Title	Validation and comparison of nomograms in predicting disease-specific survival for papillary thyroid carcinoma
Author(s)	Lang, HHB; Wong, CKH
Citation	World Journal of Surgery, 2015, v. 39 n. 8, p. 1951-1958
Issued Date	2015
URL	http://hdl.handle.net/10722/209312
Rights	The final publication is available at Springer via http://dx.doi.org/10.1007/s00268-015-3044-2

Original Article

Validation and comparison of nomograms in predicting disease-specific survival for papillary thyroid carcinoma

Running head: The new nomogram predicts disease-specific survival

Brian Hung-Hin LANG[1], MS, FRACS

Carlos KH WONG[2], PhD

[1]Department of Surgery, The University of Hong Kong, Hong Kong SAR, China

[2]Department of Family Medicine and Primary Care, University of Hong Kong, 3/F Ap Lei Chau Clinic, 161 Main Street, Ap Lei Chau, Hong Kong

Address for Correspondence:

Dr Brian HH Lang

Division of Endocrine Surgery, Department of Surgery, Queen Mary Hospital, 102 Pokfulam Road, Hong Kong SAR, China. Tel.: (852) 22554232, Fax No.: (852) 28172291

Email: blang@hku.hk

Total text: 2772 words

Keywords: nomogram; thyroid carcinoma; thyroidectomy; disease-specific death; cancer death; predicting prognosis

ABSTRACT

Background:

Nomogram could estimate individualized prognosis in papillary thyroid carcinoma (PTC). We aimed to create and validate a new nomogram and compare it with other published nomograms using a large patient cohort.

Methods:

Eight-hundred and forty-nine PTC patients with ≥7 years follow-up were randomly assigned to the development (n=425) and validation (n=424) groups. The former was used for developing a nomogram for disease-specific survival (DSS) while the latter was for validating the nomogram by discrimination (or Area under Curve (AUC)). AUC of the newly-developed nomogram was compared to other published nomograms.

Results:

The 5-year and 10-year risk of dying from PTC were 1.4% and 3.3% respectively while dying from non-PTC-related causes were 2.3% and 5.1% respectively. The new nomogram was developed from age, tumor size, multifocality, nodal status and distant metastases. The discrimination was excellent (AUC (95%CI) for 5- and 10-year DSS were 0.896 (0.683-0.971) and 0.919 (0.871-0.967), respectively). Its predictability was similar to other published nomograms (p>0.05). Based on the new nomogram, a total score of < 28 meant 99.72% chance of surviving from PTC at 10 years while a score of ≥28 meant 9.09% chance of dying from PTC at 10-years.

Conclusions:

Using variables from the current *TNM* staging system, a new nomogram was developed. It exhibited excellent discriminatory ability and accuracy in predicting 10-year DSS relative to other published nomograms. However, given the excellent prognosis of PTC, the new nomogram was better at ruling out than predicting PTC-related death. Further validation by an external cohort is required.

INTRODUCTION

Differentiated thyroid carcinoma is the most common type of thyroid carcinoma and its incidence has doubled over the past two decades[1,2]. Although it is generally associated with an excellent outcome, its prognosis depends on the presence of certain clinicopathologic characteristics[3,4]. Numerous staging systems have been available for stratifying risk in papillary (PTC) and follicular thyroid carcinoma and the American Joint Committee on Cancer (AJCC) Tumor Node Metastasis *(TNM)* staging system has been the most accurate[3,4]. However, although it works well for a patient population, it is less useful in predicting prognosis for an individual patient. Nomogram is a powerful clinical tool designed in predicting outcome of an individual patient by utilizing multiple clinical variables[5,6]. A well-balanced nomogram could provide individualized estimation of prognosis and help counselling cancer patients' management selection and optimizing therapeutic approaches[5,6]. To PTC, this could potentially mean tailoring the extent of surgery, decision on radioiodine (RAI) ablation afterwards and follow-up intensity[3,4]. To our knowledge, there have been few nomograms reported in the literature designed for thyroid cancers[7-9]. Furthermore, given the fact that most were developed from population-based cohorts, tumor variables described appeared less conventional and therefore, were more difficult to apply clinically than those used in the existing *TNM* staging or other risk-stratification systems[10]. Furthermore, since they were developed from the Caucasians population, they may be less applicable to our predominantly ethnic-Chinese population. Given these issues, we aimed to develop and validate a new nomogram and then compare it with other nomograms published in the literature using a large patient cohort from at our institution. Since we anticipated that a considerable number of patients may die from non-PTC related causes, a competing-risk model was used for analysis.

