European Review for Medical and Pharmacological Sciences 2024; 28: 2351-2362

Development and validation of a nomogram model for predicting distant metastasis of aged ≥50 patients with thyroid carcinoma: a SEER database analysis

J.-W. YU, R. PANG, B. LIU, L. ZHANG, L.-Y. KONG

Department of Head and Neck Thyroid, Harbin Medical University Cancer Hospital, Harbin, Heilongjiang, China

Abstract. – OBJECTIVE: This work aimed to construct and validate a model for predicting distant metastasis (DM) in thyroid carcinoma (TC) patients aged≥50.

PATIENTS AND METHODS: The research data were collected from the Surveillance, Epidemiology, and End Results (SEER) program databases via SEER*Stat software (https://seer.cancer.gov/). Logistics regression was used to screen the independent risk factors for TC patients. The nomogram was constructed and validated based on the logistics regression results for predicting DM occurrence in TC patients. Moreover, the characteristic curves (ROC) were used to assess the predictive performance. The decision analysis curve (DCA) and the calibration curve were used to test this nomogram's accuracy and discrimination. Additionally, we analyzed survival and risk scores in TC patients with metastasis using the Kaplan-Meier (KM) method.

RESULTS: A total of 11,166 TC patients were divided into a training set and a validation set. The results showed that topography (T), lymph node metastasis (N), and (grade) G were crucial risk factors for predicting DM. ROC analysis showed that the model had a good discriminative ability both in the training and validation set. The DCA curve showed greater net benefits across a range of DM risks for the nomogram in the training and validation set. Survival analyses showed that the metastasis cases with low-risk scores have shown a poorer prognosis in this study, both in the training and validation set.

CONCLUSIONS: The nomogram model had excellent predictive performance and net benefit for predicting DM of TC patients aged ≥50. The model can help doctors develop treatment plans for their patients.

Key Words:
Thyroid cancer, SEER, Nomogram, Middle-aged and elderly people.

Introduction

Thyroid cancer (TC) is the most frequent endocrine cancer in the human body. According to the Global Cancer Statistics 2020, TC is responsible for 586,000 cases worldwide, ranking in 9th place for incidence. Its mortality rates are much lower. Based on statistical data, the annual global death toll for patients with TC was 44,000 deaths, equating to 8 deaths per 1,000,000 patients with TC[1]. There are four main subtypes containing papillary thyroid carcinoma (PTC), follicular thyroid carcinoma (FTC), medullary thyroid carcinoma (MTC), and anaplastic thyroid cancer (ATC)[2]. The risk factors for thyroid cancer include ionizing radiation, obesity, hormonal exposures, environmental pollution, and so on[3]. In current clinical practice, surgery is the general treatment for thyroid cancer, and is divided into total thyroidectomy and subtotal thyroidectomy. The extent of surgical resection is determined by a variety of factors, such as the patient's age, gender, occupation, whether there is a family history of thyroid cancer, whether there is a history of neck radiotherapy, tumor size, lesion location, metastasis status, whether the patient has other underlying diseases and so on[4]. Some studies[5-7] suggested that total thyroidectomy and unilateral gland lobe (plus isthmus) resection were not different in reducing the recurrence rate and improving the survival rate of differentiated thyroid cancer (DTC) patients without any risk factors.

Lim et al[8] indicated that compared with unilateral lobectomy plus isthmus resection, total thyroidectomy was a better surgical procedure for thyroid cancer patients with lesions tumor size > 1 cm located in the isthmus. Based on the follow-up data of 61,775 DTC patients in the

Corresponding Author: Lingyu Kong, Ph.D; e-mail: 1827@hrbmu.edu.cn

National Cancer Database, Adam et al[9] found no significant difference in overall survival between DTC patients with a size of 1-4 cm after lobectomy and total thyroidectomy after multivariate adjustment. However, the subgroup analysis of tumor size of 2-4 cm showed that the HR value was 1.93, and when the tumor size was > 2 cm and the number of metastatic lymph nodes was > 2, the risk of long-term recurrence after thyroidectomy was significantly increased[10]. Conventional 131I therapy and thyroid stimulating hormone (TSH) suppression therapy were carried out to reduce the probability of recurrence according to the patient's condition after surgery. However, a proportion of patients still develop distant metastases or recurrence, reducing the long-term survival of patients. Its mortality rate and treatment burden are improved due to distant metastases.

For distant metastases of thyroid cancer, the most common ones are lung metastases and bone metastases, of which 85.6% of lung metastases in differentiated thyroid cancer[2]. Therefore, it is very important to predict the recurrence and metastasis of patients according to risk factors. Nixon et al[11] indicated that about half of the patients with well-differentiated thyroid cancer presenting with distant metastases die of the disease within 5 years of initial diagnosis despite thyroid surgery and RAI and lung distant metastases play an important role in poor prognosis in those patients aged>45. Wang et al[12] constructed a nomogram that had good discrimination and was based on surgery, tumor size, topography (T), lymph node metastasis (N), metastasis (M), age, and other clinical features to predict cancer-specific survival in older patients with PTC. Park et al[13] predict the survival prognosis of PTC patients by using machine learning methods.

The Surveillance, Epidemiology, and End Results (SEER) is a National Cancer Institute database that captures the incidence, mortality, and disease status of millions of patients with malignancies in selected states and counties in the USA. The SEER database is designed to reduce the burden of cancer in the US population, the tumor information in the database is standardized through the SEER*Stat software (https://seer.cancer.gov/) and is regularly updated and published. Oncology researchers worldwide have requested access to some of the data, which provides an excellent source for clinical researchers who lack clinical research data. In addition, the large sample size and statistical power of the SEER database allow studies based on the SEER database to have high clinical reference values. There were many retrospective studies based on the SEER database. For instance, Zhao et al[14] researched the relationship between clinicopathological characteristics and survival outcomes in patients with Paget's disease using SEER databases. Safi et al[15] examined the impact of adverse cardiotoxic reactions from immunotherapy on patients' overall survival (OS) based on data from the SEER database of patients with advanced non-small cell lung cancer.

In this study, our aim was to construct and validate the model for predicting distant metastasis (DM) in patients who developed TC and aged more than 50 based on the SEER database. This work could provide new clinical ideas for treatment.

Patients and Methods

Data Source

We collected clinical and demographic data on patients from the SEER database (https://seer.cancer.gov/). It is an authoritative source of cancer statistics in the USA. SEER is the Surveillance, Epidemiology, and End Results program that provides information on cancer statistics to reduce the cancer burden in the US population. SEER is supported by the Surveillance Research Program (SRP) in NCI's Division of Cancer Control and Population Sciences (DCCPS). It covers 34.6% of the US population and collects case information from 18 population-based cancer registries[16]. It is a public database, and the researcher can access the data after the application.

