BMC Cancer

RESEARCH Open Access

Development of a clinical-molecular prediction model for central lymph node metastasis in cN0 stage papillary thyroid microcarcinoma: a retrospective study

Jinqiu Wang[1], Weida Fu[1], Jin Luo[1], Mingze Wei[1] and Yongping Dai[1*]

Abstract

Background Identifying occult central lymph node metastasis (CLNM) is essential for guiding prophylactic lymph node dissection (PLND) in patients with cN0 stage papillary thyroid microcarcinoma (PTMC). This study aimed to identify molecular prognostic biomarkers associated with PTMC and develop a clinical-molecular prediction model for CLNM.

Methods Differentially expressed genes (DEGs) in PTMC were identified through bioinformatics analysis of the TCGA database. Prognostic DEGs were selected using Cox and LASSO regression analyses, and a risk-scoring model was constructed based on these genes. The prognostic value of the model was validated using Kaplan-Meier survival analysis and ROC curves. DEG expression levels were compared between patients with CLNM and those without (NCLNM). Clinical data and surgical specimens were collected from 404 patients with cN0 stage PTMC treated at the First Affiliated Hospital of Ningbo University in 2022. The cohort was randomly divided into a derivation cohort ($n=323$) and a validation cohort ($n=81$). DEG expression was quantified using RT-qPCR. Univariate and multivariate logistic regression analyses were conducted in the derivation cohort to identify predictors of CLNM and develop a predictive model. The model's performance was evaluated using the Hosmer-Lemeshow test, ROC curves, calibration curves, and decision curve analysis (DCA).

Results In the TCGA database, FN1, MT-1 F, and TFF3 were identified as prognostic biomarkers. Risk scores based on these genes achieved AUCs of 0.789 (5 years) and 0.674 (10 years) for predicting disease-free survival. Furthermore, FN1, MT-1 F, and TFF3 expression levels were significantly higher in the CLNM group compared to the NCLNM group. Among the 404 PTMC patients, the incidence of CLNM was 42.6% ($n=172$). RT-qPCR analysis demonstrated significantly elevated expression of FN1 in both PTMC tissues compared to normal tissues and in the CLNM group relative to the NCLNM group, while MT-1 F and TFF3 exhibited markedly reduced expression levels. In the derivation cohort, FN1, MT-1 F, TFF3, tumor size ≥5 mm, calcification, multifocality, and extrathyroidal extension were independent predictors of CLNM. The prediction model based on these factors showed AUCs of 0.736 (derivation

*Correspondence:
Yongping Dai
daiyongp1969@163.com

Full list of author information is available at the end of the article

cohort) and 0.813 (validation cohort). Moreover, calibration curves, the Hosmer-Lemeshow test ($\chi^2 = 2.411$, $P = 0.966$), and DCA confirmed the model's robust performance and clinical utility.

Conclusion FN1, MT-1 F, and TFF3 are valuable prognostic biomarkers for PTMC. The clinical-molecular prediction model incorporating these genes provides a basis for personalized PLND decision-making in cN0 stage PTMC patients.

Trial registration number Not applicable.

Keywords Papillary thyroid microcarcinoma, Lymph node metastasis, Molecular markers, Prediction model

Introduction

The global incidence of thyroid cancer has increased significantly in recent years [1]. According to the latest Chinese cancer statistics, thyroid cancer is the fourth most common malignancy among women in China, with an annual growth rate of approximately 20% [2]. Papillary thyroid carcinoma (PTC) is the predominant histological subtype, accounting for 90% of all thyroid cancer cases [3]. The World Health Organization defines papillary thyroid microcarcinoma (PTMC) as PTC with a maximum diameter ≤1.0 cm [4]. Although PTMC generally follows an indolent course with a favorable prognosis, 20-50% of patients develop cervical lymph node metastasis, primarily in the central compartment [5, 6]. As metastasis progresses, the 5-, 10-, and 15-year survival rates decline to 82.2%, 63.8%, and 23.9%, respectively [7]. Notably, 60-80% of lymph node metastases in PTC are occult, and preoperative ultrasound sensitivity for detecting central lymph node metastases (CLNM) remains low at 20-30%, making most occult metastases difficult to identify [8, 9].

Guidelines regarding prophylactic lymph node dissection (PLND) in cN0 stage PTMC patients—those without evident signs of cervical lymph node metastasis prior to surgery—vary across regions. Several guidelines, including Chinese guidelines, recommend active PLND [10]. While PLND reduces tumor recurrence and complications associated with reoperation, it carries risks such as permanent hypocalcemia and recurrent laryngeal nerve injury [11]. Therefore, accurately predicting CLNM before surgery is essential to optimize patient management, minimize unnecessary surgical interventions, and reduce procedure-related complications.

Although various predictive tools have been developed, most rely primarily on clinical features and radiomic signatures derived from ultrasound or computed tomography, with limited inclusion of molecular biomarkers [12, 13]. Previous studies by Lu et al. [14] and Xiao et al. [15] identified several metastasis-related genes in PTC through bioinformatics analysis, suggesting genetic alterations in lymph node metastasis. Integrating sensitive molecular markers that accurately reflect tumor biological behavior into existing clinicopathological models could enhance personalized monitoring and risk stratification. While some researchers have developed prediction models incorporating risk factor genes associated with CLNM, these models are primarily limited to commonly tested mutations such as BRAF V600E and TERT promoter mutations [16–18]. Given that CLNM is a crucial prognostic factor for PTMC [19], we hypothesized that differentially expressed genes (DEGs) associated with prognosis may serve as potential biomarkers for CLNM. Therefore, this study aimed to identify molecular prognostic biomarkers in PTMC and investigate their application in constructing a clinical-molecular prediction model for CLNM in PTMC patients.

Methods

Data source and RNA-seq data analysis

RNA-seq transcriptome data and prognostic information for 371 PTMC samples and 59 normal tissue samples were obtained from The Cancer Genome Atlas (TCGA) database (https://cancergenome.nih.gov/). The median follow-up time was 31 months (range: 2-169 months), with a missing data rate of 0.27%. Sample selection criteria included a confirmed PTMC diagnosis and no evidence of distant metastasis.

To identify differentially expressed mRNAs, we utilized the edgeR package, applying multiple testing corrections with the false discovery rate (FDR). The cut-off criteria were $|\log_2 \text{fold change (FC)}| > 2$ and FDR < 0.01.

Participants

This study included PTMC patients who underwent radical thyroidectomy at the First Affiliated Hospital of Ningbo University between January and December 2022. The inclusion criteria were as follows: (1) age ≥ 18 years; (2) diagnosis of clinically node-negative (cN0 stage) PTMC and treatment with radical thyroidectomy; and (3) availability of surgical specimens and postoperative pathological reports. The exclusion criteria were: (1) coexisting malignancies or severe diseases, such as liver or renal dysfunction, that could affect prognosis; and (2) incomplete clinical data.

A total of 647 PTMC patients underwent radical thyroidectomy at our hospital from January to December 2022. After excluding 132 patients with preoperative imaging indicating central or lateral neck lymph node metastasis, 105 patients with severe comorbidities, and 6

patients with incomplete data, the final cohort comprised 404 patients. Among them, 287 (71.0%) were female and 117 (29.0%) were male, with a median age of 48 years (range: 21–74 years). CLNM was identified in 172 cases (42.6%).

