimilarity index (*D*-index) were calculated. Finally, the model was externally validated using a separate cohort of 87 MTC patients. The model demonstrated very strong performance, with an AUC of 0.94, a *C*-index of 0.876, and a *D*-index of 19.06. When applied to the external validation cohort, the model had an AUC of 0.9.
Conclusions: Using well-established clinicopathological prognostic variables, we developed and externally validated a robust multivariate prediction model for RFS in patients with resected MTC. The model demonstrates excellent predictive capability and may help guide decisions on patient management. The nomogram is freely available online at https://nomograms.shinyapps.io/MTC_ML_DFS/.

Keywords: medullary thyroid carcinoma, prognostic nomogram

Introduction

Medullary thyroid carcinoma (MTC) is an uncommon neuroendocrine malignancy arising from the calcitonin-producing C cells of the thyroid gland. Despite accounting for only 2% of thyroid malignancies, MTC is responsible for 8% of thyroid cancer-related deaths owing to its aggressive biological behaviour and propensity for early metastasis.[1] Approximately 25% of cases occur in the setting of multiple endocrine neoplasia type 2 (MEN2) syndrome and are often detected at an early stage through biochemical or genetic screening programs.[2] However, the majority of cases are sporadic and often pursue an aggressive clinical course. Among patients presenting with a palpable thyroid mass, the rate of cervical lymph node involvement at diagnosis is as high as 75%, and the rate of distant metastasis is up to 10%.[3-5]

The mainstay of treatment for MTC is total thyroidectomy and central compartment lymph node dissection; however, the benefit of lateral node dissection in patients with thyroid-confined disease remains uncertain.[6] Until recently, clinicians relied primarily on the American Joint Committee on Cancer (AJCC) TNM staging system to predict patient outcomes and help guide treatment decisions. However, the limitations of this staging system have been recognized for many years and, over the past few years, there have been numerous attempts to create and validate various grading schemes to further improve prognostication.[7,8] In 2022, the International Medullary Thyroid Carcinoma Grading System (IMTCGS) was developed and validated using a cohort of 327 patients with MTC from five centres across the USA, Europe, and Australia.[9] This two-tiered grading scheme, which incorporates mitotic count, Ki67 proliferative index, and tumour necrosis, has been shown to be a robust independent prognostic tool for patients with MTC in both the original multi-institutional cohort, and in multiple subsequent independent validation cohorts.[10-12]

As advances in artificial intelligence and machine-learning techniques continue to revolutionize medical practice, predictive modelling has become a popular technique for developing prognostic nomograms for various types of cancer. Nomograms have significant advantages over traditional prognostic techniques, as they can synthesize large numbers of complex variables in a systematic, unbiased manner. Previous studies have developed prognostic nomograms for patients with MTC based on relatively small, single-institution cohorts.[13,14] Moreover, larger studies using Surveillance, Epidemiology, and End Results (SEER) data[15-17] have not included histological variables such as proliferative activity (mitotic count and Ki67 index) and tumour necrosis that are now known to be critical prognostic factors in MTC.[7,8] We therefore sought to build upon the findings from the IMTCGS study to develop a novel prognostic nomogram that incorporates these important histopathological variables with the aim of further improving risk stratification for patients with MTC.

Methods

STUDY COHORT

The initial training and discovery cohort was based on 300 patients with MTC from five academic

centres (Royal North Shore Hospital [RNSH], Sydney, Australia: $n = 58$; Institut Gustave Roussy [IGR], Villejuif, France: $n = 70$; Memorial Sloan Kettering Cancer Center [MSKCC], New York, NY: $n = 68$; University of Bologna Medical Center [BO], Bologna, Italy: $n = 64$; and Brigham and Women's Hospital [BWH], Boston, MA: $n = 40$) which had previously contributed to the IMTCGS study.[9] In addition to *k*-fold validation (described below), as an independent validation cohort we included data from an additional 87 MTC patients from Emory University Midtown Hospital, Atlanta, GA.[18] Although patients from these cohorts were included in prior studies,[7,8,9,18,19] prognostic factors were independently assessed, and survival data updated for this study. All cases were previously reviewed at individual participating sites by at least one specialist endocrine pathologist to confirm the diagnosis of MTC as part of the preceding studies. Data for all patients were obtained from pathology reports and electronic medical records at each institution and included: age, sex, tumour size, lymphovascular invasion, microscopic extrathyroidal extension (ETE), and margin status. Lymphovascular space invasion was defined as the presence of either lymphatic or vascular space invasion. The study was approved by the Institutional Review Board of each participating site.

EVALUATION OF MORPHOLOGICAL VARIABLES

Grading was based on the primary tumour (that is, the thyroid and any accompanying lymph nodes resected at the time of primary surgery). The mitotic count and Ki67 proliferative index were evaluated using the same methods described previously.[9] Briefly, both measurements were obtained in the area showing the highest proliferative activity (so called 'hot spots'). The mitotic count was assessed per 2 mm^2. For the Ki67 proliferative index, 500–2000 tumour cells were counted manually per tumour. Necrosis was defined as tumour necrosis associated with degenerating cytoplasm and punctate karyorrhectic nuclear debris regardless of its extent. Infarct-like necrosis, which is usually associated with a fibroblastic stromal reaction, haemorrhage, or identifiable fine-needle aspiration tract, was not considered necrosis for the purpose of this study. The histological features of MTC are shown in Figure 1.