Data Extraction

The screening criteria contained (1) primary site in the thyroid gland, (2) age greater than or equal to 50 years. We collected patients' information on age, ethnicity, T, N, M, subtype, vital status, stage, sex, grade (G), cancer-specific survival (CSS), and survival months. The population consists of the following age groups: 50-54 years, 75-79 years, 65-69 years, 55-59 years, 80-84 years, 70-74 years, 60-64 years, and 85+ years. The primary site of tumors in the population is the thyroid gland. The patients contain a variety of thyroid cancer subtypes for this study. The histopathological subtype was based on Third Revision Histopathological codes (ICD-O-3). The ICD-O-3 were listed in the **Supplementary Table I**.

Statistical Analysis

In this study, a nomogram was used to build models to predict the risk of metastasis events in

thyroid cancer patients. SPSS version 23.0 (IBM Corp., Armonk, NY, USA) software was used for data collection and analysis. The incidence rates of metastasis and no metastasis were calculated by dividing the total number of patients both in the training set and validation set. Categorical variables were expressed as a percentage (%) and compared using the Chi-square test.

Quantitative data conforming to normal distribution and homogeneity of variance were expressed by mean±standard deviation and the *t*-test for comparison between the two groups (M1 and M0). Variables with statistically significant differences between groups were included in the univariate and multivariate logistic regression analysis to determine the independent influencing factors for the occurrence of DM, with a test level of 0.05. The results were presented as adjusted odds ratios (OR) with 95% confidence intervals (CI).

Then, the risk score of patients was calculated based on risk factors and the corresponding logistics regression coefficient as follows: Risk score=β1× the value of risk factors 1+β2×the value of factors 1+.....+βn× the value of risk factors n, where β is the multivariate regression coefficient of the corresponding risk factor, and the value of risk factors is the value of the corresponding risk factors.

The Kaplan-Meier (KM) method was used to analyze the relationship between risk score and the survival of patients with M1 and examined differences in survival between the patients of high-risk and low-risk groups by Log-rank tests. Otherwise, the nomogram was constructed based on key risk factors using R software. Then, the three methods containing ROC, DCA, and calibration curves were used to evaluate the predictive performance of the model. Concretely, the ROC package of R software was used to plot the ROC curve, which is a tool for assessing the predictive performance of the model. A decision analysis curve (DCA), which is a new evaluation algorithm, was used to plot for calculating the net benefit of the prediction model. In addition, the study design flow chart is shown in the **Supplementary Figure 1.**

Results

Patient Characteristics

A total of 11,166 patients with thyroid cancer, including 3,274 (29.32%) males and 7,892 (70.68%) females were collected from the SEER databases. 2,698 (24.16%) of them were within the age range of 50-54 years. 687 (6.15%) of them were within the age range of 75-79 years. 1,677(15.02%) were within the age range of 65-69 years. 2,424 (21.71%) were within the age range of 55-59 years. 348 (3.12%) were within the age range of 80-84 years. 1,052 (9.42%) were within the age range of 70-74 years. 2,022 (18.11%) were within the age range of 60-64 years. 258 (2.31%) aged more than 85. The white patients, the black patients, and patients of other ethnicities were 9,091 (81.42%), 1,375 (12.31%), and 608 (5.45%), respectively. In this study, most patients were in stage I (5,933; 53.13%) and stage III (2,357; 21.11%), followed by stage IV (1,550; 13.88%), and the smallest numbers were in stage II (1326; 11.88%). All patients were assessed for neoplasm disease lymph node stage, of whom 8,929 (79.97%) were N0 and 2,237 (20.03%) were N1. The number of patients at stages T0, T1, T2, T3, and T4 was 16 (0.14%), 6,706 (60.06%), 1,579 (14.30%), 2,213 (19.82%), and 634 (5.68%), respectively. Of those patients who were followed up, 9,536 (85.4%) survived, and 1,630 (14.6%) died. The number of patients at M0 and M1 was 10,850 (97.170%) and 316 (2.83%), respectively. The number of patients alive and those who died of other causes was 10,590 (94.84%), and those who died due to thyroid cancer were 561 (5.02%). The number of patients at G I, G II, G III, and G IV was 1,931, 396, 163, and 223, respectively. The number of patients at M0 and M1 was 10,850 (97.170%) and 316 (2.83%), respectively.

Those patients were divided into training sets containing 7,443 patients and validation sets containing 3,723 patients. The propensity score matching analysis was conducted to exclude differences in patient age and gender between the two sets. Table I shows that some clinic information had no significant difference between the training set and the validation set.

Univariate Analysis

The results of the univariate analysis showed that those variables were different between the M0 group and M1 group, and they were statistically significant both in training sets and validation sets. Those variables contained age, ethnicity, T, N, CSS status, overall survival (OS) status, grade (G), sex, subtype, and survival months. The differences were statistically significant ($p<0.05$) when comparing the groups. The results of the analysis of variables are given in Table II.

Table I. Baseline clinical characteristics of patients with thyroid cancer.