RNA extraction and gene expression analysis

Surgical specimens, including tumor and normal tissues from 404 PTMC patients, were stored at -80°C post-excision. Tissue samples were pretreated, and RNA was extracted using Trizol (15596-026, Invitrogen Company). Reverse transcription quantitative PCR (RT-qPCR) was performed using 2×ChamQ Universal SYBR qPCR Master Mix (Q711, Vazyme) following standard protocols. Primer sequences, designed using data from the National Center for Biotechnology Information website (https://www.ncbi.nlm.nih.gov/), are listed in Supplementary Table S1 [20]. Gene expression levels were analyzed using the $2^{\wedge(-\Delta\Delta Ct)}$ method.

Data collection

Clinical and pathological data from the 404 patients were collected from medical records, including sex, age, tumor size, multifocality, tumor boundary, aspect ratio, blood flow, calcification, extrathyroidal extension, and Hashimoto's thyroiditis (HT).

Statistical analysis

Statistical analyses were performed using SPSS 26.0 and R 4.1.3 software. In the TCGA dataset, univariate Cox regression analysis was used to identify DEGs significantly associated with disease-free survival (DFS) in PTMC patients. Feature selection was conducted using the least absolute shrinkage and selection operator (LASSO) method, followed by stepwise multivariate Cox regression analysis to construct a prognostic risk score model. Patients were stratified based on risk scores, and Kaplan-Meier survival curves and receiver operating characteristic (ROC) curves were generated. One-way ANOVA was used to compare DEG expression levels across different groups.

In the 404-patient dataset, the Wilcoxon rank-sum test was applied to compare gene expression levels across different tissue types. Patients were randomly assigned to derivation and validation cohorts in a 4:1 ratio [21, 22]. In the derivation cohort, univariate and multivariate logistic regression analyses were performed to identify independent factors associated with CLNM. A clinical-molecular prediction model incorporating these factors was developed and visualized as a nomogram. Model performance was assessed using ROC curves, calibration curves, decision curve analysis (DCA), and the Hosmer-Lemeshow test in both cohorts. Statistical significance was set at $P < 0.05$.

Ethics consideration

This study was conducted in accordance with the Declaration of Helsinki and was approved by the Ethics Committee of the First Affiliated Hospital of Ningbo University (Approval No. 2021-R140). The TCGA data used in this study are publicly accessible through the National Cancer Institute's Genomic Data Commons portal (https://portal.gdc.cancer.gov/). Written informed consent was obtained from all participants.

Results

Differentially expressed mRNAs associated with the prognosis of PTMC

We retrieved genetic and clinical data from 371 PTMC patients in the TCGA database. Differential expression analysis identified 1510 differentially expressed mRNAs between tumor and normal tissues, including 704 up-regulated and 806 down-regulated genes (Fig. 1A-B). Univariate Cox analysis revealed that 212 of these genes were significantly associated with DFS ($P < 0.05$). Subsequently, LASSO regression identified three differentially expressed mRNAs with prognostic significance: FN1, MT-1 F, and TFF3 (Fig. 1C). Among these, higher FN1 expression correlated with poorer prognosis, while lower expression of MT-1 F and TFF3 was also associated with worse outcomes (Fig. 1D-F).

A prognostic model was developed using stepwise multivariate Cox regression analysis based on these three genes. The risk score formula was defined as follows: Risk Score = 0.1297 × FN1 + (-0.1511) × MT-1 F + (-0.0052) × TFF3. Based on the risk scores, PTMC patients were stratified into high-risk and low-risk groups (Fig. 2A). Survival analysis demonstrated that the low-risk group exhibited significantly better DFS compared to the high-risk group ($P = 0.015$, Fig. 2B). Additionally, the time-dependent ROC curves for 5- and 10-year DFS yielded AUCs of 0.789 and 0.674, respectively (Fig. 2C).

Correlation of FN1, MT-1 F, and TFF3 expression with CLNM in PTMC patients

The poor prognosis in PTMC patients is significantly associated with CLNM [23–25]. Previous studies have showed that lymph node metastasis increases the risk of recurrence in PTMC patients by threefold [26]. Consequently, we further investigated the relationship between FN1, MT-1 F, TFF3, and CLNM using TCGA dataset. As shown in Fig. 3A-C, the cohort included 208 patients with CLNM and 163 patients without CLNM (NCLNM). FN1 expression was significantly upregulated in the CLNM group versus both NCLNM and normal tissue groups ($P < 0.001$), while MT-1 F and TFF3 were markedly down-regulated ($P < 0.001$). Moreover, as depicted in Fig. 3D-F, the AUCs for FN1, MT-1 F, and TFF3 in discriminating CLNM were 0.642, 0.656, and 0.652, respectively. These

