

## OPEN ACCESS

## EDITED BY

Carlotta Giani,  
University of Pisa, Italy

## REVIEWED BY

Valeriano Leite,  
Instituto Português de Oncologia Francisco  
Gentil, Portugal  
Sriram Gubbi,  
National Institute of Diabetes and Digestive  
and Kidney Diseases (NIH), United States

## \*CORRESPONDENCE

Giulia Sapuppo  
✉ giuliasapuppo@hotmail.it

## SPECIALTY SECTION

This article was submitted to  
Cancer Endocrinology,  
a section of the journal  
Frontiers in Endocrinology

RECEIVED 21 December 2022

ACCEPTED 23 February 2023

PUBLISHED 13 March 2023

## CITATION

Sapuppo G, Grasso S, Di Benedetto G,  
Belfiore A and Pellegriti G (2023)  
Prospective study and proposal of an  
outcome predictive nomogram in a  
consecutive prospective series of  
differentiated thyroid cancer based on the  
new ATA risk categories and TNM.  
*Front. Endocrinol.* 14:1128963.  
doi: 10.3389/fendo.2023.1128963

## COPYRIGHT

© 2023 Sapuppo, Grasso, Di Benedetto,  
Belfiore and Pellegriti. This is an open-access  
article distributed under the terms of the  
[Creative Commons Attribution License  
\(CC BY\)](#). The use, distribution or  
reproduction in other forums is permitted,  
provided the original author(s) and the  
copyright owner(s) are credited and that  
the original publication in this journal is  
cited, in accordance with accepted  
academic practice. No use, distribution or  
reproduction is permitted which does not  
comply with these terms.

# Prospective study and proposal of an outcome predictive nomogram in a consecutive prospective series of differentiated thyroid cancer based on the new ATA risk categories and TNM

Giulia Sapuppo <sup>1\*</sup>, Sonia Grasso<sup>1</sup>, Guenda Di Benedetto<sup>1</sup>,  
Antonino Belfiore<sup>1</sup> and Gabriella Pellegriti<sup>1,2</sup>

<sup>1</sup>Department of Clinical and Experimental Medicine, University of Catania, Garibaldi-Nesima Medical Center, Catania, Italy, <sup>2</sup>Researcher in Oncology, Department of Clinical and Experimental Medicine, University of Catania, Catania, Italy

**Introduction:** The personalized management of differentiated thyroid cancer (DTC) is currently based on the postoperative TNM staging system and the ATA risk stratification system (RSS), both updated in 2018 and 2015, respectively.

**Purpose:** We aimed to evaluate the impact of the last two editions of TNM and ATA RSS in the prediction of persistent/recurrent disease in a large series of DTC patients.

**Patients and methods:** Our prospective study included 451 patients undergone thyroidectomy for DTC. We classified the patients according to TNM (both VIII and VII ed.) and stratified them according to the ATA RSS (both 2015 and 2009). We then evaluated the response to the initial therapy after 12-18 months according to the ATA "ongoing" risk stratification, and analyzed the variables associated with persistent/recurrent disease by multivariate analysis.

**Results:** The performance of the last two ATA RSSs was not significantly different. By staging patients according to the VIII or VII TNM editions, we found significant differences only in the distribution of patients with structural disease classified in stages III and IV. At multivariate analysis, only T-status and N-status were independently associated with persistent/recurrent disease. Overall, ATA RSSs and TNMs showed low predictive power in terms of persistent/recurrent disease (by Harrell's test).

**Conclusions:** In our series of DTC patients, the new ATA RSS as well as the VIII TNM staging provided no additional benefit compared to the previous editions. Moreover, the VIII TNM staging system may underestimate disease severity in patients with large and numerous lymph node metastases at diagnosis.

#### KEYWORDS

differentiated thyroid cancer, thyroid cancer outcome, thyroid cancer risk factors, persistent or recurrent disease, TNM staging system, risk stratification system TNM staging and ATA risk stratification systems

## 1 Introduction

Thyroid cancer (TC) is the most common endocrine malignancy, representing about 90% of cases. TC is the most rapidly increasing cancer in the United States, where its incidence increased by 211% in the years 1975-2013; however, due in part to the adoption of more conservative diagnostic criteria, the incidence rate declined by 2.5% per year from 2014 to 2018 (1). The mortality from TC is very low and the death rate increased slightly during 2008 to 2017 (0.6% per year) despite earlier diagnosis and better treatment. In recent years TC death rate appears fairly stable (1). The prognosis of differentiated thyroid cancer (DTC) is excellent with a 5-year survival rate of 99-100% for localized, 98% for regional and 53% for metastatic disease (2).

In 2015 the American Thyroid Association (ATA) guidelines for DTC management (3) introduced a new risk stratification system, which included some supplementary prognostic factors such as lymph node characteristics (number, size and extranodal extension), mutational status, and foci's number of vascular invasion. However, as stated in the ATA guidelines "the incremental benefit of adding these specific prognostic variables to the 2009 initial risk stratification system has not been established" and the added value of this modification has not yet been validated.

The TNM classification has also been updated in 2018 to better predict DTC survival (4). As a result of the changes, the eighth edition of the TNM results in the downstaging of a significant percentage of patients to more accurately indicate the low specific risk of death. Currently, the appropriate clinical-therapeutic management of TC requires post-surgical TNM staging to predict survival, and the assessment of the risk of persistent/recurrent disease according to the ATA risk stratification system. During the follow-up this risk is also regularly re-evaluated according to the "ongoing risk stratification" (3) based on the measurement of thyroglobulin (Tg), thyroglobulin antibodies (TgAb), neck ultrasound, post-131-I WBS, other imaging evaluation as required.

As expected, a more personalized and accurate assessment of the risk of persistent or recurrent disease and death from TC has a significant impact on the initial treatment decision (extension of thyroid and/or lymph node surgery, need for radioactive iodine ablation/therapy, need for TSH suppressive therapy) and

appropriate management strategies during short and long-term follow-up.

In the present study we prospectively evaluated, in a continuous series of 451 DTC patients, the variables associated with persistent/recurrent disease (biochemical and structural) and the predictive power of the different ATA risk categories and TNM staging systems. Our findings may help establishing a tailored treatment management based on tumor and patient characteristics.

## 2 Patients and methods

We studied a consecutive series of 451 DTC patients undergone thyroidectomy with or without lymphnode dissection from October 2017 to February 2020 and followed up at our thyroid outpatients' clinic. The median follow-up was 20.5 (IQR 14.7-27.4) months.