Variables	Set (n=11166)		χ^2/F	p
	Validation set (n=3,723)	Training set (n=7,443)		
Age			**0.005**	**1.000**
50-54 years	900 (24.2%)	1,798 (24.2%)		
75-79 years	229 (6.2%)	458 (6.2%)		
65-69 years	560 (15.0%)	1,117 (15.0%)		
55-59 years	808 (21.7%)	1,616 (21.7%)		
80-84 years	116 (3.1%)	232 (3.1%)		
70-74 years	350 (9.4%)	702 (9.4%)		
60-64 years	674 (18.1%)	1,348 (18.1%)		
85+ years	86 (2.3%)	172 (2.3%)		
Ethnicity			**36.205**	**0.030**
White	2,956 (79.4%)	6,135 (82.4%)		
Other	511 (13.7%)	864 (11.6%)		
Black	203 (5.5%)	405 (5.4%)		
Unknown	53 (1.4%)	39 (0.5%)		
T			**5.553**	**0.235**
T2	506 (13.6%)	1,091 (14.7%)		
T4	223 (6.0%)	411 (5.5%)		
T1	2,221 (59.7%)	4,485 (60.3%)		
T3	769 (20.7%)	1,444 (19.4%)		
T0	4 (0.1%)	12 (0.2%)		
N			**7.646**	**0.006**
N0	2,922(78.5%)	6,007 (80.7%)		
N1	801 (21.5%)	1,436 (19.3%)		
Subtype			**0.783**	**0.941**
PTC	3,287 (88.3%)	6,544 (87.9%)		
FTC	166 (4.5%)	357 (4.8%)		
MTC	69 (1.9%)	145 (1.9%)		
ATC	76 (2.0%)	149 (2.0%)		
Others	125 (3.4%)	248 (3.3%)		
Vital status			**0.065**	**0.799**
Alive	3,184 (85.5%)	6,352 (85.3%)		
Dead	539 (14.5%)	1,091 (14.7%)		
Stage			**10.799**	**0.013**
III	821 (22.1%)	1,536 (20.6%)		
IV	554 (14.9%)	996 (13.4%)		
I	1,936 (52.0%)	3,997 (53.7%)		
II	412 (11.1%)	914 (12.3%)		
Sex			**0.013**	**0.908**
Male	1,089 (29.3%)	2,185 (29.4%)		
Female	2,634 (70.7%)	5,258 (70.6%)		
M			**0.315**	**0.574**
M0	3,613 (97.0%)	7,237 (97.2%)		
M1	110 (3.0%)	206 (2.8%)		
Grade			**108.119**	**0.001**
I	837 (22.5%)	1,094 (14.7%)		
III	53 (1.4%)	110 (1.5%)		
II	109 (2.9%)	287 (3.9%)		
IV	74 (2.0%)	149 (2.0%)		
Unknown	2,650 (71.2%)	5,803 (78.0%)		
CSS			**5.853**	**0.055**
Alive or dead of other cause	3,516 (94.4%)	7,074 (95.0%)		
Dead (attributable to this cancer dx)	198 (5.3%)	363 (4.9%)		
Unknown	9 (0.2%)	6 (0.1%)		
Survival months	50.561±30.707	50.791±30.94	0.137	0.711

CSS: cancer-specific survival.

Table II. Comparison of basic variables between the M0 group and the M1 group.

	Training set (N=7443)				Validation set (N=3723)			
Variables	M0 (N=7237)	M1 (N=206)	χ^2/F	p	M0 (N=3613)	M1 (N=110)	χ^2/F	p
Age			**167.94**	**<0.001**			**91.385**	**<0.001**
85+ years	146 (2.0%)	26 (12.6%)			896 (24.8%)	4 (3.6%)		
50-54 years	1,777 (24.6%)	21 (10.2%)			216 (6.0%)	13 (11.8%)		
55-59 years	1,593 (22.0%)	23 (11.2%)			545 (15.1%)	15 (13.6%)		
80-84 years	219 (3.0%)	13 (6.3%)			795 (22.0%)	13 (11.8%)		
70-74 years	684 (9.5%)	18 (8.7%)			103 (2.9%)	13 (11.8%)		
75-79 years	425 (5.9%)	33 (16.0%)			327 (9.1%)	23 (20.9%)		
65-69 years	1,079 (14.9%)	38 (18.4%)			654 (18.1%)	20 (18.2%)		
60-64 years	1,314 (18.2%)	34 (16.5%)			77 (2.1%)	9 (8.2%)		
T			**970.104**	**<0.001**			**509.237**	**<0.001**
T3	1,388 (19.2%)	56 (27.2%)			495 (13.7%)	11 (10.0%)		
T2	1,073 (14.8%)	18 (8.7%)			164 (4.5%)	59 (53.6%)		
T1	4,464 (61.7%)	21 (10.2%)			2,210 (61.2%)	11 (10.0%)		
T4	302 (4.2%)	109 (52.9%)			742 (20.5%)	27 (24.5%)		
T0	10 (0.1%)	2 (1.0%)			2 (0.1%)	2 (1.8%)		
N			**176.794**	**<0.001**			**129.6**	**<0.001**
N1	1,322 (18.3%)	114 (55.3%)			729 (20.2%)	72 (65.5%)		
N0	5,915 (81.7%)	92 (44.7%)			2,884 (79.8%)	38 (34.5%)		
Cause-specific death			**1502.669**	**<0.001**			**741.932**	**<0.001**
Dead	235 (3.2%)	128 (62.1%)			3,475 (96.2%)	41 (37.3%)		
Alive or dead of other cause	6,997 (96.7%)	77 (37.4%)			129 (3.6%)	69 (62.7%)		
Unknown	5 (0.1%)	1 (0.5%)			9 (0.2%)	0 (0.0%)		
Grade			**965.714**	**<0.001**			**529.12**	**<0.001**
III	90 (1.2%)	20 (9.7%)			820 (22.7%)	17 (15.5%)		
Unknown	5,695 (78.7%)	108 (52.4%)			2,607 (72.2%)	43 (39.1%)		
I	1,085 (15.0%)	9 (4.4%)			105 (2.9%)	4 (3.6%)		
IV	87 (1.2%)	62 (30.1%)			43 (1.2%)	31 (28.2%)		
II	280 (3.9%)	7 (3.4%)			38 (1.1%)	15 (13.6%)		
Ethnicity			**11.871**	**0.008**			**18.502**	**<0.001**
White	5,981 (82.6%)	154 (74.8%)			2,884 (79.8%)	72 (65.5%)		
Other	826 (11.4%)	38 (18.4%)			483 (13.4%)	28 (25.5%)		
Black	391 (5.4%)	14 (6.8%)			193 (5.3%)	10 (9.1%)		
Unknown	39 (0.5%)	0 (0.0%)			53 (1.5%)	0 (0.0%)		
Stage			**1,371.371**	**<0.001**			**648.381**	**<0.001**
IV	790 (10.9%)	206 (100.0%)			3,240 (89.7%)	47 (42.7%)		
I	3,997 (55.2%)	0 (0.0%)			156 (4.3%)	10 (9.1%)		
II	914 (12.6%)	0 (0.0%)			61 (1.7%)	8 (7.3%)		
III	1,536 (21.2%)	0 (0.0%)			47 (1.3%)	29 (26.4%)		
Vital status			**572.834**	**<0.001**			**310.62**	**<0.001**
Dead	941 (13.0%)	150 (72.8%)			459 (12.7%)	80 (72.7%)		
Alive	6,296 (87.0%)	56 (27.2%)			3,154 (87.3%)	30 (27.3%)		
Sex			**43.536**	**<0.001**			**6.329**	**0.012**
Female	5,155 (71.2%)	103 (50.0%)			2,569 (71.1%)	66 (60.0%)		
Male	2,082 (28.8%)	103 (50.0%)			1,046 (29.0%)	44 (40.0%)		
Subtype			**1007.83**	**<0.001**			**420.802**	**<0.001**
PTC	6,451 (89.1%)	93 (45.1%)			3,225 (89.3%)	47 (42.7%)		
FTC	337 (4.7%)	20 (9.7%)			156 (4.3%)	10 (9.1%)		
MTC	133 (1.8%)	12 (5.8%)			61 (1.7%)	9 (8.2%)		
ATC	84 (1.2%)	65 (31.6%)			47 (1.3%)	29 (26.4%)		
Others	232 (3.2%)	16 (7.8%)			124 (3.4%)	5 (4.5%)	6.329	0.012
Survival months	56.938±34.14	49.15±33.15	10.438	0.001	50.469±30.63	53.564±33.08	1.084	0.298

CSS: cancer-specific survival.