SURVIVAL OUTCOMES AND DATA PREPROCESSING

The primary outcome of this study was recurrence-free survival (RFS), defined as the interval length of time after treatment during which a patient survives without any structural locoregional recurrence or distant metastasis. For the purpose of this study,

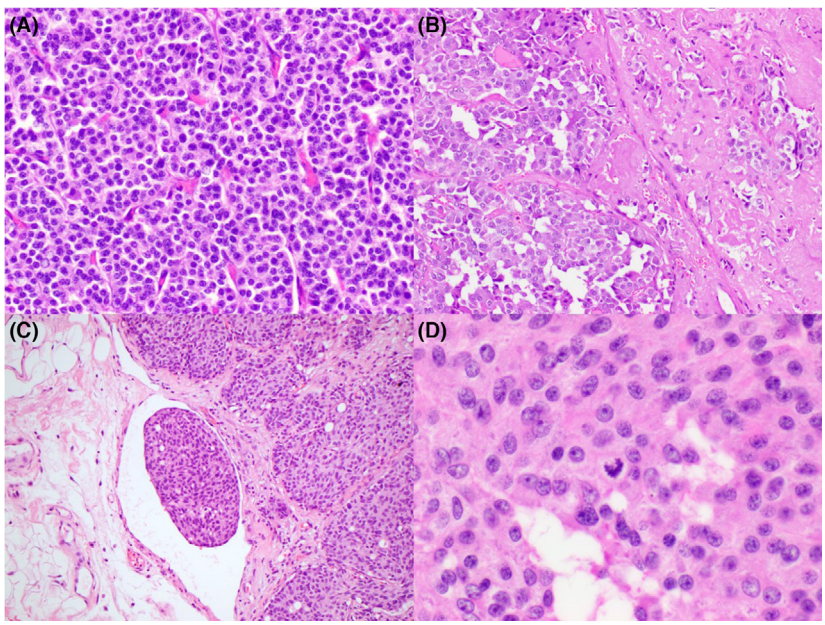

Figure 1. Histological features of medullary thyroid carcinoma (MTC). **A**: Tumour cells have round to oval nuclei with speckled chromatin. **B**: Cytoplasm can be abundant and amyloid deposition is often present. **C**: Lymphovascular invasion and **D**: mitotic activity are important histological prognostic factors.

patients with elevated calcitonin but without structural recurrence (so-called biochemical only recurrence) were considered recurrence free. All statistical analyses were performed using R software.[20] A comprehensive list of R packages used in this study is provided in Table S1. No missing data were encountered in the dataset. Categorical variables were converted into a binary format using the one-hot encoding method, while continuous variables were normalized to ensure a uniform data distribution and enhance the performance of the machine-learning model.

MODEL DEVELOPMENT

The predictive model was developed based on data from the 300 patients provided by the five international centres that had contributed to the IMTCGS study, using Cox Proportional Hazards (CPH) modelling for regression analyses using the 'rms' package in R. Initially, a univariate CPH model was employed to investigate the association of each variable with RFS. A pairwise correlation matrix was used to assess collinearity between variables. Variables that displayed statistical significance in the univariate analysis were incorporated into a multivariate CPH model. This approach was taken to identify and consider the variables that have the greatest influence on outcome. The results of the univariate and multivariate survival analyses are shown in Table 1.

MODEL VALIDATION AND EVALUATION

A *k*-fold cross-validation method was applied to the training cohort to fit the multivariate Cox model and to assess model stability. The entire cohort was divided into a training cohort and a validation cohort based on a 9:1 split using simple random sampling. A 10-fold cross-validation, dictated by the variable n_cv, was performed on the training cohort for both fitting the multivariate Cox model and assessing the model stability. The 10-fold cross-validation involves partitioning the original sample into 10 equal size subsamples. Of the 10 subsamples, a single subsample (10% of the cohort) was retained as validation data for testing the model, and the remaining nine subsamples (90% of the cohort) were used for training. This process was repeated 10 times, with each of the 10 subsamples used exactly once as validation data. The 10 results were then averaged to produce a single estimation, enhancing the reliability of the model performance evaluation.

Hazard ratios for each variable, along with their 95% confidence intervals (CI), were computed across the 10 folds to assess the stability and reliability of the predictors (Figures S5 and S6). Prediction errors

Table 1. Univariate and multivariate survival analyses of prognostic factors for recurrence-free survival for 300 medullary thyroid carcinoma patients

	Univariate*			Multivariate*		
Variable	HR	95% CI	*P*-value	HR	95% CI	*P*-value
Age	1.007	0.994–1.019	0.312	0.999	0.985–1.014	0.936
Sex	0.321	0.212–0.487	**<0.01**	1.042	0.661–1.644	0.860
AJCC stage	5.055	3.34–7.651	**<0.01**	3.309	2.107–5.198	**<0.01**
Tumour size (cm)	1.392	1.283–1.511	**<0.01**	0.959	0.857–1.074	0.470
Mitotic count	1.245	1.179–1.315	**<0.01**	0.975	0.887–1.071	0.593
Necrosis	7.083	4.649–10.791	**<0.01**	1.785	1.044–3.052	**0.034**
Ki67	1.117	1.09–1.144	**<0.01**	1.038	0.993–1.086	0.101
Vascular invasion	10.944	6.391–18.74	**<0.01**	1.952	1.055–3.611	**0.033**
Extrathyroidal extension	10.497	6.755–16.313	**<0.01**	1.607	0.928–2.784	0.091
Margin	6.815	4.549–10.208	**<0.01**	1.937	1.194–3.145	**0.007**

CI, confidence interval; HR, hazard ratio.

Note: Bold values indicate statistically significant.

*Cox proportional hazards model.

were estimated for each model to provide insight into the model's accuracy.

Additionally, receiver operating characteristic (ROC) curves and confusion matrices were used to measure the model's performance. The model's discrimination capacity was evaluated using the area under the curve (AUC) of the ROC, *C*-index, and *D*-index. Furthermore, the model's calibration was assessed with calibration plots of expected versus observed survival.

EXTERNAL VALIDATION

The predictive model was externally validated using a separate cohort of patients with MTC (n = 87) from Emory University Midtown Hospital to test the model's generalizability. The clinical and pathological features of this cohort have been previously reported,[18] but were not used in any way for model training. ROC curves for the two cohorts are illustrated in Figure S8.

MODEL PERFORMANCE

Model performance was assessed using Decision Curve Analysis (DCA). The DCAs were plotted to visually inspect the net benefits of each model over a range of threshold probabilities. Two Cox models were generated, one using all variables, and the other solely based on the AJCC stage. The survival probabilities at distinct time-points (12, 24, 36, 120, and 240 months) were calculated for both models. Subsequently, DCA was conducted using the 'dcurves' package in R to compare the net benefits of each model at these time-points. The net benefit curves were plotted for each time-point and model, providing a visual comparative analysis (Figure S7). Multivariate analyses were also conducted, which enabled the evaluation of multiple variables simultaneously in order to understand their independent effects on the outcome.