We staged the tumors according to the 7th and 8th TNM editions: T (the greatest size of the primary tumor) and N (regional lymphnodes metastases) through histology. We indicated N0 when all removed lymphnodes were negative or if there is no radiological or clinical evidence of lymphnode metastasis, N1a if only the central compartment (levels VI-VII) is involved and N1b if latero-cervical nodes (levels I-V) were positive. M (distant metastases) was evaluated according to the post-surgical 131I-whole body scan (WBS) and/or other imaging evidence.

Disease status was evaluated in all patients through neck ultrasound and Tg and TgAb determination after surgery and periodically during the follow-up (every 3-6 or 12 months).

After 12-18 months from the first evaluation, the response to initial therapy (surgery with or without iodine treatment) was assessed with neck ultrasound and both serum Tg and TgAb measurements, either basal or TSH stimulated (L-thyroxine, LT4 withdrawal or rhTSH injection) in radioiodine treated patients. According to their response to the initial therapy (ongoing stratification) patients were re-classified as excellent response (ER), indeterminate (IR), biochemical incomplete (BIR) or structural incomplete response (SIR).

Subsequent follow-up was modulated based on the initial risk evaluation and the first treatment response.

The level of TSH-suppression was based according to the ATA guidelines: substitutive LT4 therapy in low-risk patients without

evidence of disease; mild TSH suppression (0.1-0.4 mU/L) in intermediate risk patients or biochemical disease; complete suppression (< 0.1 mU/L) in high risk patients and/or if structural disease.

All patients with the evidence of persistent or recurrent disease underwent additional morphological exams (computed assisted tomography (CT scan), magnetic resonance imaging, bone scan, positron emission tomography). If patients were not cured, further treatments such as radioiodine treatment therapy, other surgeries or different therapies) were brought off.

### 3 Statistical analysis

Categorical variables were expressed as frequencies and percentages; quantitative normally distributed ones as mean  $\pm$  standard deviation (SD) and non-normally distributed ones as median with interquartile range (IQR). The normality was verified through the Kolmogorov-Smirnov test.

The Chi-square test with Yates's correction or Fisher's test were used to analyze the categorical variables. Multivariate analysis was outbrought off through the logistic regression including only significant variables for recurrent/persistent disease at univariate analysis.

The predictive power of the different ATA risk categories and TNM staging systems at short-term re-evaluation was assessed by the Harrell's C concordance index (C-index).

A nomogram was implemented based on the parameters that resulted significantly associated with the risk of recurrent/persistent disease at the multivariate logistic regression analysis, with a risk score for predicting the probability of persistent and recurrent disease.

TABLE 1 Clinical and histopathological characteristics of the 451 DTC patients.

	n.	(%)
Patients (n.)	451	
Follow-up median (IQR) (months)*	20.5 (14.7-27.4)	
Age median (IQR) (y)	47.5 (36.9-57.8)	
Gender		
F/M (ratio)	324/127 (2.6/1)	
Histotypes		
Papillary	433	96.0
Follicular	12	2.7
Papillary-follicular	4	0.9
Poorly-differentiated	2	0.4
M1	11	2.4
Minimal extrathyroid extension (ETE)	97	21.5
Multifocality	176	39.0

\* Follow-up minimum 6 months.

A p-value <0.05 was considered statistically significant for all analyses. Data analysis was performed using the Stata software version 16.

## 4 Results

Clinical and histopathological characteristics are shown in Table 1. Most patients were females with a F/M ratio of 2.6/1.0. Median age at diagnosis was 47.5 yrs (IQR 36.9-57.8). Histotype was papillary in almost all cases (96%). 14.2% of patients had an aggressive PTC variant (diffuse sclerosing, tall cell, insular or columnar).

Eleven patients (2.4%) had distant metastases at diagnosis (10 lung metastases and 1 with lung and bone metastases).

Minimal extrathyroid extension was present in 21.5% and multifocality in 39.0% of patients.

Most patients fell into the T1 and T2 categories, especially with the new TNM staging system; 26.8% of patients had central (N1a) and 13.7% latero-cervical lymphnode metastases (N1b).

Two hundred seventy-seven (61.4%) patients were treated with 131I with different activities: 85 patients with 30 mCi, 186 patients with 100 mCi, 3 patients with 50 mCi, 1 patient with 70 mCi, 1 patient with 150 mCi e another 1 with 200 mCi. 156 patients were treated after L-T4 withdrawal and 121 after rhTSH administration. Six (1.3%) patients required another surgical treatment in the neck and 6 (1.3%) patients underwent to a second 131I treatment after about 12 months from the first.

### 4.1 TNM staging system and risk stratification system

TNM, staging (VII and VIII editions) and Risk stratification classifications are shown in Table 2.

Applying VII TNM staging system 60% tumors were T1a and T1b, instead applying VIII TNM more than ¾ of the patients (77.4%) fall into this category. This different percentage depends mainly on the removal of minimal extrathyroid extension from T3 classification.

Regarding the lymph node status, applying VII TNM staging 12.0% patients were N0 47.5% Nx; instead applying VIII TNM staging almost 60% of cases were N0a or N0b (268, 59.4%) (due to the removal of Nx in the new TNM system). Thesame percentages were in N1a and N1b classes.

Six patients presented latero-cervical metastases without involvement of the central compartment (skip metastases).

Patients were also staged comparing VII vs VII TNM staging system (Figure 1): most patients fell into stage I using both classifications, respectively 71.9% vs 86.5%. Using VIII TNM edition there was a significative downstaging in all categories (about 30%), mostly from stage III and IVA into stage I and II.

At initial evaluation, patients were subdivided into three different risk categories with few variations in percentage due mainly to the lymph node number and size categorization using 2009 or 2015 ATA risk stratification: 42.8% and 45.5% low, 54.8%

TABLE 2 TNM, staging (VII and VIII editions) and risk categories (2009 and 2015) of the 451 DTC patients.

		n.	(%)			n.	(%)
TNM (VII ed.)				TNM (VIII ed.)			
T status (T)				T status (T)			
T1a		166	36.8	T1a		206	45.7
T1b		105	23.3	T1b		143	31.7
T2		56	12.4	T2		73	16.2
T3		113	25.1	T3a		5	1.1
T4a		5	1.1	T3b		13	2.9
Tx		6	1.3	T4a		5	1.1
				Tx		6	1.3
N status (N)				N status (N)			
N0		54	12.0	N0a		124	27.5
N1a		121	26.8	N0b		144	31.9
N1b		62	13.7	N1a		121	26.8
Nx		214	47.5	N1b		62	13.7
Staging (VII ed.)				Staging (VIII ed.)			
<45	I	202	44.8	<55	I	298	66.1
	II	4	0.9		II	10	2.2
≥45	I	122	27.1	≥55	I	92	20.4
	II	19	4.2		II	39	8.6
	III	62	13.7		III	5	1.1
	IVA	34	7.5		IVA	0	
	IVB	0			IVB	7	1.6
	IVC	8	1.8				
Risk categories at first evaluation 2009 guidelines				Risk categories at first evaluation 2015 guidelines			
Low		193	42.8	Low		205	45.5
Intermediate		247	54.8	Intermediate		231	51.2
High		11	2.4	High		15	3.3

and 51.2% intermediate and 2.4% and 3.3% high risk according to respectively 2009 and 2015 ATA risk system.