Test of Logistic Regression Model

The above statistically significant variables were included in the logistic regression to explore independent factors affecting metastase events in thyroid cancer patients. The results of univariable logistic regression showed that age, ethnicity, T, N, stage, sex, subtype, and G were independent factors affecting the occurrence of metastases. Those variables were included in multivariate logistic regression for calculation. The T, N, and G were independent factors affecting the occurrence of metastases (Table III).

Construction and Assessment of the Nomogram

According to the key risk factors screened by logistic regression, the nomogram containing T, N, and G was established and shown in Figure 1. The model showed a great discriminative ability by ROC analysis both in the training set (Figure 2A, AUC: 0.914; 95%CI: 0.881-0.946) and validation set (Figure 2B, AUC: 0.889; 95%CI: 0.847-0.932). The calibration plots of the nomogram showed consistency between the actual observations and the predicted DM, both in the training (Figure 2C) and validation cohorts (Figure 2D). The DCA curve showed greater net benefits across a range of DM risks for the nomogram in the training set (Figure 3A) and validation (Figure 3B).

The Survival Analyses of TC Patients with Metastasis

We analyzed the survival and risk scores in TC patients with metastasis by the Kaplan-Meier method and compared them using the log-rank test on OS. According to the KM plots, the metastasis cases with low-risk scores have shown a poorer prognosis in this study both in the training set (p<0.001, Figure 4A) and validation set (p<0.001, Figure 4B). The correlation between the survival and risk scores in TC patients with metastasis was also investigated by using Kaplan-Meier analysis with the log-rank test on CSS both in the training set and validation set. The results showed that the metastasis cases with low-risk scores have shown a poorer prognosis both in the training set (p<0.001, Figure 4C) and validation set (p<0.001, Figure 4D).

Discussion

The incidence of thyroid cancer was the seventh highest cancer according to the Global Cancer Survey 2020[1,17]. Patients with thyroid cancer are

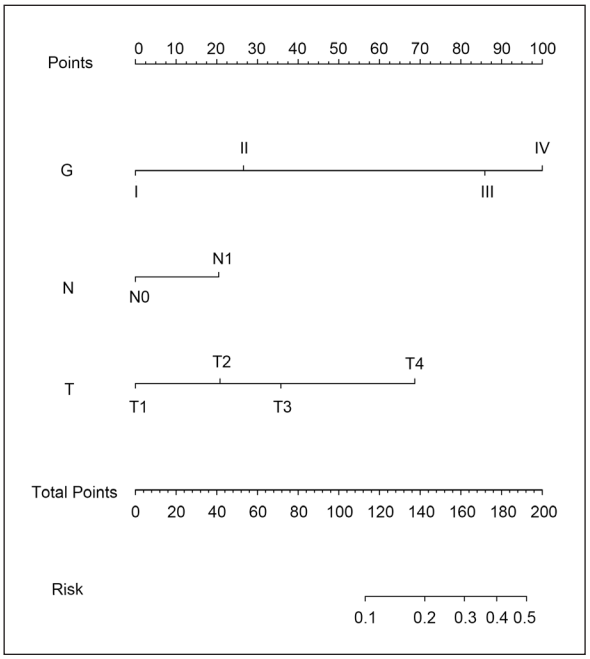

Figure 1. The nomogram for predicting the occurrence of metastases in patients with thyroid cancer.

Table III. Univariate and multivariate logistic regression analysis of the risk factors in patients with M0.

Variables	Univariate logistic regression analysis			Multivariate logistic regression analysis		
	B	p	OR (95% CI)	B	p	OR (95% CI)
Age	0.351	<0.001	1.421 (1.323-1.519)	0.054	0.386	1.055 (0.935-1.191)
Ethnicity	0.301	<0.001	1.351 (1.133-1.611)	0.312	0.065	1.366 (0.981-1.903)
T	1.468	0.002	4.342 (3.684-5.118)	0.539	0.009	1.713 (1.142-2.571)
N	1.713	<0.001	5.544 (4.186-7.344)	0.613	0.019	1.846 (1.107-3.079)
Stage	17.036	0.950	/			
Sex	-0.907	<0.001	0.404 (0.306-0.533)	-0.373	0.142	0.689 (0.419-1.134)
Subtype	0.695	<0.001	2.003 (1.842-2.179)	0.174	0.113	1.189 (0.959-1.476)
G	1.521	<0.001	4.577 (3.669-5.711)	0.838	<0.001	2.311 (1.585-3.371)

OR: odds ratio; CI: Confidence interval.