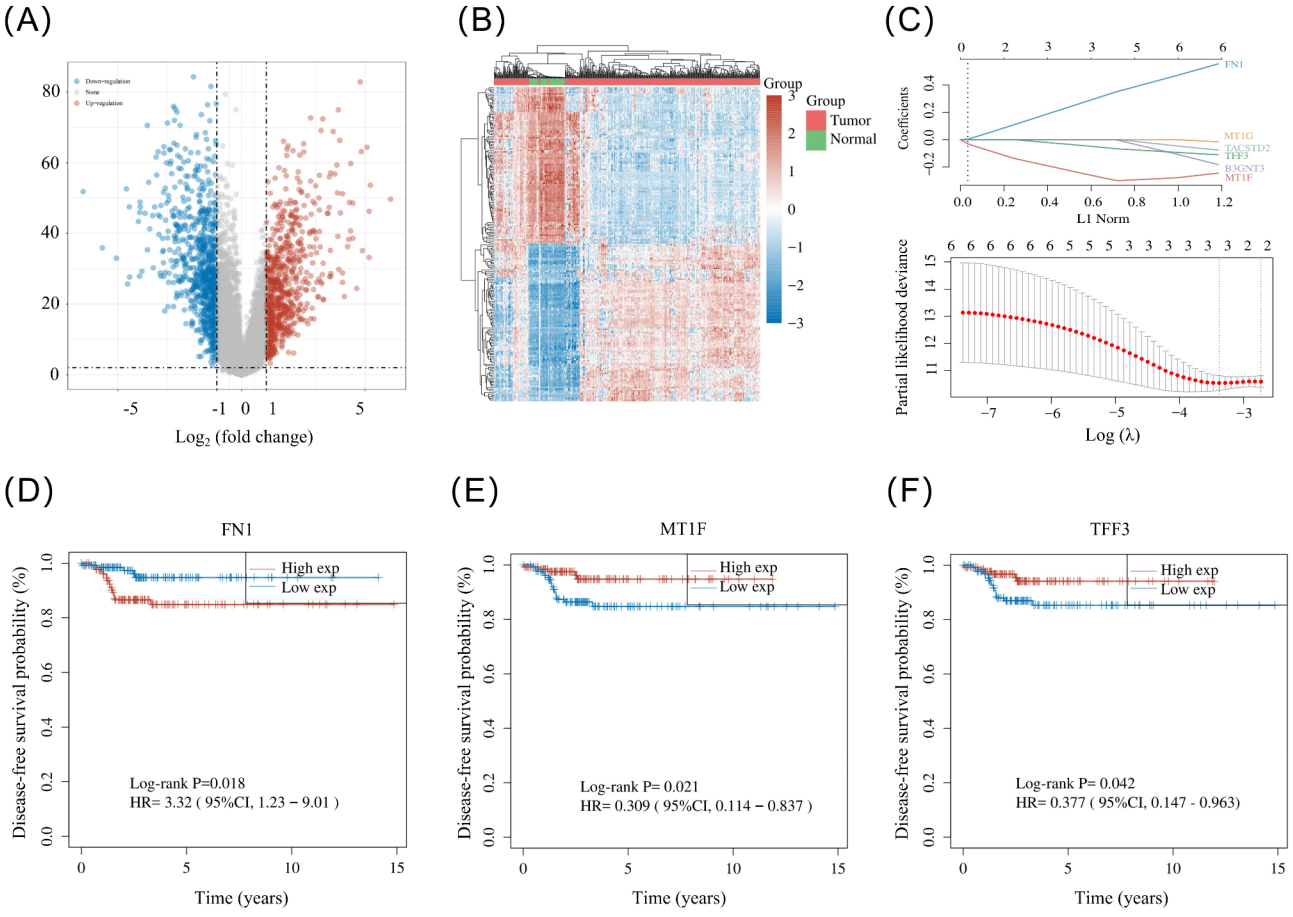

Fig. 1 Differentially expressed mRNAs associated with the prognosis of PTMC. (**A**) Volcano plot of differentially expressed mRNAs between tumor tissues and normal tissues. Red spots represent up-regulated genes, and blue spots represent down-regulated genes. (**B**) Heatmap of differentially expressed mRNAs between tumor tissues and normal tissues. (**C**) OS-related mRNA selection using the least absolute shrinkage and selection operator (LASSO) regression. (**D-F**) Kaplan-Meier survival curves for PTMC patients with high or low expression of FN1 (**D**), MT-1 F (**E**), and TFF3 (**F**) in the TCGA dataset

findings suggest that these genes may serve as potential biomarkers for CLNM in PTMC.

Expression of FN1, MT-1 F, and TFF3 in tissue samples

Analysis of the TCGA dataset identified FN1, MT-1 F, and TFF3 as prognostic genes associated with PTMC. In tumor tissues, high FN1 expression and low MT-1 F and TFF3 expression were correlated with CLNM and worse DFS. To validate these findings, tissue samples from 404 PTMC patients were collected and analyzed. RT-qPCR results indicated that FN1 expression in tumor tissues (2.15 ± 0.22) was significantly higher than that in normal tissues (1.11 ± 0.15, $P < 0.001$). Conversely, MT-1 F expression in tumor tissues (0.70 ± 0.11) was significantly lower than that in normal tissues (0.84 ± 0.24, $P < 0.001$). Similarly, TFF3 expression in tumor tissues (0.47 ± 0.26) was markedly lower than in normal tissues (1.22 ± 0.61, $P < 0.001$). Subsequently, based on the postoperative pathological results, patients were divided into CLNM and NCLNM groups. Consistent with the TCGA findings, the expression levels of FN1, MT-1 F, and TFF3 in tumor tissues showed significant differences between the CLNM and NCLNM groups (Table 1).

Influencing factor analysis of CLNM in PTMC patients

Next, we assessed whether the three genes were independent factors influencing CLNM and could be used to develop a predictive model for CLNM. The data from 404 patients were divided into a derivation cohort ($n = 323$) and a validation cohort ($n = 81$). The characteristics of patients in both cohorts are summarized in Table 2. In the derivation cohort, univariate logistic analysis revealed that tumor size ≥ 5 mm, unclear boundaries, calcification, extrathyroidal extension, high FN1 expression, low MT-1 F expression, and low TFF3 expression were factors associated with CLNM ($P < 0.05$), whereas sex, age, HT, multifocality, aspect ratio ≥ 1, and blood flow were not ($P > 0.05$, Table 3). These significant factors were then incorporated into multivariate logistic regression, which identified all seven variables as independent predictors of CLNM ($P < 0.05$, Table 3).

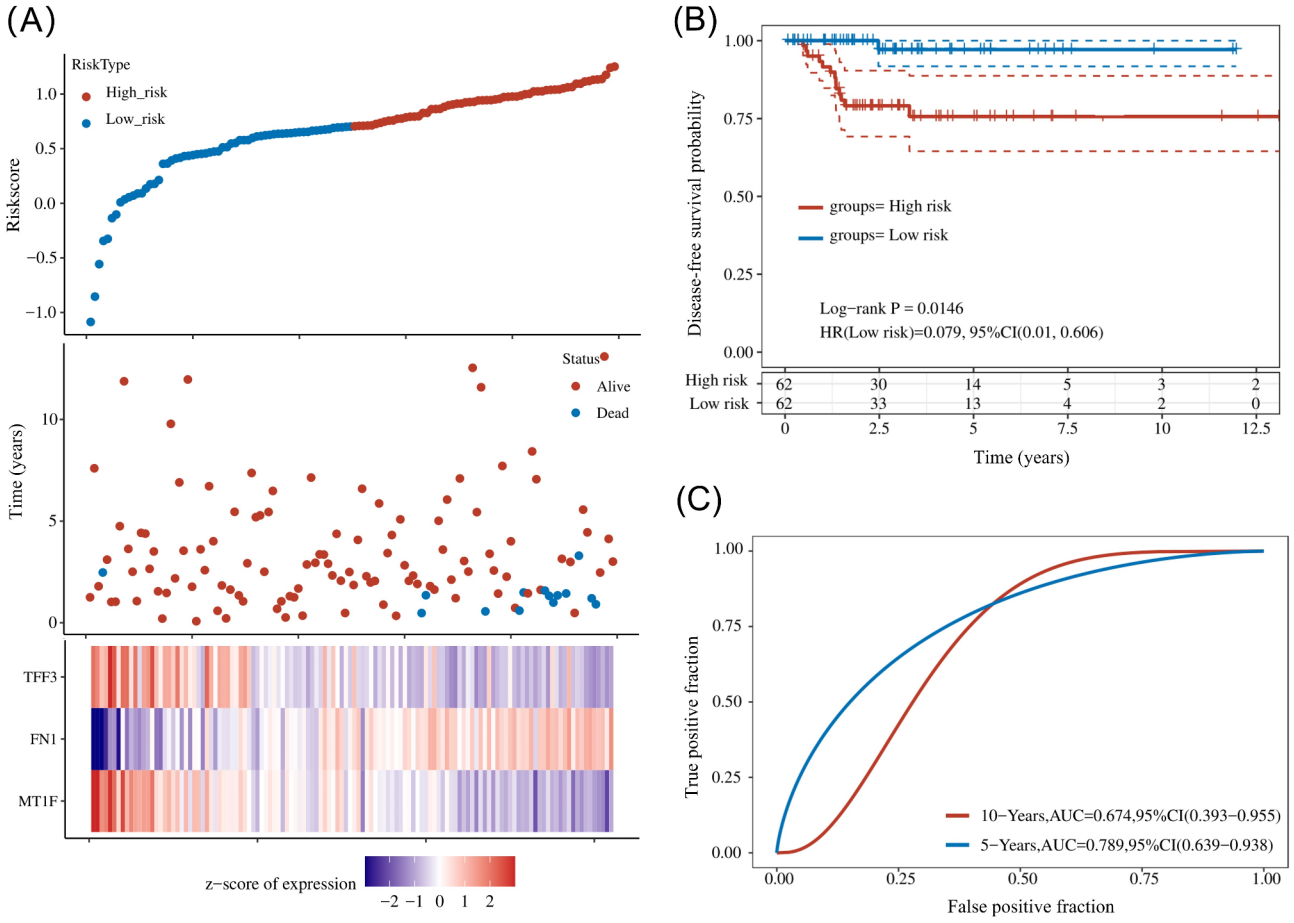

Fig. 2 Risk scores based on the differentially expressed mRNAs. (**A**) Risk scores, survival time, and survival status in the TCGA dataset. Top: scatterplot of risk scores from low to high; middle: scatterplot distribution of survival time and survival status corresponding to risk scores of different samples; bottom: heat map of gene expression in the prognostic model. (**B**) Kaplan-Meier survival curves for high-risk and low-risk PTMC patients in the TCGA dataset. (**C**) Time-dependent ROC curve for 5-year and 10-year OS of the PTMC patients in TCGA dataset

Construction of a clinical-molecular predictive model for CLNM in PTMC patients

A logistic regression model was developed based on the seven aforementioned factors, with the equation: Logit(*P*) = -3.098 + 0.626 × Tumor size ≥ 5 mm + 0.696 × Unclear boundaries + 0.858 × Calcification + 0.847 × Extrathyroidal extension + 0.951 × High FN1 expression + 0.880 × Low MT-1 F expression + 0.836 × Low TFF3 expression. To enhance clinical applicability, this model was visualized as a nomogram (Fig. 4). When using the nomogram, the corresponding score for each variable is determined based on the patient's characteristics. These scores are summed to obtain a total point, which provides an estimated risk of CLNM.