## 4.2 Response to initial therapy, 12-18 months after initial treatment, in all patients

After initial treatment, 63.9% of patients presented with excellent response. However, 35 patients (not ablated) had basal Tg between 0.2 and 1 ng/mL, stable and compatible with small thyroid remnant. After 12-18 months from initial treatment, just over a third of patients (36.1%) were not cured.

In particular 82 patients presented an indeterminate response (68 patients had indeterminate Tg or TgAb and 14 patients non-specific

finding at neck ultrasound or CTscan). Five patients had biochemical incomplete response and 76 patients structural incomplete response (52 had lymph node metastases, almost all small in number and size; 10 lung metastases; 1 only bone metastases and 5 lung and bone metastases, 5 lung and LN metastases, 2 lung, bone and local disease).

### 4.2.1 Response to initial therapy according to ATA risk categories

As expected, the percentage of patients with an excellent response decreased through risk categories, being less frequent in intermediate- and mostly high-risk patients with both ATA risk categories (Figure 2).

In low risk patients, no significant difference was found in both TNM: approximately 72% had an ER, 19% IR, 1% BIR and 7.8% SIR.

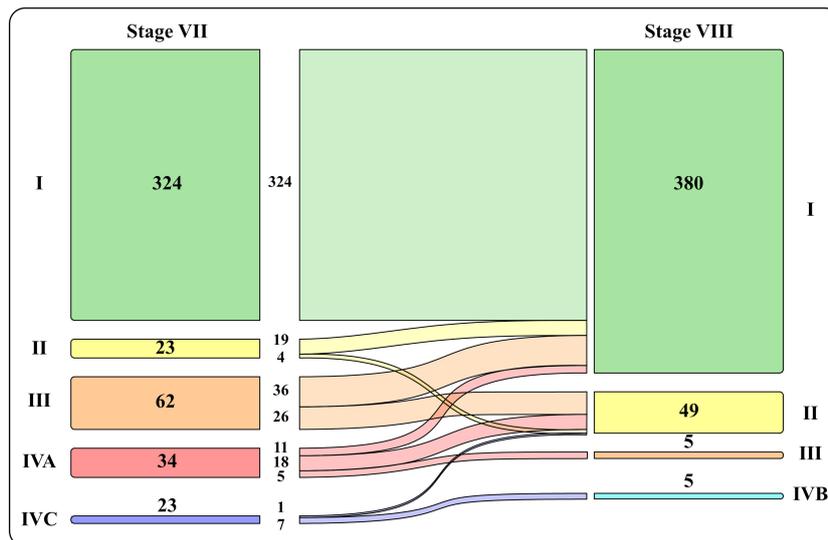


FIGURE 1 Graphic representation of patients' distribution into the four stages with VII vs VIII TNM staging system.

Also in intermediate risk patients, no significant difference was found: approximately 60% ER, 17.5% IR, 1% BIR and 21% SIR.

Instead in high risk patients, 20% vs 9.1% had ER, 13.3% vs 9.1% an IR and 66.7% vs 81.8% SIR, respectively using 2009 and 2015 ATA risk stratification.

The results did not change after analysing classic PTC subtype and all DTC separately.

Taking into account only structural incomplete response (SIR) at univariate analysis, the factors associated to the presence of persistent disease are, beyond the same risk factors of above, were male gender and multifocality. At multivariate analysis T status and lateral-lymph node were independent predictors of disease (Table 4).

Since the multivariate analysis showed that lymph node metastases have a significant impact to predict persistent/

### 4.3 Structural disease after 12-18 months after initial treatment according to VII vs VIII TNM staging

As expected, the percentage of patients with structural disease increases through the stages mostly using VIII TNM ed. (Figure 3).

Comparing the two TNM editions, a similar percentage of disease was observed in stage I and II both using VII or VIII TNM edition instead a significant difference in stage III and IV, respectively p=0.004 and 0.04.

### 4.4 Predictors of persistent/recurrent disease at 12-18 months after first treatment

Risk factors of persistent disease (either morphologic or biochemical) at last disease assessment are presented in Table 3.

At univariate analysis, the factors associated to the presence of disease were: T status, the presence of lymph node metastasis, both in the central compartment and in lateral compartments (N1a and N1b status), the presence of more than five lymph node metastases, ATA risk intermediate or high and radioiodine treatment therapy.

At multivariate analysis, both T status and lateral lymph node metastasis were patient features independently predicting persistent/recurrent disease (higher O.R. for N1b= 3.07, p<0.001).

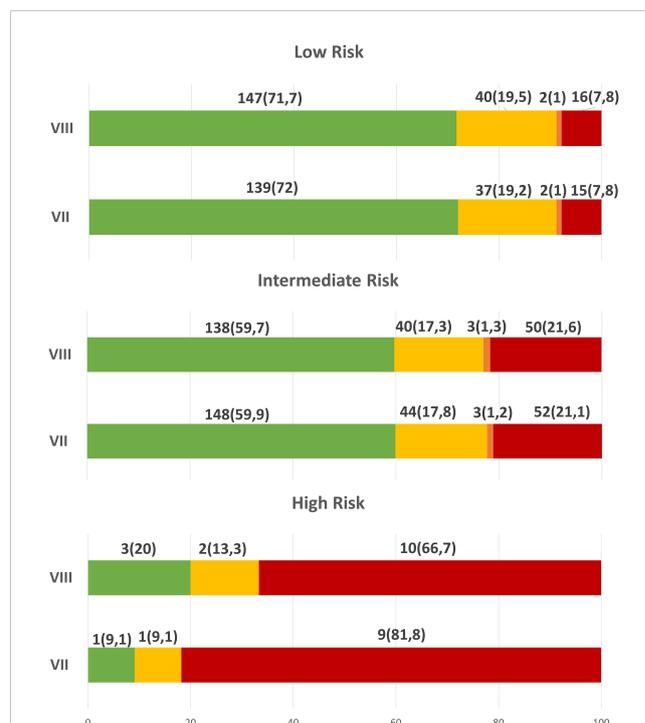


FIGURE 2 Response to initial therapy, 12-18 months after initial treatment, according to ATA risk categories.