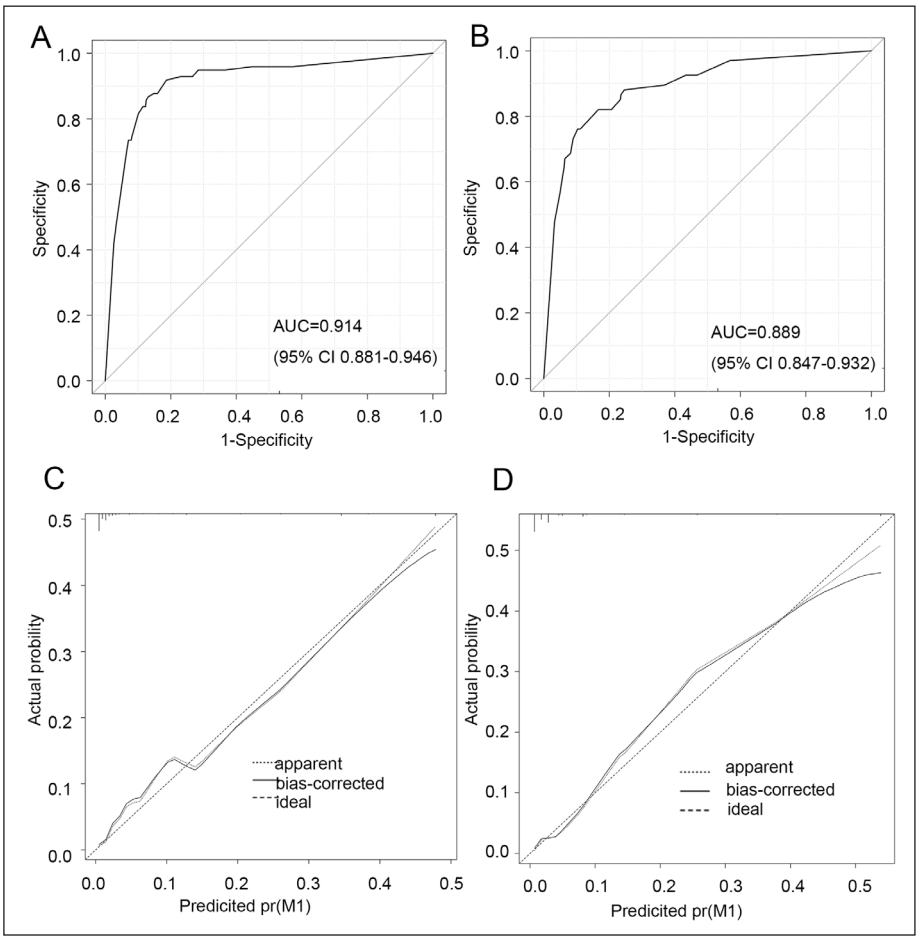

Figure 2. AUC for predicting the occurrence of metastases in the training set (**A**) and validation set (**B**). Calibration curve of the nomogram in the training set (**C**) and validation set (**D**).

very unlikely to develop distant metastases (including lung, bone, liver, and brain metastases), but they have a severely reduced survival time when they do[18-20]. Metastasis of TC was crucial for survival in older patients due to a positive correlation with age[12,21]. Scholars[22] have shown that the frequency of metastasis in elderly patients was higher than in young patients, and they often have a poorer prognosis than young patients. However, studies on DM in elderly patients with TC were rare. SEER currently collects and publishes cancer incidence and survival data from population-based cancer registries covering approximately 48.0 percent of the US population. There have been many discoveries based on the SEER database. For example, radical surgery can achieve better results for non-metastatic bladder cancer than radiotherapy23. Compared with other interventional radiology methods for the treatment of non-HCC liver cancer, heat-radio-frequency ablation may improve the survival rate[24]. Radiotherapy was associated with poor prognosis in patients with ovarian cancer regardless of pathology or stage[25]. Zhang et al[26] indicated that the disease of young patients with metastatic breast cancer was more aggressive, but the prognosis was better, and the prognosis of the young patients was better than the prognosis of the older group. Some prediction models were constructed based on the SEER database. For instance, Wang et al[27] constructed a prognostic nomogram for cervical cancer patients. Yao et al[28] included age, tumor location, histological type, T stage, carcinoembryonic antigen level, tumor deposits, Log odds of metastatic lymph nodes, and extraperitoneal metastasis to establish a risk model for peritoneal metastasis of rectal cancer.

Hence, we constructed the model to predict DM for them based on the SEER databases. Specifically, we collected and analyzed the clinical characteristics and demographic information of TC patients. Then, we established and validated the nomogram based on key metastasis risk factors for TC patients aged ≥50. We also used three methods for assessing the predictive performance of the nomogram model. The results showed that

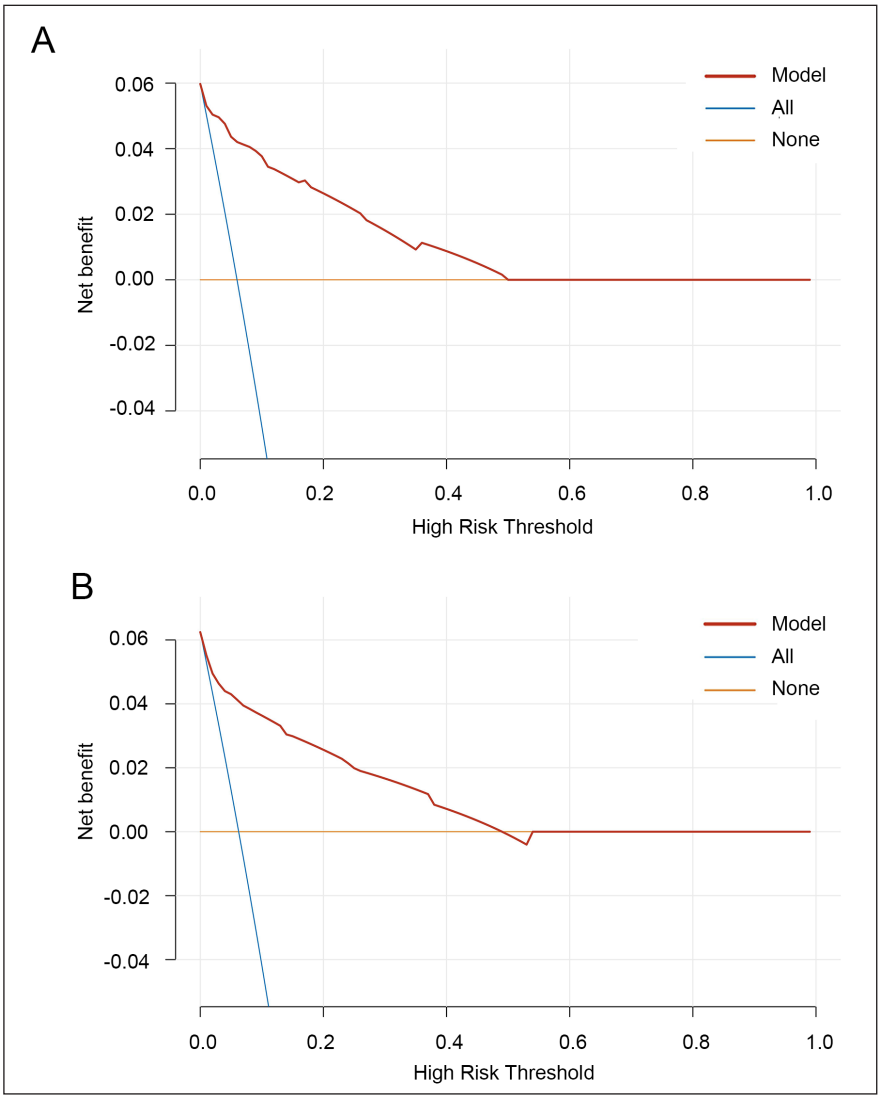

Figure 3. DCA of the nomogram in the training set (**A**) and validation set (**B**).