CONSTRUCTION OF THE WEB-BASED NOMOGRAM

Based on the final model, a nomogram was constructed to predict RFS at specific time intervals post-diagnosis. The nomogram serves as a visual tool that quantifies risk and generates individual probabilities of survival by allocating points to each feature of the model. To facilitate its application, a web-based dynamic nomogram was developed using the 'Shiny' package in R.[21] This interactive tool, hosted online, is intended to help clinicians estimate the RFS probabilities in individual patients with MTC based on their unique clinical and pathological characteristics. The prognostic calculator is freely available online at: https://nomograms.shinyapps.io/MTC_ML_DFS/.

Results

CLINICAL AND PATHOLOGICAL CHARACTERISTICS

Our nomogram was developed using a cohort of 300 patients with MTC distributed across five sites: BO, BWH, IGR, MSKCC, and RNSH. The variables taken into account for our analysis included patient demographic factors, pathological findings, and survival data. The median age at surgery was 58 years (mean 55 years; range 3–88 years). 130 (43.3%) patients were male. Mean tumour size was 2.42 cm (median 1.9 cm; range 0.13–11 cm). According to the AJCC stage, 99 (33%) patients were classified as stage 1, 37 (12.3%) as stage 2, 42 (14%) as stage 3, and 122 (40.7%) as stage 4. Tumour necrosis was present in 43 (14.3%) cases. Lymphovascular invasion was observed in 125 (41.7%) cases, microscopic ETE in 89 (29.7%) cases, and margin involvement in 54 (18%) cases. Mitotic counts ranged from 0 to 15 per 2 mm^2, with a mean of 1.5/2 mm^2 and a median of 1/2 mm^2. Ki67 proliferative indices ranged from 0% to 30%, with a mean of 3.7% and a median of 2%. There was no strong collinearity or multicollinearity between any of the variables (Tables S2–S4).

Figure 2 showcases density plots of patient characteristics across the five participating institutions. These plots were instrumental in preliminary data exploration, allowing us to visually assess the distribution of patient characteristics and ensure the absence of significant cohort differences. While the representation of binary categorical variables results in two peaks, the relative height and spread of these peaks provided insights into the distribution of these categories across institutions.

A principal component analysis (PCA) was performed to help visualize the high-dimensional data, to reveal how patients clustered based on their characteristics, and to identify any potential structure in the data associated with the different collection sites (Figure S1). We also implemented t-Distributed Stochastic Neighbour Embedding (t-SNE), a dimensionality reduction technique that is particularly suitable for high-dimensional datasets (Figure S2). The purpose of employing t-SNE was to reveal any nonlinear patterns within the high-dimensional data that PCA might miss, and to further help in identifying clusters or groups of patients with similar characteristics.

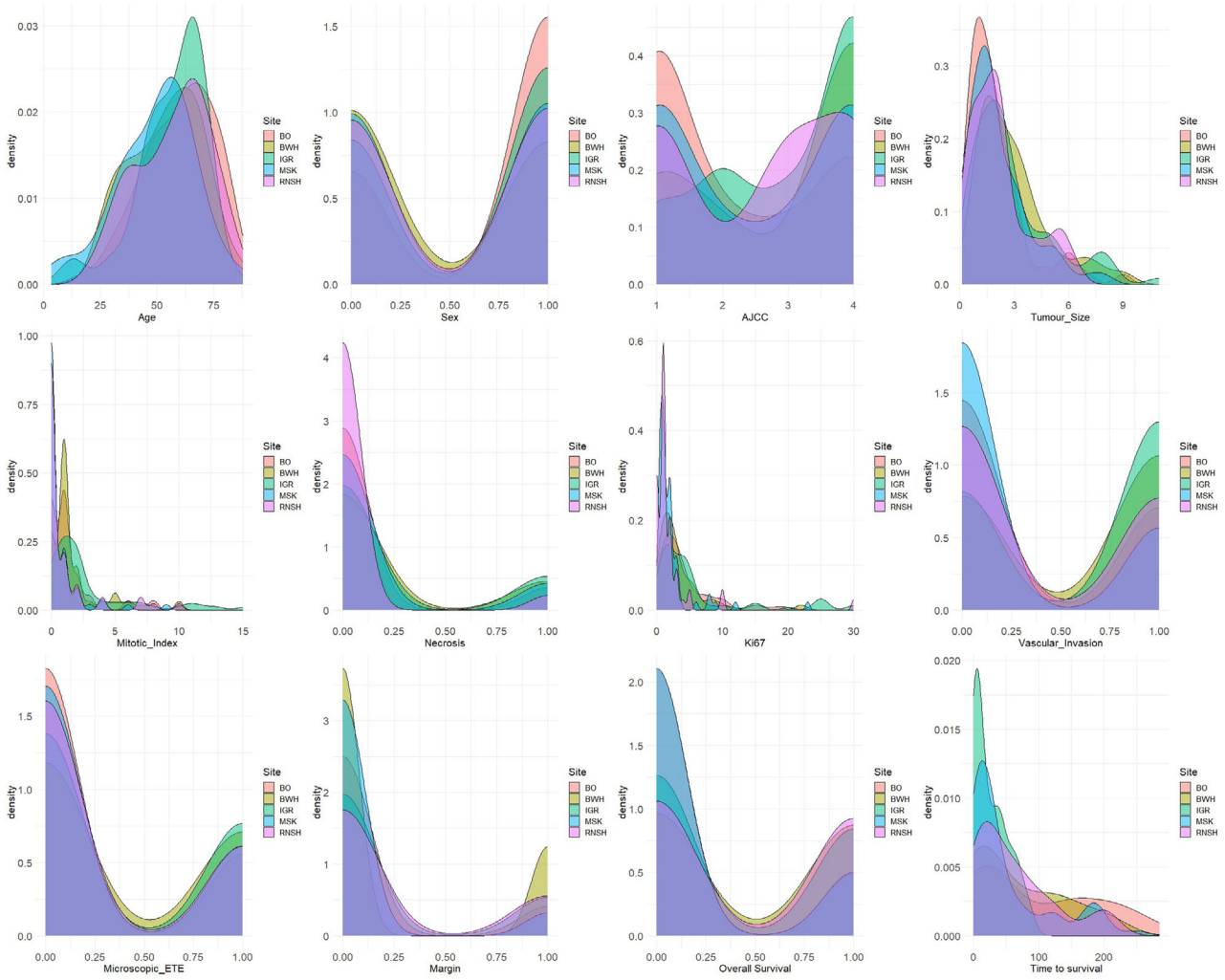

Figure 2. Density plots of patient characteristics across five sites. The figure comprises 12 panel plots, each showing the distribution of a different patient characteristic variable (age, sex, AJCC, tumour size, mitotic index, necrosis, Ki67, lymphovascular invasion, microscopic extrathyroidal extension [ETE], margin, recurrence-free survival, time-to-survival event) across five sites (BO, BWH, IGR, MSK, and RNSH).