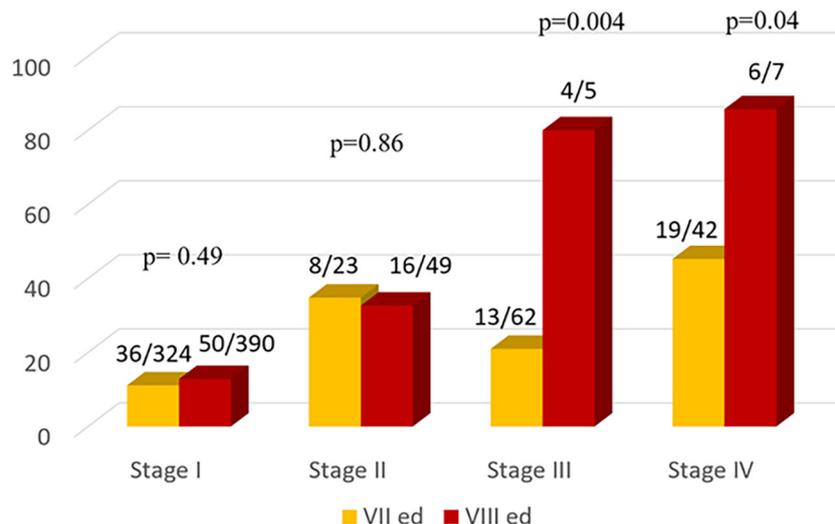


FIGURE 3

Comparison of the Number of Patients with structural disease according to the Seventh (yellow bars) and Eighth (red bars) TNM Editions.

recurrent disease, LN characteristics (localization and number) have been investigated in detail.

As expected not only the presence of positive LN (N1a + N1b vs N0/Nx), but also the number of nodal metastases and their location (N1b vs N1a) were relevant risk factors for persistent/recurrent disease persistence at 12-18 months after first treatment.

Persistent disease was significantly higher in N1 vs N0a/N0b patients (47% vs 28.7%,  $p < 0.001$ ).

Analyzing the number of nodal metastases, using the cut off of 5 positive nodes, the likelihood of persistent disease was higher for patients with  $> 5$  vs  $\leq 5$  metastatic lymph nodes or negative lymph node (respectively 64.6%, 40.8% and 28.7%) ( $p < 0.001$ ) (Table 3).

As for lymph node location (N1a or N1b), the frequency of disease at 12-18 months after first treatment progressively increased from 28.7% in N0a/N0b to 40.5% in N1a and to 59.7% in N1b (OR 1.69 and 3.67 and  $p = 0.02$  and  $< 0.001$ , respectively) (Table 3).

Analyzing the number and location of nodal metastases, for 121 N1a patients, the probability of persistent disease was higher when more than 5 lymph nodes were involved (7/13 cases, 53.8%) compared with  $\leq 5$  lymph nodes involved (42/108, 38.9%); this difference was not statistically significant ( $p = 0.30$ ).

For 62 N1b patients, the probability of persistent disease was sensibly higher when more than 5 lymph nodes were involved (24/35, 68.5%) compared with  $\leq 5$  lymph nodes involved (13/27, 48.1%); also this difference was not statistically significant ( $p = 0.10$ ).

#### 4.5 Predictive power of risk stratification system (ATA 2009 vs. ATA 2015) and TNM staging system (VII vs. VIII edition)

To evaluate the accuracy of the prediction of the different ATA risk stratification systems (2009 vs. 2015) we calculated the Harrell C concordance index (C-index) considering the disease events at 18-24 months of re-evaluation. We found no differences in

predictive power between the ATA 2009 and ATA 2015 criteria for biochemical and structural disease (C-index = 0.576 and 0.574, respectively for ATA 2009 and ATA 2015).

This still held true when only patients with structural disease were considered (C-index = 0.656 and 0.657, respectively for ATA 2009 and ATA 2015).

Similarly, the VII and the VIII TNM staging system resulted to have a similar and low predictive power (C-index = 0.560 vs 0.570, for the VII and the VIII edition, respectively).

#### 4.6 Development of a nomogram to predict persistent/recurrent disease

A nomogram incorporating all the significant parameters was constructed based on the multivariate logistic model identified in Table 3. For each parameter we obtained a corresponding prognostic points as shown in Figure 4. The point values for all predictor variables were summed to reach a total score. This value was plotted on the total score axis and a vertical line drawn from this axis straight up that indicates the patient's probability of persistent/recurrent disease at re-evaluation after the first therapy.

### 5 Discussion

Differentiated thyroid cancer (DTC) is the most common endocrine malignancy (5). The most frequent histotype (over 85%) is papillary (PTC). DTC incidence was quite stable until 1990s and in recent decades has rapidly grown, more than any other cancer (6), mainly due low-risk thyroid cancer increasing (<https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/cancer-facts-figures-2022.html>) (1).

DTC's prognosis is generally excellent, but in some patients tumors are aggressive with poor outcome. During 2008 to 2017

(0.6% per year) TC's death rate increased slightly but appears to have stabilized in recent years.

It is not known if the increasing incidence of PTC is true or whether it is due to the overdiagnosis of indolent PTCs.

Finding the variables identifying these tumors is a major issue to decide the appropriate management.

The American Thyroid Association (3) introduced a new risk stratification system with additional prognostic variables. Moreover, to better predict DTC survival, also the TNM classification was updated in 2018. A significant number of

patients were downstaged by the 8th edition (TNM-8) into lower stages to more accurately reflect the low risk of dying, but underestimating the risk of persistent/recurrent disease and death in some patients due to the fact that all young patients without distant metastases fall into stage I.

The changed risk stratification and TNM staging have a significant consequences on the earliest therapeutic judgment and subsequent follow-up management.

Concerning the risk stratification, different data have been published. In Steinschneider et al. data (7) approximately 70% of

TABLE 3 Risk factors of disease (biochemical and structural) at 12-18 months after primary treatment.