T, N, and G, were key risk factors for DM in elderly TC patients. The ROC curve showed that this model had a good predictive performance. The calibration plots of the nomogram showed good agreement between the actual observations and the predicted DM. Moreover, the DCA curve showed greater net benefits across a range of DM risks for the nomogram.

There was a set of studies about risk factor predicting models for patients with TC. Vuong et al[29] indicated that many clinical parameters affect DM in patients with TC, for instance, age, sex, gender, tumor size, vascular invasion, and lymph node metastasis via a meta-analysis of relevant studies. Of those factors, age, vascular invasion, extrathyroidal extension, and lymph node metastasis were important risk factors for DM in both the PTC subtype and FTC subtype. In this study, lymph node metastasis was one of the significant factors for DM in patients with TC. This was also consistent with prior studies by Qiao et al[30]. Their research indicated that age, gender, ethnicity, marital status, histological type, capsular invasion, and number of lymph node metastases were key factors of DM and constructed a predicted DM model including those factors for patients with TC based on the SEER database. Lymph node metastasis is the first step of DM in a variety of human cancers. Lymphatic metastasis-competent cancer cell lines can upregulate the expression of interferon-inducible genes (ISGs), such as MHC-I and PD-L1, which help tumor cells evade killing by NK cells and T cells, and thus successfully colonizing lymph nodes. Then, lymph node tumor cells

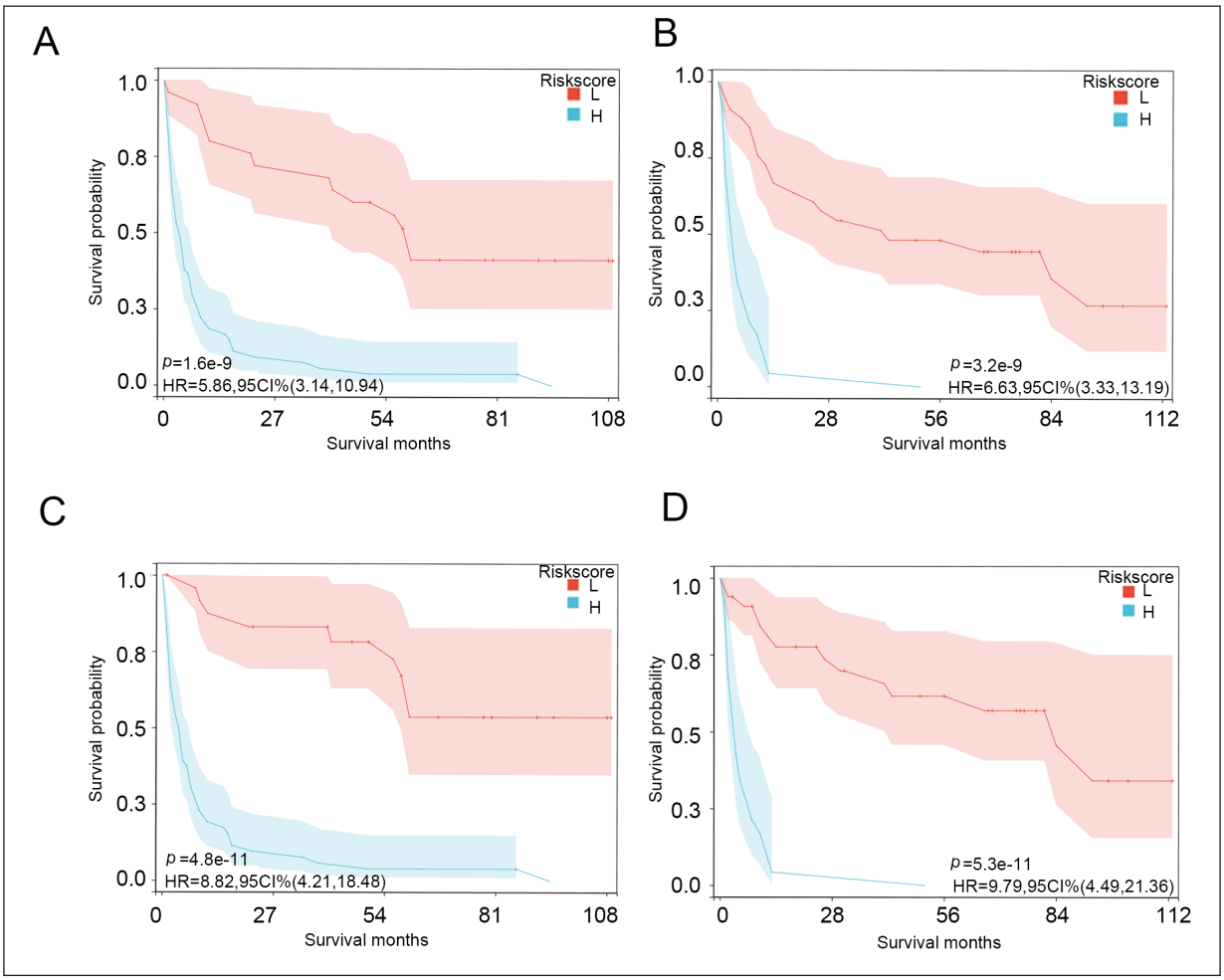

Figure 4. Kaplan-Meier curves for the value of risk-score in metastasis patients for the OS in the training set (**A**) and validation set (**B**). Kaplan-Meier curves for the value of risk-score in metastasis patients for the CSS in the training set (**C**) and validation set (**D**).

promote Treg differentiation by increasing TGF-β, which generates a tumor-specific immune tolerance microenvironment that subsequently facilitates distant tumor colonization[31-34]. Nonetheless, age was not the crucial factor for the prediction of DM because the participants were more than 49 years old. The age of 45 years serves as a critical threshold for tumor metastasis in patients with TC, as indicated by numerous studies in the literature[35-37]. To the best of our knowledge, it is the first time that G has been included in a predictive model of tumor metastasis in patients with TC. The tumor G score means the tumor differentiation grade of heterogeneity. The higher the degree of differentiation, the closer it is to normal cells. The low grade of tumor differentiation and large size of the tumor reflects the malignancy and growth time of the tumor respectively, both of which are related to the tendency of the tumor to invade and metastasize[38].