PATIENTS AND METHODS

Patient cohort

From 1970 - 2006, 1038 consecutive patients with PTC underwent total or near-total thyroidectomy in our institution. Of these, 35 (3.4%) with incomplete patient data and 154 (14.8%) with occult microcarcinoma (<1cm) were excluded. Therefore, 849 (81.8%) were included. All patients had a minimum follow-up of 7 years and were managed by a standardized protocol described previously[11]. In brief, total or near-total thyroidectomy was preferred. Simultaneous therapeutic central (level VI) +/- lateral (levels II-V) selective neck dissection was performed for clinically-proven nodal metastasis. Prophylactic central neck dissection was not routinely performed. Decision for RAI ablation was based on factors such as tumor size > 1.5cm, nodal metastasis, age >45 years old, extrathyroidal extension, residual disease and distant metastasis. To ensure updated patient status, a careful search in the territory-wide Clinical Management System (CMS) was performed on all 849 patients. The CMS links up all public hospitals and covers over 90% of all inpatient medical records in the whole territory. The latest date of follow-up or date of death and the cause of death were retrieved from the CMS. All causes of death were later confirmed by examination of the medical record, autopsy report and / or death certificate. Disease-specific survival (DSS) was defined as the time from PTC diagnosis to the date when the patient died of PTC. All relevant clinicopathological and perioperative data were collected prospectively after 1994 and regularly updated in a computerized database. The present study protocol was approved by the local institutional review board.

Development, validation and comparison of the new nomogram

Our patient cohort was randomly divided into two groups, namely the development set (n=425) and the validation (n=424) set. The former was used to develop a risk algorithm or nomogram for 5- and 10-year DSS using a proportional sub-distribution hazards model. This model accounted for the competing-risk of dying from non-PTC related causes. The validation set was used for testing out the newly-developed nomogram by discrimination and calibration. Lastly, the performance of the newly-developed nomogram was compared with other published nomograms in the literature using the validation set.

Published nomograms and their interpretation

Using the search terms ("nomogram", "predictive model" and "thyroid cancer") at the Scopus and Medline (PubMed) electronic databases, three independent nomograms were found[7-9] but since only two were applicable to PTC, only two were selected for comparison. The first nomogram was reported by Yang *et al*[8] using the Surveillance, Epidemiology and End Results (SEER) data program. They found age (years), tumor size (mm), extent of primary tumor (localized, regional or distant), lymph node involvement (none, regional or distant) and histology (papillary, follicular, medullary and anaplastic) to be significant independent variables. For this study, we assumed "distant extent of primary tumor" and "distant lymph node involvement" were equivalent to distant metastases by the *TNM* system. The second nomogram was reported by Pathak *et al.*[7] using a large patient cohort from the Manitoba (Canada) Cancer Registry. Age (years), sex (male or female), histology (papillary, non-papillary or medullary), distant metastases (yes or no), tumor stage (T1, T2, T3 or T4) and post-treatment residual disease (yes or no) were significant independent variables. Since our validation set comprised PTC only, the score on "histology" in both nomograms was given a score of zero. However, since both papers did not provide a more precise numeric cut-off value for each independent variable, the first

and/or corresponding authors were contacted directly and asked for more detailed information regarding their nomogram before analysis.

Statistical analysis

Descriptive statistics of baseline characteristics between development and validation sets were compared using independent *t*-test for continuous variables and Chi-square test for categorical variables.

Using the development set, we established a multivariable competing-risk subhazard models for predicting risk of dying from PTC, account for the competing-risks of dying from non-PTC-related causes based on Fine and Gray modelling approach[12]. Significant risk factors with $p<0.10$ in stepwise competing-risk subhazard model were retained in the point system. Each risk factor was assigned a weighting in the point system using the respective β-coefficients multiplied by 10 and rounded to the nearest whole number. The risk score for each subject is the sum of risk score contributed by each risk factor identified by the final point system. A final simple point system was constructed in scoring and predicting an individual 5-year and 10-year risk dying from PTC. Youden's index was used to determine the optimal cutoff value in predicting of the development of DSS in the next 5 or 10 years.

Data from the validation set were entered into each risk algorithm / nomogram to produce area under the receiver-operating characteristic curve (AUC). This was a measure of discrimination. The AUCs were compared using the method of Delong test[13]. A calibration plot was applied to plot the observed probabilities of PTC-related death against the nomogram-predicted probabilities of PTC-related death. Accuracy was also measured by means of sensitivity, specificity, positive predictive value and negative predictive value using the optimal cut-off. All statistical analyses were performed using SPSS version 18.0 (SPSS, Inc., Chicago, IL, USA) and

the STATA software version 13 (STATA Corp, College Station, Texas).