Validation of the CLNM prediction model for PTMC patients

The performance of the prediction model was evaluated using data from the derivation and validation cohorts. No significant differences were observed between the seven key features in the two cohorts (Table S2), confirming their comparability. In the derivation cohort, the ROC curve resulted in an AUC of 0.736 (95% confidence interval [CI]: 0.681–0.790), with a specificity of 71.7%, sensitivity of 66.2%, and accuracy of 69.3% (Fig. 5A). In the validation cohort, the AUC was 0.813 (95% CI: 0.720–0.905), with a specificity of 83.3%, sensitivity of 69.7%, and accuracy of 77.8% (Fig. 5D). Calibration curves demonstrated good agreement between predicted and actual probabilities in both derivation and validation cohort (Fig. 5B&E), further supported by the Hosmer-Lemeshow test (χ^2 = 2.411, P = 0.966). DCA indicated that the model offered a clinical net benefit across a broad threshold probability range (0.20–0.85 in the derivation cohort, Fig. 5C; and 0.1–0.90 in the validation cohort, Fig. 5F).

Discussion

In recent years, there has been a surge in the detection rates of PTMC, largely due to increased public health awareness and advancements in medical diagnostic

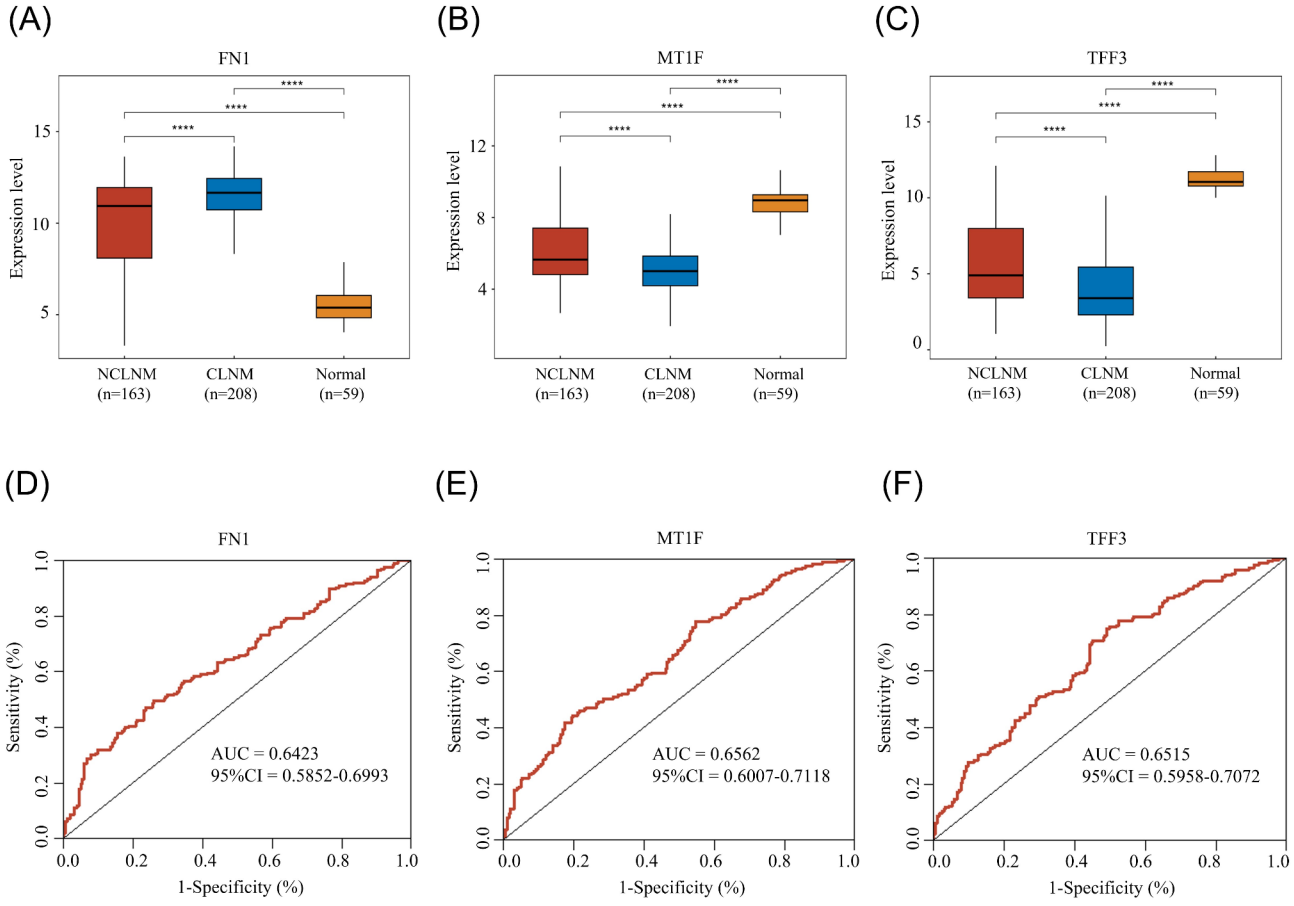

Fig. 3 (**A-C**) Comparison of expression levels of FN1 (**A**), MT-1 F (**B**), and TFF3 (**C**) in CLNM, NCLNM, and normal tissues of PTMC patients. (**D-F**) ROC curves of FN1 (**D**), MT-1 F (**E**), and TFF3 (**F**) for predicting CLNM in PTMC patients

Table 1 Comparison of FN1, MT-1 F, and TFF3 expression levels in the different groups ($n=404$)

Categories	FN1	MT-1 F	TFF3
Tumor tissue ($n=404$)	2.151 ± 0.221	0.695 ± 0.109	0.467 ± 0.262
Normal tissue ($n=404$)	1.107 ± 0.147	0.835 ± 0.235	1.222 ± 0.607
P-value	< 0.001	< 0.001	< 0.001
CLNM group ($n=172$)	2.190 ± 0.210	0.678 ± 0.099	0.419 ± 0.258
NCLNM group ($n=232$)	2.122 ± 0.225	0.708 ± 0.114	0.503 ± 0.259
P-value	0.001	0.005	0.001

CLNM, central lymph node metastasis; NCLNM, non-central lymph node metastasis

technologies [2]. Although the majority of PTMC cases progress slowly and often remain indolent, a notable subset of patients still faces the risk of developing CLNM, which profoundly impacts patient outcomes [26–28]. While numerous studies have explored methods to predict lymph node metastasis, the emergence of precision medicine highlights the critical need for molecular-level predictions of CLNM [29–31].