The results showed that the metastasis cases with low-risk scores have a poorer prognosis in this work. This result is consistent with previous findings[12]. Older patients are more likely to develop complications in treatment due to poor physical exercise. In addition, TSH inhibition in older patients with thyroid cancer is associated with an increased risk of atrial fibrillation, cardiovascular disease, and CSS[12,39,40].

One major merit of our study was that the SEER database provides a very large number of samples to construct the DM prediction model for TC patients aged ≥50 years.

Limitations

However, like previous retrospective case-control studies, causal inference is limited. Therefore, our results and conclusions should only be used to assess the risk of distance metastases in

TC patients aged ≥50 and should be validated by rigorously designed follow-up cohort studies. Clinical features and personal details were limited. Future studies with a wider range of variables are needed to validate our findings further. This study solely concentrated on the demographics of the United States, potentially restricting its applicability to other populations. Although our column line graphs were validated by the internal dataset, future prospective worldwide studies still need to externally validate our results.

Conclusions

To better predict the occurrence of DM in TC patients aged ≥50, we have established and validated a nomogram. The proposed nomogram contained three risk factors: T, N, and G. In addition, patients with DM and lower risk scores had higher survival probability. Our nomogram can predict the DM of patients with TC and provide a reference value for doctors to formulate patients' treatment plans. Compared with the nomogram established in other articles already reported, this nomogram showed greater predictive performance and was simpler.

Data Availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Ethics Approval and Informed Consent

The Ethics Committee of Harbin Medical University Cancer Hospital agreed to submit the study for review and has waived the need for ethical approval, as the data in this study are from public databases.

Conflict of Interest

The authors declare that they have no conflict of interest to disclose.

Funding

No funding was used in this study.

Authors' Contributions

J.-W. Yu and R. Pang contributed to the conception and design; J.-W. Yu and B. Liu contributed to the collection and assembly of data; J.-W. Yu, L. Zhang and L.-Y. Kong analyzed and interpreted the data; All authors wrote and approved the final manuscript.

ORCID ID

Lingyu Kong: 0009-0004-2159-1142.

References

1) Sung H, Ferlay J, Siegel RL., Laversanne M, Soerjomataram I, Jemal A, Bray F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin 2021; 71: 209-249.
2) Cabanillas ME, McFadden DG, Durante C. Thyroid cancer. Lancet 2016; 388: 2783-2795.
3) Lortet-Tieulent J, Franceschi S, Dal Maso L, Vaccarella S. Thyroid cancer "epidemic" also occurs in low- and middle-income countries. Int J Cancer 2019; 144: 2082-2087.
4) Hauch A, Al-Qurayshi Z, Randolph G, Kandil E. Total thyroidectomy is associated with increased risk of complications for low- and high-volume surgeons. Ann Surg Oncol 2014; 21: 3844-3852.
5) Lee J, Park JH, Lee CR, Chung WY, Park CS. Long-term outcomes of total thyroidectomy versus thyroid lobectomy for papillary thyroid microcarcinoma: comparative analysis after propensity score matching. Thyroid 2013; 23: 1408-1415.
6) Zhang C, Li Y, Li J, Chen X. Total thyroidectomy versus lobectomy for papillary thyroid cancer: A systematic review and meta-analysis. Medicine (Baltimore) 2020; 99: e19073.
7) Kim SK, Park I, Woo JW, Lee JH, Choe JH, Kim JH, Kim JS. Total thyroidectomy versus lobectomy in conventional papillary thyroid microcarcinoma: Analysis of 8,676 patients at a single institution. Surgery 2017; 161: 485-492.
8) Lim ST, Jeon YW, Suh YJ. Correlation Between Surgical Extent and Prognosis in Node-Negative, Early-Stage Papillary Thyroid Carcinoma Originating in the Isthmus. World J Surg 2016; 40: 344-349.
9) Adam MA, Pura J, Gu L, Dinan MA, Tyler DS, Reed SD, Scheri R, Roman SA, Sosa JA. Extent of surgery for papillary thyroid cancer is not associated with survival: an analysis of 61,775 patients. Ann Surg 2014; 260: 601-605.
10) Hwangbo Y, Kim JM, Park YJ, Lee EK, Lee YJ, Park DJ, Choi YS, Lee KD, Sohn SY, Kim SW, Chung JH, Lim DJ, Kim MH, Kim MJ, Jo YS, Shong MH, Koong SS, Hahm JR, Jung JH, Yi KH. Long-Term Recurrence of Small Papillary Thyroid Cancer and Its Risk Factors in a Korean Multicenter Study. J Clin Endocrinol Metab 2017; 102: 625-633.
11) Nixon IJ, Whitcher MM, Palmer FL, Tuttle RM, Shaha AR, Shah JP, Patel SG, Ganly I. The impact of distant metastases at presentation on prognosis in patients with differentiated carcinoma of the thyroid gland. Thyroid 2012; 22: 884-889.