IDENTIFICATION OF PROGNOSTIC FACTORS

A comprehensive survival analysis was carried out utilizing both multivariate analysis and Kaplan–Meier estimates. The results are illustrated in Figure 3.

MODEL CONSTRUCTION AND PERFORMANCE

The variables included in the model are: age, sex, AJCC stage, tumour size, mitotic index, presence of necrosis, Ki67 index, lymphovascular invasion, microscopic ETE, and surgical margin status. The results of the multivariate CPH regression model are demonstrated in the forest plot (Figure 4). The final nomogram is shown in Figure 5.

MODEL VALIDATION

We evaluated the performance of the proposed CPH model through a series of tests, which included calculating the AUC of the ROC curve, the *C*-index, and the *D*-index.

Our model demonstrated strong performance, with an AUC of 0.94, indicating a high ability to accurately distinguish between patients with different outcomes. Further quantification of the discriminative ability of the model showed a *C*-index of 0.876,

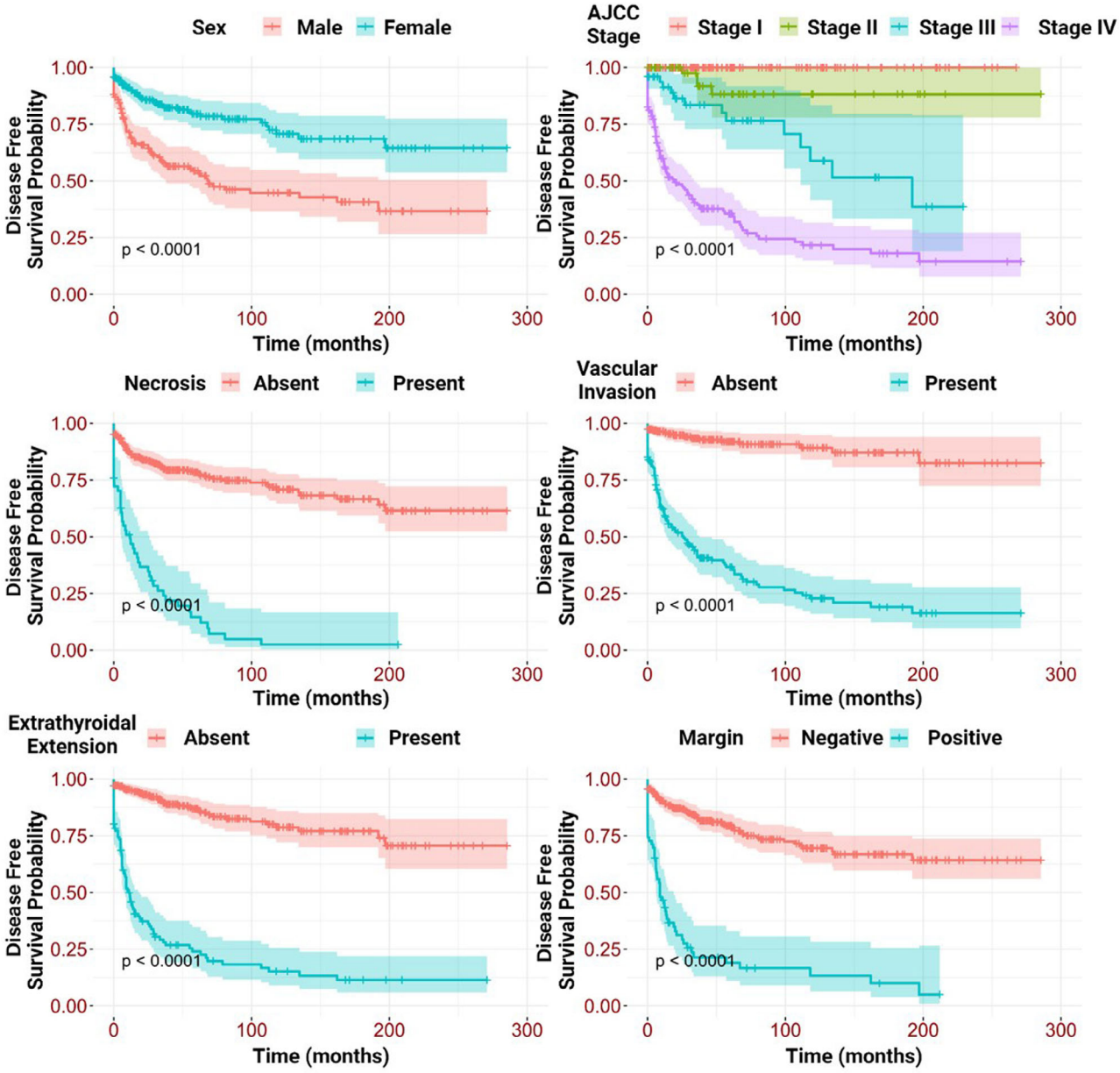

Figure 3. Kaplan–Meier survival curves stratified by sex, AJCC stage, necrosis, lymphovascular invasion, microscopic ETE, and margin status.

which is considered excellent in survival models. The *D*-index, which quantifies the model's discriminatory power and potential clinical usefulness,[21-25] was found to be 19.06, suggesting strong discrimination between patients with varying risk (Figure S3).