In this study, we conducted bioinformatics analysis using data from the TCGA database and identified FN1, MT-1 F, and TFF3 as prognostic biomarkers for PTMC. Among them, high FN1 expression was associated with poorer prognosis, while low expression of MT-1 F and TFF3 correlated with worse outcomes. In our real-world cohort of 404 PTMC patients, the incidence of CLNM was 42.6% ($n=172$), slightly higher than the 38.7% rate reported by Shi et al. [6]. RT-qPCR analysis of surgical specimens from these patients validated the findings from the TCGA dataset. Compared with adjacent normal tissues, tumor tissues exhibited higher expression of FN1 and lower expression of MT-1 F and TFF3. Moreover, we found that FN1 was further overexpressed in the PTMC group with CLNM compared to the group without CLNM, while MT-1 F and TFF3 showed the opposite pattern. These findings suggested that FN1 may promote tumorigenesis and CLNM in PTMC, while MT-1 F and TFF3 may play inhibitory roles [32].

Fibronectin (FN) is an extracellular matrix protein synthesized by fibroblasts, and its family member FN1 is recognized as one of the important negative prognostic markers in PTC patients [33]. Previous studies have demonstrated elevated FN1 expression in PTC, with aberrant overexpression of FN1 linked to decreased recurrence-free survival [34, 35]. Our sample analysis

Table 2 Characteristic of patients in derivation and validation cohort

Variables	Derivation cohort ($n=323$)			Validation cohort ($n=81$)		
	CLNM group ($n=139$)	NCLNM group ($n=184$)	*P*-value	CLNM group ($n=33$)	NCLNM group ($n=48$)	*P*-value
Sex			0.808			0.912
Male	41 (29.5%)	52 (28.3%)		10 (30.3%)	14 (29.2%)	
Female	98 (70.5%)	132 (71.7%)		23 (69.7%)	34 (70.8%)	
Age, year			0.313			0.503
< 45	49 (35.3%)	75 (40.8%)		10 (30.3%)	18 (37.5%)	
≥ 45	90 (64.7%)	109 (59.2%)		23 (69.7%)	30 (62.5%)	
HT			0.368			0.654
Yes	55 (39.6%)	82 (44.6%)		16 (48.5%)	25 (52.1%)	
No	84 (60.4%)	102 (55.4%)		17 (51.5%)	23 (47.9%)	
Tumor size, mm			0.022			0.055
< 5	34 (24.5%)	67 (36.4%)		7 (21.2%)	20 (41.7%)	
≥ 5	105 (75.5%)	117 (63.6%)		26 (78.8%)	28 (58.3%)	
Multifocality			0.620			0.524
Yes	62 (44.6%)	77 (41.8%)		14 (42.4%)	17 (35.4%)	
No	77 (55.4%)	107 (58.2%)		19 (57.6%)	31 (64.6%)	
Boundary			0.034			0.414
Clear	90 (64.7%)	139 (75.5%)		22 (66.7%)	36 (75.0%)	
Unclear	49 (35.3%)	45 (24.5%)		11 (33.3%)	12 (25.0%)	
Aspect ratio			0.309			0.863
< 1	76 (54.7%)	111 (60.3%)		20 (60.6%)	30 (62.5%)	
≧ 1	63 (45.3%)	73 (39.7%)		13 (39.4%)	18 (37.5%)	
Blood flow			0.581			0.136
Yes	72 (51.8%)	101 (54.9%)		20 (60.6%)	21 (43.8%)	
No	67 (48.2%)	83 (45.1%)		13 (39.4%)	27 (56.3%)	
Calcification			0.037			0.393
Yes	79 (56.8%)	83 (45.1%)		19 (57.6%)	23 (47.9%)	
No	60 (43.2%)	101 (54.9%)		14 (42.4%)	25 (52.1%)	
Extrathyroidal extension			0.008			0.337
Yes	27 (19.4%)	17 (9.2%)		5 (15.2%)	4 (8.3%)	
No	112 (80.6%)	167 (90.8%)		28 (84.8%)	44 (91.7%)	
FN1 expression*			0.003			0.007
High	103 (74.1%)	107 (58.2%)		25 (75.8%)	22 (45.8%)	
Low	36 (25.9%)	77 (41.8%)		8 (24.2%)	26 (54.2%)	
MT-1 F expression*			< 0.001			0.004
High	30 (21.6%)	72 (39.1%)		5 (15.2%)	22 (45.8%)	
Low	109 (78.4%)	112 (60.9%)		28 (84.8%)	26 (54.2%)	
TFF3 expression*			< 0.001			0.057
High	63 (45.3%)	119 (64.7%)		15 (45.5%)	32 (66.7%)	
Low	76 (54.7%)	65 (35.3%)		18 (54.5%)	16 (33.3%)	

HT, Hashimoto's thyroiditis

* The cutoff values for FN1, MT-1 F, and TFF3 were 2.0350, 0.7655, and 0.4205, respectively

further confirmed that FN1 is significantly upregulated in tumor tissues of PTMC patients with CLNM. This may be attributed to the role of FN1 in promoting cell proliferation, migration, invasion, and inducing epithelial-mesenchymal transition [36, 37]. Metallothioneins (MT) are a family of low-molecular-weight, cysteine-rich proteins with metal-binding properties, widely expressed in human tissues and organs [38]. Notably, the expression and role of MT in different types of malignancies remain controversial. Several studies have reported increased MT expression in human tumors of the kidney, uterus, and breast, whereas decreased MT expression has been observed in tumors of the thyroid, liver, and stomach [39, 40]. To date, little research has specifically focused on MT-1 F in the context of PTMC. However, a study by Chen et al. identified MT-1 M, a homolog of MT-1 F, as an independent risk factor for lymph node metastasis in PTC. Their in vitro experiments demonstrated that

Table 3 Univariate and multivariate logistic regression analysis in the derivation cohort

Variables	Univariate analysis		Multivariate analysis	
	OR (95% CI)	*P*-value	OR (95% CI)	*P*-value
Male	1.062 (0.653–1.726)	0.808		
Age < 45	1.264 (0.801–1.993)	0.314		
HT	0.814 (0.521–1.274)	0.369		
Tumor size ≧ 5 mm	1.768 (1.084–2.886)	0.022	1.871 (1.093-3.200)	0.022
Multifocality	1.119 (0.717–1.745)	0.620		
Unclear boundary	1.682 (1.037–2.728)	0.035	2.006 (1.176–3.421)	0.011
Aspect ratio ≧ 1	1.260 (0.807–1.969)	0.309		
Blood flow	0.883 (0.568–1.373)	0.581		
Calcification	1.602 (1.028–2.497)	0.037	2.359 (1.420–3.918)	< 0.001
Extrathyroidal extension	2.368 (1.233–4.547)	0.010	2.334 (1.144–4.761)	0.020
High FN1 expression	2.059 (1.275–3.325)	0.003	2.588 (1.519–4.409)	< 0.001
Low MT-1 F expression	2.336 (1.415–3.855)	< 0.001	2.410 (1.403–4.140)	0.001
Low TFF3 expression	2.209 (1.408–3.465)	< 0.001	2.307 (1.408–3.779)	< 0.001