Variable	Biochemical and structural disease	Univariate analysis OR [95%CI]	p	Multivariate analysis OR [95%CI]	p
<b>Age</b> - < 55 years - ≥ 55 years	105/308 (34.1) 58/143 (40.6)	1.32 (0.88-1.99)	0.18		
<b>Gender</b> - Female - Male	112/324 (34.6) 51/127 (40.2)	1.27 (0.83-1.94)	0.27		
<b>Lymph node surgery at primary treatment</b> - Performed - Not performed	117/307 (38.1) 46/144 (31.9)	1.31 (0.86-2.0)	0.20		
<b>Aggressive histology</b> - No - Yes	139/387 (35.9) 24/64 (37.5)	1.07 (0.62-1.85)	0.80		
<b>T status</b> - T1a - T1b - T2 - T3a - T3b - T4a	61/206 (29.6) 49/143 (34.3) 38/73 (52.1) 0/5 8/13 (61.5) 3/5 (60.0)	1.24 (0.78-1.96) 2.58 (1.49-4.46) 3.80 (1.20-12.09) 3.57 (0.58-21.88)	0.36 <0.001 0.01 0.14	0.99 (0.61-1.61) 2.11 (1.19-3.74) 2.54 (0.77-8.35) 2.96 (0.48-18.46)	0.98 0.009 0.12 0.24
<b>Multifocal</b> - No - Yes	92/274 (33.6) 70/176 (39.8)	1.31 (0.88-1.93)	0.18		
<b>Extra thyroidal invasion</b> - No - minimal	123/355 (34.6) 40/96 (41.7)	1.35 (0.85-2.14)	0.20		
<b>Lymph node metastases</b> - Absent - Present	77/268 (28.7) 86/183 (47.0)	2.20 (1.48-3.26)	<0.0001	Not included	
<b>Number of N1 at primary surgery</b> - N0a/N0b - ≤5 N1 - >5 N1	77/268 (28.7) 55/135 (40.7) 31/48 (64.6)	1.71 (1.10-2.65) 4.52 (2.37-8.65)	0.016 0.0000	Not included	
<b>Location of N1 at primary surgery</b> - N0a/N0b - N1a - N1b	77/268 (28.7) 49/121 (40.5) 37/62 (59.7)	1.69 (1.08-2.65) 3.67 (2.07-6.51)	0.02 0.0000	1.40 (0.87-2.25) 3.07 (1.68-5.62)	0.16 <0.001
<b>ATA risk stratification</b> - Low - Intermediate - High	58/205 (28.3) 93/231 (40.3) 12/15 (80.0)	1.71 (1.14-2.55) 10.14 (2.76-37.24)	0.008 0.0000	Not included	
<b>Radioiodine treatment</b> - Not performed - Performed	37/174 (21.3) 126/277 (45.5)	3.09 (2.00-4.77)	0.0000	Not included	

TABLE 4 Risk factors of structural disease at the 12-18 months assessment after primary treatment.

Variable	Structural disease n.76	Univariate analysis OR [95%CI]	p	Multivariate analysis OR [95%CI]	p
<b>Age</b>	51/308 (16.6)	1.07 (0.63-1.81)	0.80		
- < 55 years	25/143 (17.5)				
- ≥ 55 years					
<b>Gender</b>	46/324 (14.2)	1.87 (1.12-3.13)	0.016	1.63 (0.92-2.90)	0.09
- Female	30/127 (23.6)				
- Male					
<b>Lymph node surgery</b>	57/307 (18.6)	0.67 (0.38-1.17)	0.155		
- Performed	19/144 (13.2)				
- Not performed					
<b>Aggressive histology</b>	63/387 (16.3)	1.31 (0.67-2.55)	0.42		
- No	13/64 (20.3)				
- Yes					
<b>T status</b>	23/206 (11.2)	1.37 (0.73-2.58)	0.33	1.12 (0.58-2.18)	0.73
- T1a	21/143 (14.7)	3.21 (1.65-6.26)	<0.001	2.76 (1.37-5.56)	0.005
- T1b	21/73 (28.8)	4.97 (1.50-16.49)	0.004	3.18 (0.82-12.34)	0.09
- T2	0/5 (0.0)	11.93 (1.89-75.22)	0.001	10.85 (1.60-73.52)	0.01
- T3a	5/13 (38.5)				
- T3b	3/5 (60.0)				
- T4a					
<b>Multifocal</b>	38/274 (13.9)	1.65 (1.00-2.72)	0.04	1.16 (0.66-2.04)	0.60
- No	37/176 (21.0)				
- Yes					
<b>Extra thyroidal invasion</b>	54/355 (15.2)	1.66 (0.95-2.89)	0.07		
- No	22/96 (22.9)				
- minimal					
<b>Lymph node metastases</b>	30/268 (11.2)	2.66 (1.61-4.42)	<0.001	Not included	
- Absent	46/183 (25.1)				
- Present					
<b>Number of N1 at primary surgery</b>	30/268 (11.2)	5.46 (3.26-9.15)	<0.001	Not included	
- N0a/N0b	53/130 (40.8)	14.47 (7.16-29.22)	<0.001		
- ≤5 N1	31/48 (64.6)				
- >5 N1					
<b>Location of N1 at primary surgery</b>	30/268 (11.2)	2.39 (1.35-4.22)	0.002	1.59 (0.84-3.03)	0.15
- N0a/N0b	28/121 (23.1)	2.76 (1.39-5.47)	0.002	4.04 (2.00-8.13)	<0.001
- N1a	16/62 (25.8)				
- N1b					
<b>ATA risk stratification</b>	16/205 (7.8)	3.26 (1.79-5.94)	<0.001	Not included	
- Low	50/231 (21.6)	23.63 (7.20-77.55)	<0.001		
- Intermediate	10/15 (66.7)				
- High					
<b>Radioiodine treatment</b>	8/174 (4.6)	6.75 (3.16-14.44)	<0.001	Not included	
- Not performed	68/277 (24.5)				
- Performed					

patients were low risk, 25% intermediate and 5.2% high risk patients. Instead, in our data showed fewer patients in low category (45%) and higher in intermediate risk (51%) than in literature with no significant difference between 2009 and 2015 risk stratification.

Regarding the staging, a large proportion of patients were downstaged in the 8th edition (30-40%) vs the 7<sup>th</sup> one, mostly due to the increasing of the age cut-off to 55 years, the downclassification of T3 disease, and the overall downstaging of lymphnode metastases (4, 7-9), with a minimal impact on the expected 10-year disease-specific survival despite the large

proportion of shifted patients to stage I and II. Kim et al. (4) found that 41% of patients were downstaged and inevitably more patients with recurrences or deaths were found in the lower stages: 17% of patients downstaged from stage III to II had recurrent disease, 25% and 13.6% died in the group downstaged respectively from stage IV to III and from stage IV to II.