We plotted a graph of the false-positive rate (FPR) and true-positive rate (TPR) against different classification thresholds, pinpointing an ideal threshold for optimizing accuracy (Figure S4). After applying this threshold for binary classification, we used a confusion matrix to evaluate our model's performance, which demonstrated an overall accuracy rate of 0.86, with a kappa statistic of 0.69, indicating substantial agreement beyond chance. The 95% CI for accuracy was (0.816–0.897). The accuracy null hypothesis (H0: Accuracy = 0.667) was rejected with a *P*-value of less than 0.001, which further attests to the validity of our model (Table S5).

The robustness of our model was demonstrated by its successful validation using the independent external cohort from Emory—described in detail elsewhere.[18] The AUC for our model when applied to

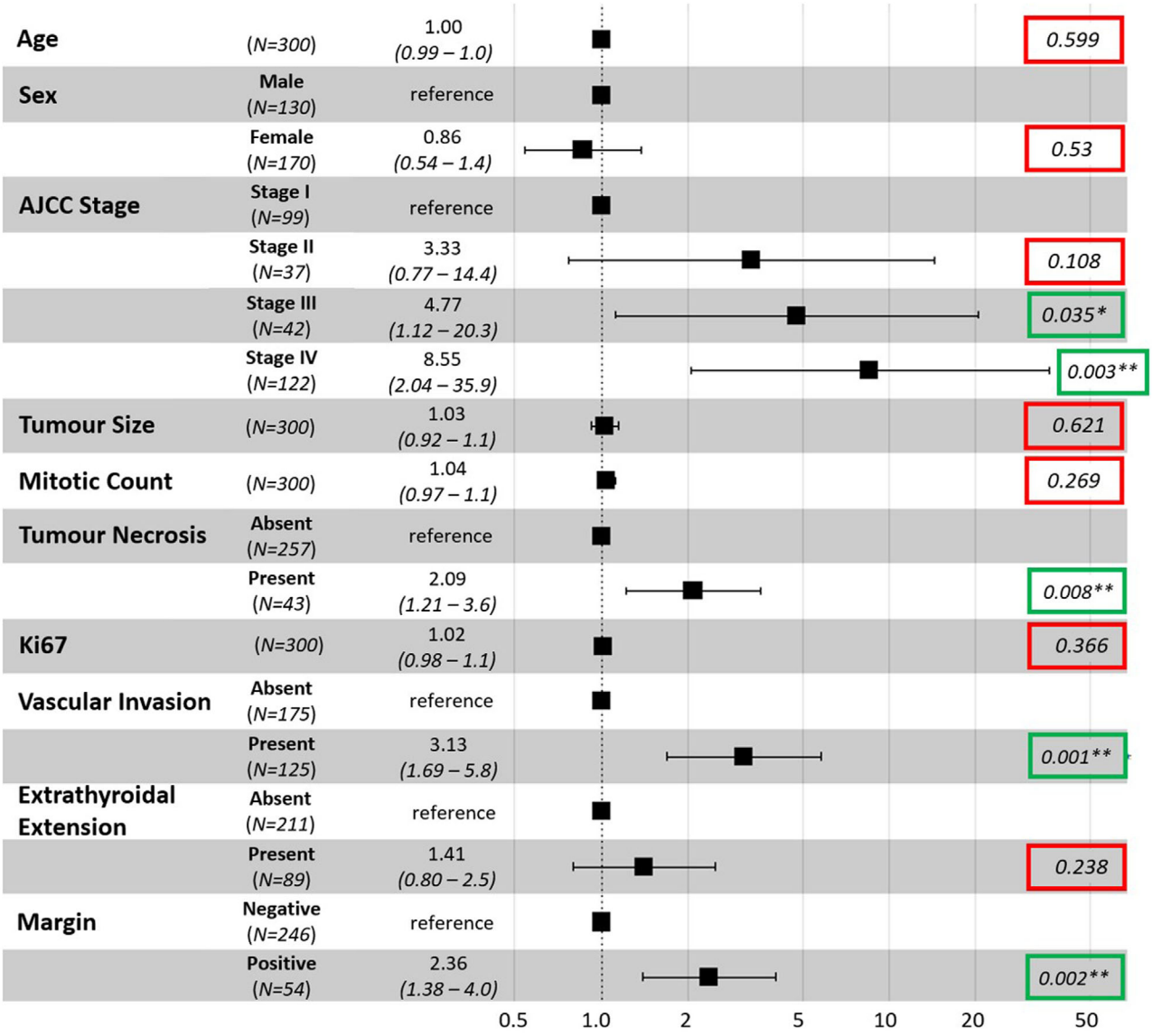

Figure 4. Forest plot of Cox proportional hazards model for recurrence-free survival. Green boxes indicate *P*-values less than 0.05; red boxes indicate *P*-values greater than 0.05.

the Emory dataset was 0.9, confirming its excellent predictive power. This consistency in AUC between our model applied to the MTC cohort and the Emory cohort, as supported by the nonsignificant *P*-value (0.3192), indicates the successful transferability of our model to other datasets, and hence its high generalizability.

The ROC curves for our model, utilizing both the initial 'MTC' cohort and the Emory validation cohort are illustrated in Figure 6. The ROC curve for the AJCC staging system is also included for comparison. Our model significantly outperformed the AJCC staging system, as demonstrated by bootstrap tests yielding *P*-values less than 2.2×10^{-16}. These very small *P*-values indicate a highly statistically significant difference in AUCs, confirming that our model adds very significant predictive power to AJCC staging alone. Further details on the validation of this model are available in Tables S6–S8.