HT, Hashimoto's thyroiditis

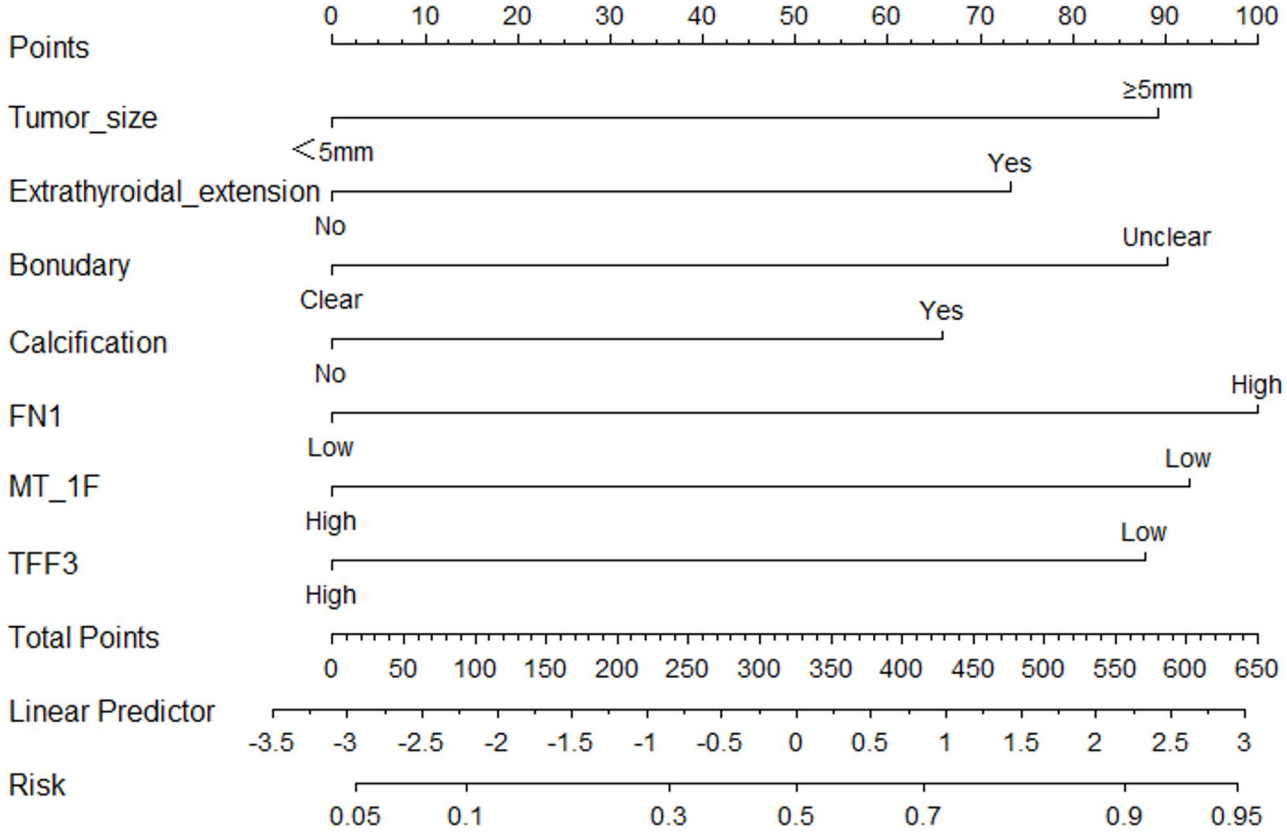

Fig. 4 The nomogram of CLNM prediction model for patients with PTMC

upregulation of MT-1 M inhibited colony formation, proliferation, migration, and invasion of PTC cell lines [41]. Our study also revealed that low MT-1 F expression is associated with poor prognosis and CLNM in PTMC, although its molecular mechanisms warrant further investigation. TFF3, a member of the trefoil factor (TFF) family, plays a role in mucosal repair, signal transduction, and regulation of apoptosis [42]. Previous studies have demonstrated that low TFF3 expression in thyroid cancer is associated with lymph node metastasis and increased tumor aggressiveness [43, 44]. Mechanistically, the silencing of TFF3 may activate the IL-6/JAK/STAT3 signaling pathway, thereby promoting cell proliferation, migration, angiogenesis, and evading immune surveillance [45].

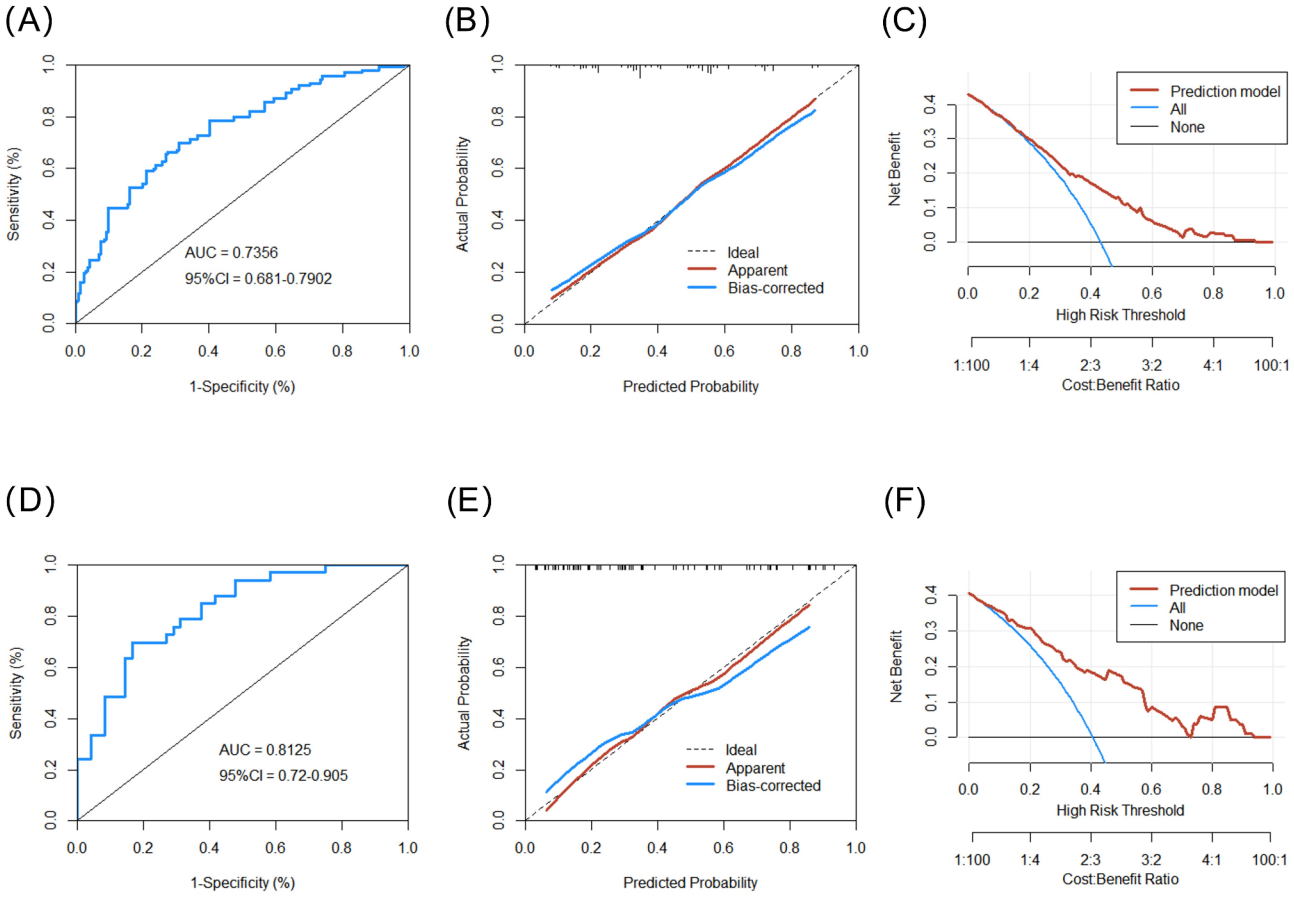

Fig. 5 Validation of the CLNM prediction model for patients with PTMC. (**A-C**) ROC curve (**A**), calibration curve (**B**), and decision curve analysis (**C**) of the CLNM prediction model in the derivation cohort; (**D-F**) ROC curve (**D**), calibration curve (**E**), and decision curve analysis (**F**) of the CLNM prediction model in the validation cohort