RESULTS

Our patient cohort comprised mostly women (79.4%) and ethnic Chinese (91.8%). The mean age was 45.6 ± 16.8 years while the mean tumor size was 2.2 ± 1.7cm. There were 210 (24.7%) patients who presented with clinically palpable cervical nodal metastases and 26 (3.0%) who presented with distant metastases. After a median follow-up of 17.1 (7.2 – 47.2) years, 61 (7.2%) patients died of PTC, 47 (5.5%) patients died of non-thyroidal malignancy and 73 (8.6%) died of a non-cancer (i.e. medical or natural) cause. The 5-year, 10-year and 15-year risks of dying from PTC were 1.4%, 3.3% and 4.4%, respectively while from non-PTC-related cause were 2.3%, 5.1% and 8.5%, respectively. At the time of analysis, 30 (3.5%) patients were still alive with detectable local and/or distant diseases while 64 (7.5%) patients were reverted to "disease-free" after treatment.

Table 1 shows a comparison of clinicopathological factors and patient status between the development and validation sets. Essentially, age at diagnosis, sex, tumor size, frequency of extrathyroidal extension, nodal status, distant metastases, completion of resection and RAI were comparable between the two sets. However, tumor multifocality was significantly more frequent in the development set than the validation set (35.5% vs. 27.6%, p=0.013). The number of PTC-related deaths in development and validation sets was 34 (8.0%) and 27 (6.4%), respectively. The 10-year risk of dying from PTC was comparable between the two sets (4.2% vs. 3.3%, p=0.367).

Table 2 shows the significant prognostic risk factors and their assigned score for 5- and 10-year DSS. Male sex, extrathyroidal extension, completion of resection and RAI were not significant factors and therefore, not shown. Only age at diagnosis, tumor size, tumor multifocality (defined ≥2 tumor foci in ipsilateral lobe), nodal status and distant metastases were significant

independent factors for both 5- and 10-year DSS. Age at diagnosis <25 years, primary tumor size <2.0cm, absence of multifocality, no nodal metastases (N0/Nx) and no distant metastases were given a risk score of zero while others were given a score ranging between 2 – 21. Therefore, the maximum total score for the 5-year and 10-year DSS nomogram were 60 and 62, respectively. Table 3 shows the respected accuracy and discrimination of the three nomograms when the development and validation sets were entered. Using data from the development set, the new nomogram had an AUC (95%CI) in predicting 5- and 10-year DSS of 0.900 (0.846-0.954) and 0.896 (0.827-0.965), respectively while when the validation set was used, the discriminatory ability was maintained at a similar level (AUC (95%CI) for 5- and 10-year DSS were 0.896 (0.822 – 0.971) and 0.919 (0.871-0.967), respectively). This demonstrated the new nomogram produced good discrimination in an independent set and confirmed its potential for exportability. AUCs for predicting 10-year DSS in Yang's nomogram and Pathak's nomogram were also excellent (0.892 and 0.859, respectively), although they appeared lower than that of the new nomogram (0.919). Nevertheless, in pairwise comparison of ROC curves, the difference between the new nomogram and Yang's (p=0.462), between the new nomogram and Pathak's (p=0.220) or between Yang's and Pathak's (p=0.616) were not statistically significant. Using the respective optimal cut-off value, each nomogram produced its best sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV). The new nomogram appeared to have comparable sensitivity, specificity, PPV and NPV as the other two published nomograms. One finding worth noting was that all the NPVs were significantly higher than PPVs and this implied the fact that all nomograms were better at ruling out than predicting PTC-related death. Based on the new nomogram, for a total score of 28 or less, the overall chance of surviving from

PTC at 10 years was 99.72% while if a total score was 28 or more, the overall chance of dying from PTC at 10-years was 9.09%.

Figure 1 shows the calibration plot for predicting 5-year and 10-year DSS. Since all the points were close to the 45-degree line, this indicated good agreement between the predicted and observed 5-year and 10-years DSS.

Figure 2a shows the newly-constructed nomogram for predicting 5-year risk of dying from PTC and Figure 2b shows the nomogram for predicting 10-year risk of dying from PTC using the development set (n=425). Each parameter has been assigned a score and the sum of these scores provides a predicted risk of dying from PTC in 5 or 10 years ranging 0.1% to 50%.

DISCUSSION

Our data showed that the 10-year probabilities of dying from PTC in the development and validation sets were 4.2% and 2.3%, respectively (or an overall rate of 3.3%). These rates were similar to those reported from the two previous population-based cohort studies[7,8]. In the SEER database[8], the 10-year probability of dying from thyroid cancer was 3.0% while in the Manitoban program, the 10-year probability of dying from PTC was 3.2%[7]. The overall 10-year probability of dying from non-PTC related causes was also similar to that reported from the SEER data (5.1% vs.5.9%, respectively)[8], although such data was not available from the Manitoban program [7]. Therefore, despite being a single institutional study, the overall probabilities of dying from PTC and non-PTC-related causes were remarkably similar to those reported from population-based studies. Secondly, over time, the probability of dying from non-PTC related causes was roughly twice that of dying from PTC itself[8]. The latter finding highlighted the importance and the rationale of using a competing-risk model in the development of a nomogram.