FREELY AVAILABLE WEB-BASED PROGNOSTIC CALCULATOR

To enhance the practical application of our prognostic model in clinical settings, we developed an intuitive, web-based calculator. This dynamic nomogram

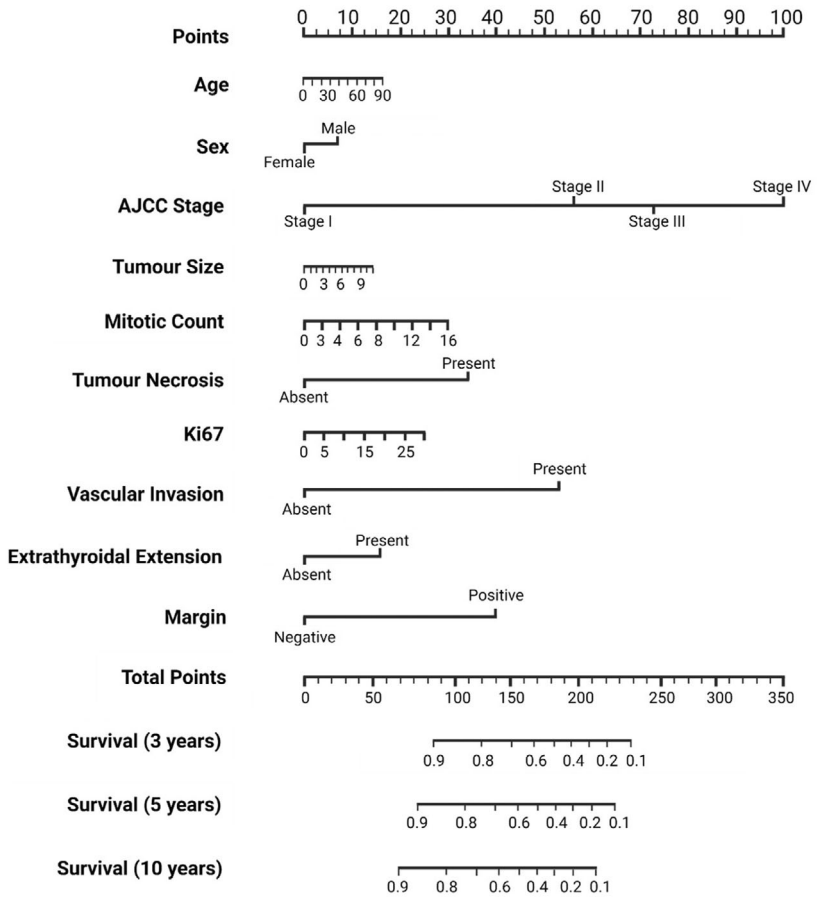

Figure 5. Nomogram for predicting 3-, 5-, and 10-year recurrence-free survival in MTC patients.

is freely accessible at https://nomograms.shinyapps.io/MTC_ML_DFS/. The online tool is designed for ease of use, allowing clinicians and researchers to input individual patient data directly into the interface. Once the necessary data are entered, the calculator uses our nomogram model to generate RFS probabilities at various time-points. These results are presented in an easy-to-interpret format, including graphical survival curves and numerical estimates of survival probability.

DEVELOPMENT AND VALIDATION OF A SECOND PROGNOSTIC NOMOGRAM FOR OVERALL SURVIVAL

Using the same approach, we then developed and validated a second prognostic nomogram for overall survival (OS), available at https://nomograms.shinyapps.io/MTC_ML_OS/. The AUC for the OS nomogram when applied to the external cohort from Emory was 0.878. The full details of this second nomogram for overall survival are available in the Supplementary Materials.

Discussion

There is now broad agreement that histological features including mitotic count, Ki67 proliferative index, and tumour necrosis are robust prognostic factors in MTC. Recently, an international consortium developed a two-tiered grading system that incorporates these three histological variables and is now endorsed by the World Health Organization (WHO) / International Agency for Research against Cancer (IARC).[9,26] In recent years, machine-learning techniques have been used to further enhance cancer prognostication in response to the discovery of increasingly complex prognostic factors. Although a small number of prognostic nomograms for MTC do exist, they incorporate relatively few variables beyond basic demographic data and TNM stage, and fail to include more recently established prognostic factors

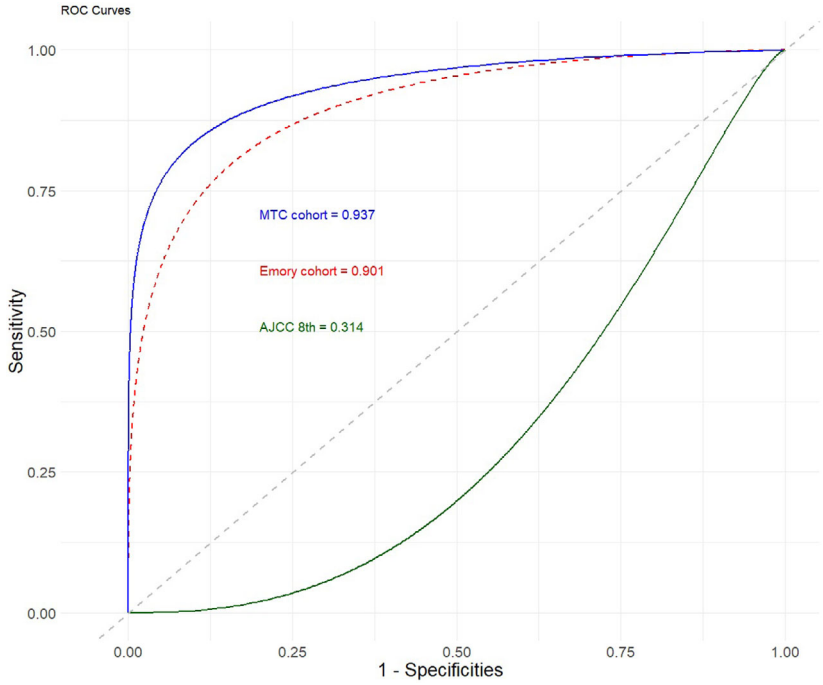

Figure 6. Receiver operating characteristic (ROC) curves and area-under-the-curve (AUC) values for the nomogram in both the original MTC cohort and the Emory validation cohort, compared with the ROC curve for AJCC stage alone.

such as mitotic count, Ki67 proliferative index, and tumour necrosis. In this study, we aimed to develop a robust and accurate multivariate prediction model for RFS in patients with resected MTC based exclusively on clinical and pathological data that is already collected in routine practice. Our discovery (n = 300) and validation (n = 87) cohorts, distributed across six different locations, constitutes one of the largest MTC studies to date, providing a diverse and representative sample for analysis. Furthermore, our model's excellent performance has also been confirmed in an independent validation cohort.