Previous multivariate analyses have identified factors such as tumor size ≥ 5 mm, calcification, multifocality, and extrathyroidal extension as significant predictors of CLNM [46]. In this study, these clinicopathological features, along with three key genes (FN1, TFF3, and MT-1 F), were found to be independent predictors of CLNM. The CLNM prediction model developed based on these features differs from existing systems in the following aspects: Firstly, the model overcomes the limitations of traditional approaches. Current CLNM risk prediction tools mainly fall into two categories: clinical factor-based models and radiomics-based models [47–49]. However, models relying solely on clinical factors exhibit relatively low accuracy, and some tumor imaging features are susceptible to imaging parameters and interpreter experience [21]. While a limited number of studies have incorporated common gene mutations associated with PTMC, such as BRAF V600E and TERT promoter mutations, the roles of these mutations in CLNM remains controversial [50, 51]. This study, conducted on a medium-sized cohort, quantitatively measured the expression of prognostic genes identified from the TCGA database and integrated them with clinical parameters to develop a novel clinical-molecular prediction model. Although the prognostic relevance of FN1 and TFF3 in PTMC has been reported previously, this is the first time they have been combined with MT-1 F for CLNM prediction. Secondly, the new model greatly improves prediction performance for CLNM. Due to the difficult anatomical visualization of central lymph nodes by ultrasound, traditional ultrasound detects only 20–31% of CLNM cases [9]. In contrast, the prediction model demonstrated accuracy rates of 69.3% and 78.0% in the derivation and validation sets, respectively. Furthermore, the AUCs of the model (0.736 and 0.813) slightly outperform those of other existing CLNM prediction models (AUCs: 0.656 ~ 0.757) [21, 49, 52]. The model was evaluated using ROC curves, calibration curves, and DCA in an internal validation cohort, which indicated its clinical utility. In the validation cohort, the model showed high accuracy and specificity, correctly identifying 83.3% of negative cases, which is crucial for avoiding unnecessary PLND. In clinical practice, molecular marker testing (FN1, MT-1 F, and TFF3) can be performed on preoperative fine needle

aspiration (FNA) samples. By integrating ultrasound features, the probability of CLNM can be calculated using the nomogram, enabling individualized risk assessment that informs surgical decision-making and improves patient outcomes.

However, this study has certain limitations. First of all, it was a single-center retrospective study, and the predictive model developed has not been externally validated. Although multiple evaluation metrics were employed in internal validation to ensure a robust assessment of the model's performance within the available dataset, the lack of an independent external validation cohort may limit the generalizability of our findings. Second, the molecular biomarkers selected were not further validated at the mechanistic level. Additionally, our model was developed using surgical samples. While previous studies have confirmed that detection of these genes in FNA samples can be used for preoperative diagnosis or course prediction of PTC [43, 53], their applicability to FNA samples has not been directly verified in this study. Future studies should include larger, multi-center cohorts to validate the predictive model and conduct in vitro experiments to elucidate the underlying biological functions of the identified biomarkers in the progression and metastasis of PTMC. Furthermore, prospective studies using ultrasound-guided FNA samples are needed to evaluate the model's performance in a clinical setting and its potential to guide PLND decisions.

Conclusion

FN1, MT-1 F, and TFF3 were identified as prognostic biomarkers for PTMC. Tumor size ≥5 mm, calcification, multifocality, and extrathyroidal extension, as well as these genes were also independent predictors of CLNM. The clinical-molecular predictive model integrating these factors may provide a useful basis for guiding personalized PLND decisions in patients with cN0 stage PTMC.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12885-025-14112-0.

Supplementary Material 1

Acknowledgements
Not applicable.

Author contributions
J.W. and W.F. conceptualized and designed the study. J.W., W.F. and J.L. conducted data collection and data analysis. M.W. drafted the manuscript. Y.D. supervised the project and revised the manuscript. All authors read and approved the final manuscript.

Funding
This work was supported by the Zhejiang Provincial Medical and Health Science and Technology Program (2022KY1110) and the "Xinyu" Talent Program of the First Affiliated Hospital of Ningbo University (XYJH-1-WJQ).

Data availability
The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
This study was conducted in accordance with the Declaration of Helsinki and was approved by the Ethics Committee of the First Affiliated Hospital of Ningbo University (Approval No. 2021-R140). All TCGA data used in this study are publicly available through the National Cancer Institute's Genomic Data Commons portal (https://portal.gdc.cancer.gov/). Written informed consent was obtained from all participants involved in the study.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Thyroid and Breast Surgery, The First Affiliated Hospital of Ningbo University, No.59 Liuting Street, Haishu District, Ningbo 315010, Zhejiang, China

Received: 25 December 2024 / Accepted: 8 April 2025
Published online: 14 April 2025

References

1. Bray F, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2024;74(3):229–63.
2. Li M, Dal Maso L, Vaccarella S. Global trends in thyroid cancer incidence and the impact of overdiagnosis. Lancet Diabetes Endocrinol. 2020;8(6):468–70.
3. Megwalu UC, Moon PK. Thyroid cancer incidence and mortality trends in the united States: 2000–2018. Thyroid. 2022;32(5):560–70.
4. Lee JS, et al. Aggressive subtypes of papillary thyroid carcinoma smaller than 1 cm. J Clin Endocrinol Metab. 2023;108(6):1370–5.
5. Parvathareddy SK et al. Risk factors for cervical lymph node metastasis in middle Eastern papillary thyroid microcarcinoma. J Clin Med. 2022;11(15).
6. Shi Y, et al. Clinicopathological findings associated with cervical lymph node metastasis in papillary thyroid microcarcinoma: A retrospective study in China. Cancer Control. 2022;29:10732748221084926.
7. Miftari R, et al. Management of the patient with aggressive and resistant papillary thyroid carcinoma. Med Arch. 2016;70(4):314–7.
8. Asimakopoulos P, et al. Management of the neck in Well-Differentiated thyroid cancer. Curr Oncol Rep. 2020;23(1):1.
9. Xing Z, et al. Thyroid cancer neck lymph nodes metastasis: Meta-analysis of US and CT diagnosis. Eur J Radiol. 2020;129:109103.
10. Administration NHCotPsRoCMAaH. Guidelines for the diagnosis and treatment of thyroid carcinoma (2022 version). Chin J Practical Surg. 2022;42(12):1343–63.
11. Alsubaie KM, et al. Prophylactic central neck dissection for clinically Node-Negative papillary thyroid carcinoma. Laryngoscope. 2022;132(6):1320–8.
12. Yu J, et al. Lymph node metastasis prediction of papillary thyroid carcinoma based on transfer learning radiomics. Nat Commun. 2020;11(1):4807.
13. Zhang S, et al. Ultrasound-Base radiomics for discerning lymph node metastasis in thyroid cancer: A systematic review and Meta-analysis. Acad Radiol. 2024;31(8):3118–30.
14. Lu DN, et al. Single-cell and bulk RNA sequencing reveal heterogeneity and diagnostic markers in papillary thyroid carcinoma lymph-node metastasis. J Endocrinol Invest. 2024;47(6):1513–30.
15. Xiao G, et al. Single-cell RNA-sequencing and Spatial transcriptomic analysis reveal a distinct population of APOE(-) cells yielding pathological lymph node metastasis in papillary thyroid cancer. Clin Transl Med. 2025;15(1):e70172.
16. Zhang R, et al. Targeted sequencing of DNA/RNA combined with radiomics predicts lymph node metastasis of papillary thyroid carcinoma. Cancer Imaging. 2024;24(1):75.