Also our data shows an important downstaging (about 30%) mostly from staging III and IVA to I, II and III. In particular the downstaging concerned 82.6% of stage II (into stage I), 100% of stage III (into stage I and II) and 100% of stage IV (into stage I, II and III).

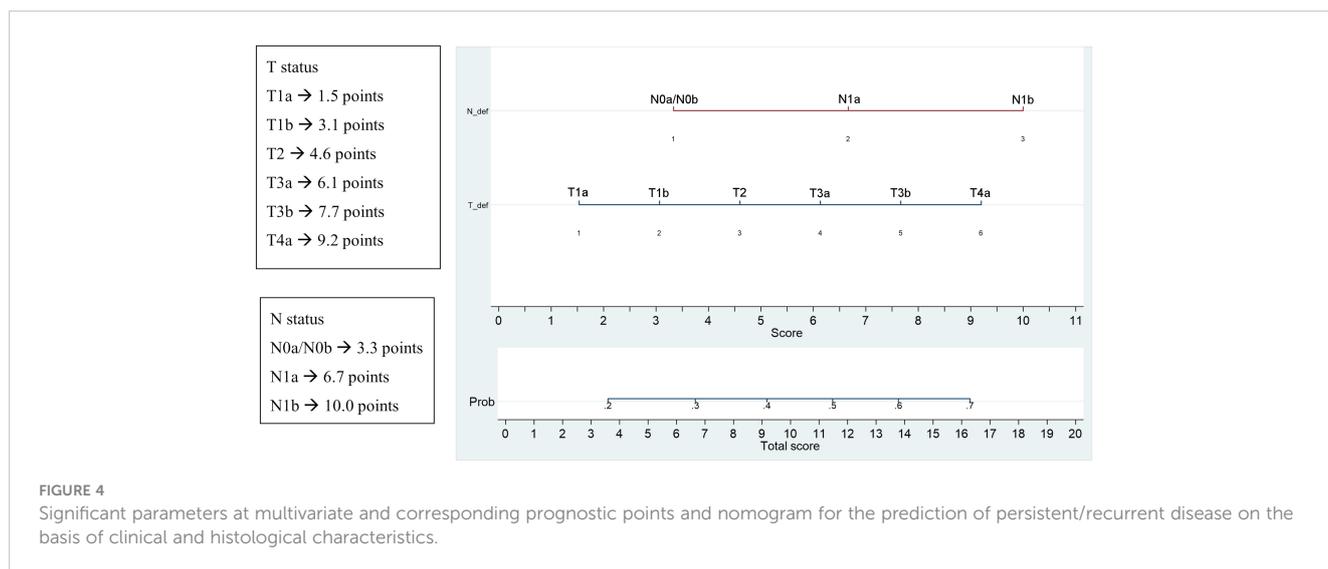


FIGURE 4

Significant parameters at multivariate and corresponding prognostic points and nomogram for the prediction of persistent/recurrent disease on the basis of clinical and histological characteristics.

Evaluating the different response at initial therapy (dynamic or ongoing risk stratification), a paper published 10 years ago by Tuttle et al. (10) found an ER in 86% of low risk patients, in 57% of intermediate-risk and in 14% in the high-risk. In 11% of low-risk patients, 22% in intermediate-risk and 14% in high-risk had BIR and 3% of patients in the low-risk, 21% in the intermediate-risk, and 72% in the high-risk had SIR.

Another paper evaluating 441 patients (7) showed that the proportion of intermediate/high-risk patients in stages I–II increased considerably according to TNM-8 and that patients downstaged in stage II with TNM-8 had more lymphnode metastases, surgeries, disease persistence and an increased disease-specific mortality (non-significant) vs to TNM-7. They found a similar rates of persistent and recurrent disease in stage I in both TNM editions, but higher in stages II ( $p = 0.05$ ) and III ( $p = 0.03$ ) in TNM-8 vs TNM-7. Therefore the new TNM guaranteed a more accurate system to assess mortality and persistence disease but that the severity, mainly in in stage II patients or in the 45–55-year old group, should not be underestimated as a result of the important downstaging of some particular groups of patients.

In the present study, after initial treatment, 63.9% of patients presented ER and 36.1% patients were not cured, of which half presented an IR, a little less cases SIR and few patients BIR.

Our data are different compared to previous data mostly for lower percentage of the ER in low risk patients (71.7%), due to an higher number of patients IR (19.5% of low risk patients, 17.3% in intermediate-risk patients and in 13.3% of the high-risk patients) and very lower percentage of patients with BIR (1% in the low-risk, 1.3% in the intermediate-risk and 0% in the high-risk group); regarding SIR the rate are similar to the other paper (in 7.8% in the low-risk, 21.6% in the intermediate-risk, and 66.7% in the high-risk group).

Moreover in our patients the rates of structural disease in stage I and II was similar in both editions and it was significantly higher in stages III ( $p = 0.004$ ) and IV ( $p = 0.04$ ) in VIII compared to VII edition.

Another issue still to be validated is the timing of the re-evaluation. Based on our data, it seems that the restaging at 12–18 months could be too early as many patients with indeterminate response could change into excellent response (for example for TgAb still positive but stable or declining). This hypothesis will be evaluated by extending the follow-up.

Concerning the risk factors, most reports show that age, gender, aggressive variants, tumor size, lymphnode metastases are the most important predictors of outcome (3).

In a recent paper in 2020, Shin et al. (11) found that at multivariate analysis tumor size > 4 cm, multifocality and nodal factors were the independent factors of recurrence free survival.

In our data only tumor size and lymph node metastases independently predicted short term outcome, instead the other risk factors were not statistically significant. In fact, tumor size is an independent risk factors from T2 up to T4a both for recurrent/persistent disease (biochemical and structural) and for only structural disease.

In a retrospective analysis of 574 patients with PTC, Tran et al. (12) found that tumor size predicted recurrence-free survival on multivariate analysis and Pellegriti et al. (13) showed that tumor size ( $\leq 1.0$  cm versus. 1.1–1.5 cm) was not predictive of recurrence. Nguyen et al. (14) showed that, in SEER database, the 10-year relative survival rates for tumors sized 1.5 cm or larger and tumors less than 1.5 cm were 95.4% and 99.8%, respectively.

Regarding lymph node metastases, their clinical relevance in PTC has been a debated matter for decades (15, 16). To date, the LN metastases impact at diagnosis on recurrence's risk is well documented in many papers, including a recent study of our group (17, 18). For many years, however, only neck lymphnode metastases, without other specified variables, was evaluated as a PTC prognostic factor.