12) Wang J, Zhanghuang C, Jin L, Zhang Z, Tan X, Mi T, Liu J, Li M, Wu X, Tian X, He D. Development and validation of a nomogram to predict cancer-specific survival in elderly patients with papillary thyroid carcinoma: a population-based study. BMC Geriatr 2022; 22: 736.
13) Park KS, Kim SH, Oh JH, Kim SY. Highly accurate diagnosis of papillary thyroid carcinomas based on personalized pathways coupled with machine learning. Brief Bioinform 2021; 22: bbaa336.
14) Zhao Y, Sun HF, Chen MT, Gao SP, Li LD, Jiang HL, Jin W. Clinicopathological characteristics and survival outcomes in Paget disease: a SEER population-based study. Cancer Med 2018; 7: 2307-2318.
15) Safi M, Kanesvaran R, Alradhi M, Al-Danakh A, Ping F, Al-Sabai N, Shan X, Liu J. Overall Survival in Heart Disease-Related Death in Non-Small Cell Lung Cancer Patients: Nonimmunotherapy Versus Immunotherapy Era: Population-Based Study. Front Oncol 2020; 22: 572380.
16) Shukla N, Osazuwa-Peters N, Megwalu UC. Association Between Age and Nodal Metastasis in Papillary Thyroid Carcinoma. Otolaryngol Head Neck Surg 2021; 165: 43-49.
17) Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2017. CA Cancer J Clin 2017; 67: 7-30.
18) Sugino K, Nagahama M, Kitagawa W, Ohkuwa K, Uruno T, Matsuzu K, Suzuki A, Tomoda C, Hames KY, Akaishi J, Masaki C, Ito K. Distant Metastasis in Pediatric and Adolescent Differentiated Thyroid Cancer: Clinical Outcomes and Risk Factor Analyses. J Clin Endocrinol Metab 2020; 105: dgaa545.
19) Tong Y, Huang Z, Hu C, Chi C, Lv M, Li P, Zhao C, Song Y. Independent risk factors evaluation for overall survival and cancer-specific survival in thyroid cancer patients with bone metastasis: A study for construction and validation of the predictive nomogram. Medicine (Baltimore) 2020; 99: e21802.
20) Li C, Wu Q, Sun S. Radioactive Iodine Therapy in Patients With Thyroid Carcinoma With Distant Metastases: A SEER-Based Study. Cancer Control 2020; 27: 1073274820914661.
21) Samra B, Jabbour E, Ravandi F, Kantarjian H, Short NJ. Evolving therapy of adult acute lymphoblastic leukemia: state-of-the-art treatment and future directions. J Hematol Oncol 2020; 13: 70.
22) Vini L, Hyer SL, Marshall J, A'Hern R, Harmer C. Long-term results in elderly patients with differentiated thyroid carcinoma. Cancer 2003; 97: 2736-2742.
23) Abdel-Rahman O. Squamous Cell Carcinoma of the Bladder: A SEER Database Analysis. Clin Genitourin Cancer 2017; 15: e463-e468.
24) Legenza A, Tupper C, Burdyny M, Peter T. Non-hepatocellular carcinoma management: A SEER database analysis. Journal of Clinical Oncology 2022; 40: e16108-e16108.
25) Yu L, Gong H, Li Q, Ren H, Wang Y, He H, Li T, Song Q. Survival Analysis of Radiation Therapy in Ovarian Cancer: A SEER Database Analysis. J Oncol 2021; 2021: 8849039.
26) Zhang W, Wu S, Liu J, Zhang X, Ma X, Yang C, Cao M, Zhang S, Liu Y. Metastasis patterns and prognosis in young breast cancer patients: A SEER database analysis. Front Oncol 2022; 12: 872862.
27) Wang C, Yang C, Wang W, Xia B, Li K, Sun F, Hou Y. A Prognostic Nomogram for Cervical Cancer after Surgery from SEER Database. J Cancer 2018; 9: 3923-3928.
28) Yao L, Shao H, Zhang X, Huang X. A novel risk model for predicting peritoneal metastasis in colorectal cancer based on the SEER database. J Cancer Res Clin Oncol 2023; 149: 15989-16000.
29) Vuong HG, Duong UNP, Pham TQ, Tran HM, Oishi N, Mochizuki K, Nakazawa T, Hassell L, Katoh R, Kondo T. Clinicopathological Risk Factors for Distant Metastasis in Differentiated Thyroid Carcinoma: A Meta-analysis. World J Surg 2018; 42: 1005-1017.
30) Qiao L, Li H, Wang Z, Sun H, Feng G, Yin D. Machine learning based on SEER database to predict distant metastasis of thyroid cancer. Endocrine 2023. doi: 10.1007/s12020-023-03657-4. Online ahead of print.
31) Reticker-Flynn NE, Zhang W, Belk JA, Basto PA, Escalante NK, Pilarowski GOW, Bejnood A, Martins MM, Kenkel JA, Linde IL, Bagchi S, Yuan R, Chang S, Spitzer MH, Carmi Y, Cheng J, Tolentino LL, Choi O, Wu N, Kong CS, Gentles AJ, Sunwoo JB, Satpathy AT, Plevritis SK, Engleman EG. Lymph node colonization induces tumor-immune tolerance to promote distant metastasis. Cell 2022; 185: 1924-1942 e23.
32) Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. Nat Rev Cancer 2012; 12: 252-264.
33) Benci JL, Xu B, Qiu Y, Wu TJ, Dada H, Twyman-Saint Victor C, Cucolo L, Lee DSM, Pauken KE, Huang AC, Gangadhar TC, Amaravadi RK, Schuchter LM, Feldman MD, Ishwaran H, Vonderheide RH, Maity A, Wherry EJ, Minn AJ. Tumor Interferon Signaling Regulates a Multigenic Resistance Program to Immune Checkpoint Blockade. Cell 2016; 167: 1540-1554 e12.
34) Sakaguchi S, Mikami N, Wing JB, Tanaka A, Ichiyama K, Ohkura N. Regulatory T Cells and Human Disease. Annu Rev Immunol 2020; 38: 541-566.
35) Sugino K, Kure Y, Iwasaki H, Ozaki O, Mimura T, Matsumoto A, Ito K. Metastases to the regional lymph nodes, lymph node recurrence, and distant metastases in nonadvanced papillary thyroid carcinoma. Surg Today 1995; 25: 324-328.
36) Goffredo P, Sosa JA, Roman SA. Differentiated thyroid cancer presenting with distant metastases: a population analysis over two decades. World J Surg 2013; 37: 1599-1605.
37) Feinmesser M, Stern S, Mechlis-Frish S, Lubin E, Feinmesser R, Kristt D. HLA-DR expression and lymphocytic infiltration in metastatic and non-metastatic papillary carcinoma of the thyroid. Eur J Surg Oncol 1996; 22: 494-501.

38) Imai Y, Ichinose M. Risk stratification for predicting postoperative recurrence/metastasis of colorectal cancer by grade of venous invasion coupled with histological subtype. BMC Gastroenterol 2022; 22: 79.
39) Flynn RW, Bonellie SR, Jung RT, MacDonald TM, Morris AD, Leese GP. Serum thyroid-stimulating hormone concentration and morbidity from cardiovascular disease and fractures in patients on long-term thyroxine therapy. J Clin Endocrinol Metab 2010; 95: 186-193.
40) Tuttle RM, Ahuja S, Avram AM, Bernet VJ, Bourguet P, Daniels GH, Dillehay G, Draganescu C, Flux G, Fuhrer D, Giovanella L, Greenspan B, Luster M, Muylle K, Smit JWA, Van Nostrand D, Verburg FA, Hegedus L. Controversies, Consensus, and Collaboration in the Use of (131)I Therapy in Differentiated Thyroid Cancer: A Joint Statement from the American Thyroid Association, the European Association of Nuclear Medicine, the Society of Nuclear Medicine and Molecular Imaging, and the European Thyroid Association. Thyroid 2019; 29: 461-470.