17. Chen H, et al. Prediction model of cervical lymph node metastasis in papillary thyroid carcinoma. Cancer Control. 2024;31:10732748241295347.
18. Zhao F, et al. A LASSO-based model to predict central lymph node metastasis in preoperative patients with cN0 papillary thyroid cancer. Front Oncol. 2023;13:1034047.
19. Wang K, et al. Potential diagnostic of lymph node metastasis and prognostic values of TM4SFs in papillary thyroid carcinoma patients. Front Cell Dev Biol. 2022;10:1001954.
20. Wang QX, et al. A panel of four genes accurately differentiates benign from malignant thyroid nodules. J Exp Clin Cancer Res. 2016;35(1):169.
21. Wang Y, et al. Predicting central cervical lymph node metastasis in papillary thyroid microcarcinoma using deep learning. PeerJ. 2024;12:e16952.
22. Zhao Y, et al. Prediction of central lymph node metastasis in patients with papillary thyroid microcarcinoma by gradient-boosting decision tree model based on ultrasound radiomics and clinical features. Gland Surg. 2023;12(12):1722–34.
23. Rui ZY, et al. A retrospective study of the risk factors and the prognosis in patients with papillary thyroid carcinoma depending on the number of lymph node metastasis. Clin Exp Med. 2021;21(2):277–86.
24. Li N, et al. Correlations of LncRNAs with cervical lymph node metastasis and prognosis of papillary thyroid carcinoma. Onco Targets Ther. 2019;12:1269–78.
25. Hu L, et al. Enhanced stiffness in peri-cancerous tissue: a marker of poor prognosis in papillary thyroid carcinoma with lymph node metastasis. Oncologist. 2024;29(9):e1132–48.
26. Didehban S, Abdollahi A, Meysamie A. Evaluation of etiology, clinical manifestations, diagnosis, Follow-up, histopathology and prognosis factors in papillary thyroid microcarcinoma: A systematic review and Meta-analysis. Iran J Pathol. 2023;18(4):380–91.
27. Nagaoka R, et al. Multifocality and progression of papillary thyroid microcarcinoma during active surveillance. World J Surg. 2021;45(9):2769–76.
28. Malik AA, et al. Characteristics and management of papillary thyroid microcarcinoma in the united Arab Emirates: experience from a large tertiary hospital. Saudi J Med Med Sci. 2022;10(1):42–8.
29. Huang H, et al. A prediction model for identifying high-risk lymph node metastasis in clinical low-risk papillary thyroid microcarcinoma. BMC Endocr Disord. 2023;23(1):260.
30. Huang Y, et al. Exploring risk factors for cervical lymph node metastasis in papillary thyroid microcarcinoma: construction of a novel population-based predictive model. BMC Endocr Disord. 2022;22(1):269.
31. Tang H, et al. The emerging era of molecular medicine. ACS Nano. 2024;18(45):30911–8.
32. Zhang Q, et al. Screening and validation of lymph node metastasis risk-factor genes in papillary thyroid carcinoma. Front Endocrinol (Lausanne). 2022;13:991906.
33. Kilicarslan S, Hiz-Cicekliyurt MM. Identification of potential biomarkers of papillary thyroid carcinoma. Endocrine 2024.
34. Jiang Y, et al. MicroRNA-142-3P suppresses the progression of papillary thyroid carcinoma by targeting FN1 and inactivating FAK/ERK/PI3K signaling. Cell Signal. 2023;109:110792.
35. Zhang S et al. Identification and analysis of genes associated with papillary thyroid carcinoma by bioinformatics methods. Biosci Rep. 2019;39(4).
36. Sun W, et al. The NEAT1_2/miR-491 axis modulates papillary thyroid cancer invasion and metastasis through TGM2/NFκb/FN1 signaling. Front Oncol. 2021;11:610547.
37. Chen C, Shen Z. FN1 promotes thyroid carcinoma cell proliferation and metastasis by activating the NF-Kb pathway. Protein Pept Lett. 2023;30(1):54–64.
38. Si M, Lang J. The roles of Metallothioneins in carcinogenesis. J Hematol Oncol. 2018;11(1):107.
39. Bizoń A, Jędryczko K, Milnerowicz H. The role of Metallothionein in oncogenesis and cancer treatment. Postepy Hig Med Dosw (Online). 2017;71(0):98–109.
40. Wojtczak B, et al. Metallothionein isoform expression in benign and malignant thyroid lesions. Anticancer Res. 2017;37(9):5179–85.
41. Chen Y, et al. Low Metallothionein 1 M (MT1M) is associated with thyroid cancer cell lines progression. Am J Transl Res. 2019;11(3):1760–70.
42. Laskou A et al. Different forms of TFF3 in the human endocervix, including a complex with IgG Fc binding protein (FCGBP), and further aspects of the Cervico-Vaginal innate immune barrier. Int J Mol Sci 2024;25(4).
43. Lukyanov SA et al. Prediction of the aggressive clinical course of papillary thyroid carcinoma based on fine needle aspiration biopsy molecular testing. Int J Mol Sci 2024;25(13).
44. Ju SH, et al. Transcriptomic characteristics according to tumor size and SUV(max) in papillary thyroid cancer patients. Sci Rep. 2024;14(1):11005.
45. Xin Y, et al. Trefoil factor 3 inhibits thyroid cancer cell progression related to IL-6/JAK/STAT3 signaling pathway. Evid Based Complement Alternat Med. 2021;2021:2130229.
46. Mao J, et al. Risk factors for lymph node metastasis in papillary thyroid carcinoma: A systematic review and Meta-Analysis. Front Endocrinol (Lausanne). 2020;11:265.
47. Gao X, et al. Predictors and a prediction model for central cervical lymph node metastasis in papillary thyroid carcinoma (cN0). Front Endocrinol (Lausanne). 2021;12:789310.
48. Peng Y, et al. Prediction of central lymph node metastasis in cN0 papillary thyroid carcinoma by CT radiomics. Acad Radiol. 2023;30(7):1400–7.
49. Ma T, et al. A clinical and molecular pathology prediction model for central lymph node metastasis in cN0 papillary thyroid microcarcinoma. Front Endocrinol (Lausanne). 2023;14:1075598.
50. Zhu D, et al. Predictive value of ultrasound imaging characteristics and a BRAF V600E nomogram for central lymph node metastasis risk in papillary thyroid microcarcinoma. Altern Ther Health Med. 2023;29(8):139–43.
51. Ren H, et al. Co-existence of BRAF(V600E) and TERT promoter mutations in papillary thyroid carcinoma is associated with tumor aggressiveness, but not with lymph node metastasis. Cancer Manag Res. 2018;10:1005–13.
52. Qiu P, et al. Development of a nomogram for prediction of central lymph node metastasis of papillary thyroid microcarcinoma. BMC Cancer. 2024;24(1):235.
53. Titov SE et al. Could SLC26A7 be a promising marker for preoperative diagnosis of High-Grade papillary thyroid carcinoma?? Diagnostics (Basel). 2024;14(23).

Publisher's note