The impact of neck metastatic LN on PTC risk stratification was better defined in 2015 ATA guidelines, in which additional informations, as the number, size or extranodal extension of metastatic lymphnodes, were included in the evaluation. To date,

these additional characteristics have not been validated and their relevance in defining the recurrence risk has yet to be quantified.

In our series the positivity of lymph node metastases, the number ( $\leq 5$  or  $>5$ ) and the location (N1a and N1b) are effective predictors of the outcome of the patients at 12-18 months after the first treatment, both for recurrent/persistent disease (biochemical and structural) and for only structural disease.

Data on the location of metastases, N1a or N1b, were judged insufficient to include this information in the clinicopathologic variables to estimate the risk in PTC (3) (recommendation #48, [B20], paragraph 1, line 16).

In a recent publication of our group (19), N1b worsened the prognosis and may be related to the appearance of distant metastases, which are considered the best surrogate index for cancer-specific death.

N1b status could be a marker of more aggressive or more advanced disease at diagnosis in PTC patients and associated with other ones (as molecular alterations) of cancer aggressiveness.

Concerning the mETE, its role on disease specific survival (DSS) and on overall survival (OS) have been evaluated by several studies but it's still consider a controversial prognostic factor. Some authors (19–22) showed similar clinical outcomes. However, Castagna et al. (23) showed poorer outcome, in term of persistent/recurrent structural disease and tumor-related death in patients with mETE vs tumors  $>1.5$  cm without extrathyroid extension (11.8% vs. 5.1%), concluding that only small mETE cancers should be consider at low-risk.

Recently, an expansion of TNM-8 has been published (Telescoping) to test the subcategories, according to the mETE, to get a better estimate of the prognosis and to plan the follow-up management. In the next few years, the impact of mETE for each tumor class will be available.

Comparing disease specific survival (DSS) between TNM VII vs VIII ed, Tam et al. (20) concluded that DSS in both TNM editions is similar, although through the updated TNM-8 the 10-year DSS appears more proper between stages. For stages I–IV with TNM-7 the 10-year DSS ranged from 100% to 82.6% ( $p < 0.001$ ) and the 10-year OS from 95.8% to 59.7% while with TNM-8 from 99.8% to 71.9% ( $p < 0.001$ ) and from 94.3% to 34.6%, respectively.

Contrasting, results from Chung et al. (21), analyzing a large series of 3,176 DTC patients, and Jeon et al. (22), investigating the predictive capability of DSS of TNM-8 compared to TNM-7 in 1,613 DTC patients, showed that TNM-8 has a higher power to differentiate patients in each stage and to predict also the DSS.

Although therefore with the TNM-8 an improved assignment of patients at high risk of dying from DTC into more advanced stages of disease seems evident, on the other hand, leads to the erroneous belief that the disease is less aggressive. Nearly 50% of the cancer-related deaths are involve patients in stages I–II, compared to none with the TNM-7 (7). Having thyroid cancer a very low mortality, in some patients the risk of death is not always related to recurrence's risk.

Recently, several studies showed a better predictability for the new TNM (6, 24). A more accurate survival predictions is suggested when TNM-8 is applied, due to the downstaging of a significant number of patients (about 30%).

In our data, being the median follow-up of 20.5 months, mortality was not assessed so an analysis evaluating the presence of distant metastases, good surrogate for predicting mortality, was carried out. The C-Harrel test to evaluated the power to predict disease after short follow-up found no difference using VII and VIII TNM editions (both for biochemical and structural disease and also for structural disease).

Lastly, generally, cancer nomograms are prediction tools to assess the risk based on specific patient's and tumor's characteristics and to predict the likely outcomes of different therapies. By integrating different prognostic variables, the nomogram has the ability to create an individual numerical probability of a clinical event, useful to improve disease prognostication and therefore a personalized follow-up.

In literature, recently, several nomograms for prediction the risk of death from thyroid cancer have been proposed (23, 25). In 2016 Lang et al. (24) validated a nomogram for PTC patients with an excellent discriminating skill and accuracy in predicting 10-years-disease-specific death and recurrence. In 2017 two nomograms were proposed (26), one for regional recurrence-free survival and one for distant recurrence-free survival prediction with a C-index of 0.72 and 0.83, respectively.

Also the nomogram elaborated by our data could be useful to plan an individualized follow-up for each patient based on the score obtained from the risk calculation. The limit of this system is the lack of external validation

In conclusion, although the new TNM-8 compared to the previous TNM edition would seem to better discriminate the disease-specific death, in some patients (as N1b at diagnosis, mainly if numerous and large) could underestimate the severity of disease due to the significant downstaging and bring to a non-negligible treatment burden. In fact while in the TNM-7 N1b patients were included in advanced risk categories, the new TNM-VIII does not discriminate the death rate according to the lymphnode location, downstaging some patients (particularly old patients and/or N1b) decreasing the discriminating ability for the few patients with a negative outcome, despite categorized as stage II. Therefore, a careful follow-up is needed for downstaged patients.

Further prospective studies are needed to better define the real effectiveness of the 2015 ATA risk stratification system and the VIII TNM staging system.

## Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## Ethics statement

The studies involving human participants were reviewed and approved by Comitato Etico Catania 2 A.O. Garibaldi in Catania Piazza Santa Maria del Gesù, 5 95124 Catania. The patients/participants provided their written informed consent to participate in this study.

## Author contributions

I) Conception and design: GS and GP. (II) Collection and assembly of data: GS, GD, MR, AB and GP. (III) Data analysis and interpretation: GS and GP. (IV) Manuscript writing: GS and GP. (V) Final approval of manuscript: All authors. All authors contributed to the article and approved the submitted version.