The strength of our model lies in its comprehensive incorporation of clinical and pathological parameters, including gender distribution, pathological findings, and survival data. By employing best-practice machine-learning techniques, such as PCA and t-SNE, we effectively visualized our high-dimensional data and identified key patterns and structures within our dataset. This approach allowed us to cluster patients based on their unique characteristics, providing valuable insights into the heterogeneity of MTC in different populations.

Through rigorous survival analyses, we identified several major prognostic factors that significantly impacted RFS in our cohort. The AJCC stage, presence of necrosis, lymphovascular invasion, microscopic ETE, and surgical margin status emerged as crucial variables influencing patient outcomes. We also incorporated several additional variables shown to be prognostically significant in other studies, despite not emerging as significant in this cohort. These variables were retained to enhance the model's generalizability and predictive accuracy across diverse clinical scenarios and patient populations. In complex diseases like MTC, a multifactorial approach that considers a broad spectrum of variables, even if they are not universally significant across all cohorts, is essential for capturing the nuanced interactions and influences that drive patient outcomes.

Our multivariate prediction model surpassed the predictive capacity of the AJCC staging system alone, demonstrating an AUC of 0.94 and a *C*-index of 0.876. These strong performance metrics, coupled with a *D*-index of 19.06, attest to the model's excellent discriminatory ability and potential for clinical applicability. Furthermore, external validation using the Emory University Midtown Hospital dataset reinforced the consistent and transferable performance of our model across different patient cohorts.

Importantly, our model's superiority over a single-variable AJCC model underscores the importance of considering multiple clinical and pathological variables in addition to stage for accurate survival

prediction. By adopting a multifactorial approach, we gain a more nuanced understanding of patient prognosis, enabling personalized treatment strategies and improved patient management.

To enhance the practical utility of our findings in clinical settings, we developed a user-friendly web-based calculator based on our predictive model. Web-based dynamic nomograms offer several advantages over static nomograms, such as ease of use in clinics, rapid translation to clinical practice, and potentially more accurate clinical information. The interactive nature of our web-based nomogram allows clinicians to input patient data directly into the interface to quickly obtain personalized RFS probabilities. This accessibility and adaptability facilitate informed decision-making and improve patient outcomes.

Despite the strengths of our study, several limitations should be acknowledged. The retrospective design of the study and the variation in follow-up duration across different sites may have influenced the results. Moreover, data on postsurgical treatment was not available for all patients. While postsurgical treatments can indeed influence survival outcomes, our model was designed to predict RFS based on clinical and pathological data available at the time of diagnosis and surgery. This approach ensures that the nomogram can be applied immediately postsurgery to guide subsequent management decisions. In future iterations of our model, we aim to incorporate postsurgical treatment data to provide an even more comprehensive prediction tool. In the meantime, we recommend interpreting the nomogram's predictions in conjunction with the individual patient's postsurgical treatment plan.

Vascular space invasion appears to be a more adverse prognostic factor than lymphatic invasion. For example, Torricelli *et al.*[12] recently demonstrated that angioinvasion was significantly correlated with metastasis and death in a cohort of 66 patients with MTC. However, some authorities find it difficult to confidently separate capillary from lymphatic space invasion in thyroid malignancies—summarized previously.[27] Furthermore, the datasets employed for generating the nomograms in this study collapsed lymphatic and vascular space invasion into one variable —'lymphovascular invasion'. Therefore, in this study we examined either lymphatic and/or blood vessel invasion together as one variable. However, we note that there may be scope to improve the nomogram in the future by evaluating these parameters separately.

It is also important to note that data on serum calcitonin levels (and calcitonin doubling time) were not available for a significant number of patients in the cohort, and therefore not included in the nomogram. We acknowledge that patients with resected MTC may present with elevated serum calcitonin levels without structural disease recurrence—so-called 'biochemical recurrence'. However, given the lack of data on calcitonin levels, we did not use biochemical recurrence as an endpoint in this study and required structural recurrence. While this may be a potential area for future research, we have previously been unable to demonstrate prognostic significance for calcitonin doubling time in smaller cohorts,[8] possibly owing to differing protocols and time periods for measuring serum calcitonin.

Additionally, the lack of comprehensive genetic data across all locations limited the inclusion of genetic and molecular aspects in our model. Future research should aim to validate and refine our model using larger and more diverse datasets to enhance its performance and generalizability. Furthermore, considering potential interactions among the variables in our model could provide deeper insights into their combined effects on patient outcomes. Incorporating comprehensive genetic and molecular data in future studies may improve the precision of our predictive model, as advances in genomic medicine continue to provide valuable prognostic information.

In conclusion, our study presents a significant advancement in the prediction of RFS in patients with MTC. The incorporation of diverse clinical and pathological variables, along with the development of a user-friendly web-based nomogram, enhances the clinical applicability of our findings. By offering personalized survival predictions, our model can aid healthcare professionals in making informed treatment decisions and ultimately improve patient care. The promising results pave the way for further research and validation, ultimately advancing personalized medicine in the management of MTC. The nomogram for RFS is freely available online at: https://nomograms.shinyapps.io/MTC_ML_DFS/.

Acknowledgement

Open access publishing facilitated by The University of Sydney, as part of the Wiley - The University of Sydney agreement via the Council of Australian University Librarians.

Author contributions

YAA, AJG, TLF performed the study concept, design, interpretation of data, statistical analysis, writing of

first draft, and review and revision of the article. All other coauthors provided data acquisition, material support, and review and revision of the article. All authors read and approved the final article.

Funding statement

ARG was funded by the Cancer Institute of NSW early career fellowship NSW Early Career Fellowship (2019/ECF1081). YAA was supported by the RCPA scholarship in pathology for medical schools.

Conflict of interest statement

KV acknowledges receipt of advisory fees from Veracyte Inc. unrelated to the study. BR is a speaker and on the advisory boards for Eisai, Exelixis, and Lilly. FN reports employment at PathAI. LL is on the tumour board for Bayer, Eisai, and Ipsen; receives honoraria from Eisai, Lilly, and Roche; and receives support for attending meetings and/or travel from Ipsen and AAA Novartis.