## References

1. Available at: <https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/cancer-facts-figures-2022.html> (Accessed 30 nov 22).
2. Available at: <https://seer.cancer.gov/statfacts/html/thyro.html>.
3. Haugen BR. American Thyroid association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer: What is new and what has changed? *Cancer* (2015) 123(3):372–81. doi: 10.1002/cncr.30360
4. Kim K, Kim JH, Park IS, Rho YS, Kwon GH, Lee DJ. The updated AJCC/TNM staging system for papillary thyroid cancer (8th edition): From the perspective of genomic analysis. *World J surgery*. (2018) 42(11):3624–31. doi: 10.1007/s00268-018-4662-2
5. Davies L, Welch HG. Current thyroid cancer trends in the united states. *JAMA otolaryngology- Head Neck surgery*. (2014) 140(4):317–22. doi: 10.1001/jamaoto.2014.1
6. Pellegriti G, Frasca F, Regalbuto C, Squatrito S, Vigneri R. Worldwide increasing incidence of thyroid cancer: update on epidemiology and risk factors. *J Cancer Epidemiol* (2013) 2013:965212. doi: 10.1155/2013/965212
7. Shteinshnaider M, Mualllem Kalmovich L, Koren S, Or K, Cantrell D, Benbassat C. Reassessment of differentiated thyroid cancer patients using the eighth TNM/AJCC classification system: A comparative study. *Thyroid* (2018) 28(2):201–9. doi: 10.1089/thy.2017.0265
8. Perrier ND, Brierley JD, Tuttle RM. Differentiated and anaplastic thyroid carcinoma: Major changes in the American joint committee on cancer eighth edition cancer staging manual. *CA: Cancer J Clin* (2018) 68(1):55–63. doi: 10.3322/caac.21439
9. Amin MB, Greene FL, Edge SB, Compton CC, Gershenwald JE, Brookland RK, et al. The eighth edition AJCC cancer staging manual: Continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. *CA: Cancer J Clin* (2017) 67(2):93–9. doi: 10.3322/caac.21388
10. Tuttle RM, Tala H, Shah J, Leboeuf R, Ghossein R, Gonen M, et al. Estimating risk of recurrence in differentiated thyroid cancer after total thyroidectomy and radioactive iodine remnant ablation: Using response to therapy variables to modify the initial risk estimates predicted by the new American thyroid association staging system. *Thyroid* (2010) 20(12):1341–9. doi: 10.1089/thy.2010.0178
11. Shin CH, Roh JL, Song DE, Cho KJ, Choi SH, Nam SY, et al. Prognostic value of tumor size and minimal extrathyroidal extension in papillary thyroid carcinoma. *Am J surgery*. (2020) 220(4):925–31. doi: 10.1016/j.amjsurg.2020.02.020
12. Tran B, Roshan D, Abraham E, Wang L, Garibotto N, Wykes J, et al. The prognostic impact of tumor size in papillary thyroid carcinoma is modified by age. *Thyroid* (2018) 28(8):991–6. doi: 10.1089/thy.2017.0607
13. Pellegriti G, Scollo C, Lumerà G, Regalbuto C, Vigneri R, Belfiore A. Clinical behavior and outcome of papillary thyroid cancers smaller than 1.5 cm in diameter: study of 299 cases. *J Clin Endocrinol Metab* (2004) 89(8):3713–20. doi: 10.1210/jc.2003-031982
14. Nguyen XV, Choudhury KR, Eastwood JD, Lyman GH, Esclamado RM, Werner JD, et al. Incidental thyroid nodules on CT: evaluation of 2 risk-categorization methods for work-up of nodules. *AJNR Am J neuroradiology*. (2013) 34(9):1812–7. doi: 10.3174/ajnr.A3487

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

15. Lang BH, Ng SH, Lau LL, Cowling BJ, Wong KP, Wan KY. A systematic review and meta-analysis of prophylactic central neck dissection on short-term locoregional recurrence in papillary thyroid carcinoma after total thyroidectomy: Official journal of the American thyroid association. *Thyroid* (2013) 23(9):1087–98. doi: 10.1089/thy.2012.0608
16. Wang TS, Cheung K, Farrokhyar F, Roman SA, Sosa JA. A meta-analysis of the effect of prophylactic central compartment neck dissection on locoregional recurrence rates in patients with papillary thyroid cancer. *Ann Surg Oncol* (2013) 20(11):3477–83. doi: 10.1245/s10434-013-3125-0
17. Sapuppo G, Palermo F, Russo M, Tavarelli M, Masucci R, Squatrito S, et al. Latero-cervical lymph node metastases (N1b) represent an additional risk factor for papillary thyroid cancer outcome. *J endocrinological Invest* (2017) 40(12):1355–63. doi: 10.1007/s40618-017-0714-y
18. Randolph GW, Duh QY, Heller KS, LiVolsi VA, Mandel SJ, Steward DL, et al. The prognostic significance of nodal metastases from papillary thyroid carcinoma can be stratified based on the size and number of metastatic lymph nodes, as well as the presence of extranodal extension. *Thyroid* (2012) 22(11):1144–52. doi: 10.1089/thy.2012.0043
19. Sapuppo G, Tavarelli M, Russo M, Malandrino P, Belfiore A, Vigneri R, et al. Lymph node location is a risk factor for papillary thyroid cancer-related death. *J endocrinological Invest* (2018) 41(11):1349–53. doi: 10.1007/s40618-018-0865-5
20. Tam S, Boonsripitayanon M, Amit M, Fellman BM, Li Y, Busaidy NL, et al. Survival in differentiated thyroid cancer: Comparing the AJCC cancer staging seventh and eighth editions. *Thyroid* (2018) 28(10):1301–10. doi: 10.1089/thy.2017.0572
21. Kim TH, Kim YN, Kim HI, Park SY, Choe JH, Kim JH, et al. Prognostic value of the eighth edition AJCC TNM classification for differentiated thyroid carcinoma. *Oral Oncol* (2017) 71:81–6. doi: 10.1016/j.oraloncology.2017.06.004
22. Kim M, Kim WG, Oh HS, Park S, Kwon H, Song DE, et al. Comparison of the seventh and eighth editions of the American joint committee on Cancer/Union for international cancer control tumor-Node-Metastasis staging system for differentiated thyroid cancer. *Thyroid* (2017) 27(9):1149–55. doi: 10.1089/thy.2017.0050
23. Liu G, Liu Q, Sun SR. Nomograms for estimating survival in patients with papillary thyroid cancer after surgery. *Cancer Manage Res* (2019) 11:3535–44. doi: 10.2147/CMAR.S194366
24. Lang BH, Wong CK, Yu HW, Lee KE. Postoperative nomogram for predicting disease-specific death and recurrence in papillary thyroid carcinoma. *Head neck*. (2016) 38(Suppl 1):E1256–63. doi: 10.1002/hed.24201
25. Pathak KA, Lambert P, Nason RW, Klonisch T. Comparing a thyroid prognostic nomogram to the existing staging systems for prediction risk of death from thyroid cancers. *Eur J Surg Oncol* (2016) 42(10):1491–6. doi: 10.1016/j.ejso.2016.05.016
26. Ge MH, Cao J, Wang JY, Huang YQ, Lan XB, Yu B, et al. Nomograms predicting disease-specific regional recurrence and distant recurrence of papillary thyroid carcinoma following partial or total thyroidectomy. *Medicine*. (2017) 96(30):e7575. doi: 10.1097/MD.00000000000007575