Data availability statement

The datasets used and/or analysed during the current study are not publicly available but are available from the corresponding author on reasonable request.

Ethics approval

This study was approved by the Northern Sydney Local Health District Human Research Ethics Committee (ref: LNR 1312-417M) and the local Ethics Committees of all contributing institutions.

References

1. Barletta JA, Nose V, Sadow PM. Genomics and epigenomics of medullary thyroid carcinoma: from sporadic disease to familial manifestations. *Endocr. Pathol.* 2021; **32**; 35–43.
2. Kotwal A, Erickson D, Geske JR, Hay ID, Castro MR. Predicting outcomes in sporadic and hereditary medullary thyroid carcinoma over two decades. *Thyroid* 2021; **31**; 616–626.
3. Rowland KJ, Moley JF. Hereditary thyroid cancer syndromes and genetic testing. *J. Surg. Oncol.* 2015; **111**; 51–60.
4. Moley JF. Medullary thyroid carcinoma: management of lymph node metastases. *J. Natl. Compr. Canc. Netw.* 2010; **8**; 549–556.
5. Larouche V, Akirov A, Thomas CM, Krzyzanowska MK, Ezzat S. A primer on the genetics of medullary thyroid cancer. *Curr. Oncol.* 2019; **26**; 389–394.
6. Papachristos AJ, Nicholls LE, Mechera R *et al.* Management of medullary thyroid cancer: patterns of recurrence and outcomes of reoperative surgery. *Oncologist* 2023; **28**; 1064–1071.
7. Alzumaili B, Xu B, Spanheimer PM *et al.* Grading of medullary thyroid carcinoma on the basis of tumor necrosis and high mitotic rate is an independent predictor of poor outcome. *Mod. Pathol.* 2020; **33**; 1690–1701.
8. Fuchs TL, Nassour AJ, Glover A *et al.* A proposed grading scheme for medullary thyroid carcinoma based on proliferative activity (Ki-67 and mitotic count) and coagulative necrosis. *Am. J. Surg. Pathol.* 2020; **44**; 1419–1428.
9. Xu B, Fuchs TL, Ahmadi S *et al.* International medullary thyroid carcinoma grading system: a validated grading system for medullary thyroid carcinoma. *J. Clin. Oncol.* 2022; **40**; 96–104.
10. Podany P, Meiklejohn K, Garritano J *et al.* Grading system for medullary thyroid carcinoma: an institutional experience. *Ann. Diagn. Pathol.* 2023; **64**; 152112.
11. Vissio E, Maletta F, Fissore J *et al.* External validation of three available grading systems for medullary thyroid carcinoma in a single institution cohort. *Endocr. Pathol.* 2022; **33**; 359–370.
12. Torricelli F, Santandrea G, Botti C *et al.* Medullary thyroid carcinomas classified according to the international medullary carcinoma grading system and a surveillance, epidemiology, and end results-based metastatic risk score: a correlation with genetic profile and angioinvasion. *Mod. Pathol.* 2023; **36**; 100244.
13. Ho AS, Wang L, Palmer FL *et al.* Postoperative nomogram for predicting cancer-specific mortality in medullary thyroid cancer. *Ann. Surg. Oncol.* 2015; **22**; 2700–2706.
14. Jin L, Zhang X, Ni S *et al.* A nomogram to predict lateral lymph node metastases in lateral neck in patients with medullary thyroid cancer. *Front. Endocrinol. (Lausanne)* 2022; **16**; 902546.
15. Guan YJ, Fang SY, Chen LL, Li ZD. Development and validation of prognostic nomograms for medullary thyroid cancer. *Onco. Targets Ther.* 2019; **27**; 2299–2309.
16. Chen L, Wang Y, Zhao K, Wang Y, He X. Postoperative nomogram for predicting cancer-specific and overall survival among patients with medullary thyroid cancer. *Int. J. Endocrinol.* 2020; **22**; 8888677.
17. Le MK, Kawai M, Odate T, Vuong HG, Oishi N, Kondo T. Metastatic risk stratification of 2526 medullary thyroid carcinoma patients: a study based on surveillance, epidemiology, and end results database. *Endocr. Pathol.* 2022; **33**; 348–358.
18. Lubin DJ, Berhrman DB, Goyal S *et al.* Independent validation of the international grading system for medullary thyroid carcinoma: a single institution experience mod Pathol. 2023; **36**; 100235.
19. Najdawi F, Ahmadi S, Capelletti M, Dong F, Chau NG, Barletta JA. Evaluation of grade in a genotyped cohort of sporadic medullary thyroid carcinomas. *Histopathology* 2021; **79**; 427–436.
20. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing; 2023. https://www.R-project.org/. Accessed 1 November 2023.
21. Chang W, Cheng J, Allaire J, Xie Y, McPherson J. Package 'shiny'. 2015 http://citeseerx.ist.psu.edu/viewdoc/download. Accessed 1 November 2023.
22. Harrell FE, Lee KL, Mark DB. Multivariable prognostic models: issues in developing and validating a model. *Stat. Med.* 1996; **15**; 361–387.

23. Steyerberg EW, Harrell FE Jr, Borsboom GJ. Internal validation of predictive models: efficiency and optimism in bootstrapped and shrunken estimates. *Stat. Med.* 2006; **25**; 3447–3475.
24. Hilden J, Pencina M. Calibration of prognostic models: a review of methods and applications. *Stat. Methods Med. Res.* 2007; **16**; 225–241.
25. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation* 2007; **115**; 928–935.
26. Osamura RY, Gill AJ, Tallini G *et al*. Medullary thyroid carcinoma. In *WHO Classification of Tumours Editorial Board. Endocrine and neuroendocrine tumours [Internet]*. Vol. **10**. 5th ed. Lyon: International Agency for Research on Cancer, 2022 WHO classification of tumours series. Updated 14 August 2023. Available at: https://tumourclassification.iarc.who.int/chapters/53/50. Accessed 1 November 2023.
27. Ghossein R, Barletta JA, Bullock M *et al*. Data set for reporting carcinoma of the thyroid: recommendations from the international collaboration on cancer reporting. *Hum. Pathol.* 2021; **110**; 62–72.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Data S1.