One point worth noting regarding the new and the other two nomograms was that because of the excellent prognosis of PTC with few patients dying from PTC after a long period of time, it was significantly easier for either of the 3 nomograms to rule out PTC-related death than to predict it. For example, using the new nomogram, if a patient had an aggregate score of 28 or less, the overall chance of not dying from PTC was close to 100% while the chance of dying from PTC if the aggregate score was 28 or more was only 9.09% (or roughly 1 in 10). Therefore, despite using an optimal cut-off, our new nomogram could still only correctly predict a patient with a higher score may die in the future in less than 10% of the time while it could predict with certainty that a patient with a lower score will not die from PTC.

Relative to the other two nomograms,[7,8] we believe we have developed a nomogram which is more practical and friendlier for clinicians to use. This is because most of the prognostic variables (such as age, tumor size, nodal status and distant metastases) incorporated in the new nomogram appeared similar to the variables used in the existing AJCC *TNM* staging and other risk-stratification systems[3] and therefore, it would be easier to apply in actual clinical practice. However, unlike the *TNM* staging system, the new nomogram provides a better prediction for individual patients, based on statistical modeling. Interestingly, the presence of tumor multifocality turned out to be a significant independent factor in the new nomogram but not considered in the current *TNM* staging system or other two published nomograms. One possible reason for this discrepancy might have been because tumor multifocality is often accurately not recorded or coded in population-based cancer registries and therefore, not available for evaluation in most population-based studies. However, it is worth highlighting that numerous studies have found tumor multifocality to be an important prognostic factor of PTC-related death in clinically-significant (> 1cm) PTC[14-16].

Residual disease after thyroidectomy was found to be an important prognostic variable in one nomogram[7] but not in ours and the other nomogram. Although some existing stratification systems also considered residual disease to be an important prognostic factor[3,4], the variable "completeness of resection" did not turn out significant in our multivariate analysis. One possible reason might have been because this factor was largely based on the surgeon's and pathologist's judgement and so, it may vary between cases making this variable a less reliable factor.

In terms of nomogram performance, although not statistically significant, our model appeared to have better discrimination and accuracy in predicting 10-year DSS than the other two models[7,8].

However, this should not be a surprise as the independent data set (or the validation set) used was essentially derived from the same patient cohort as the development set and therefore, it was not truly externally validated. Nevertheless, given that our nomogram might be easier to apply clinically than other two nomograms, our nomogram might be a better clinical tool for predicting PTC-related survival.

Despite these findings, we acknowledged certain shortcomings. Since this was a single institution study, it may have insufficient power to identify some true prognostic variables. Also unlike other population-based analyses, it was subjected to institutional and referral biases. Nevertheless, we believe our data are of better quality than many population-based databases as data were updated prospectively and therefore the chance of misdiagnosis and death misclassification was small. We also acknowledge the fact that even though our nomogram was validated by a separate cohort, this cohort came from the same population and therefore, strictly speaking, our nomogram was not externally validated. Nevertheless, using an external cohort (i.e. patients managed outside our institution) in validating ours and other two nomograms could be done in the future. Also we would acknowledge the fact that our nomogram did not predict relapses which may arguably be a clinically more-relevant outcome than cancer-related deaths[10]. However, since response to initial therapy is an important factor in evaluating relapses, many patients from the earlier period might not be included for such analysis[17]. Lastly, since this nomogram was purely based on clinical factors, perhaps the addition of molecular markers may further improve the prediction accuracy[18,19].

Conclusion

Over time, almost twice as many patients with PTC died of non-PTC-related causes than of PTC alone. Using variables similar to those in current *TNM* staging system, we have developed a new nomogram using more clinically refined and applicable variables in predicting 5- and 10-year DSS. Relative to the other two published nomograms, the new nomogram had excellent discriminatory ability and accuracy in predicting 10-year DSS. However, due to the excellent prognosis of PTC, the new nomogram was better at ruling out than predicting PTC-related death. Perhaps, future studies could validate this new nomogram with an external cohort.

ACKNOWLEDGEMENTS

We like to acknowledge Professor K. Alok Pathak from the University of Manitoba, Canada and Dr. Limin Yang from National Research Institute for Child Health and Development, Japan for generously providing a more detailed scoring algorithm of their respected nomogram.

DISCLOSURE STATEMENT

All authors had nothing to disclose. No competing financial interests exist.

REFERENECES

1. Cancer incidence and mortality in Hong Kong 1983-2012. Hong Kong Cancer Registry, Hong Kong. Available: http://www3.ha.org.hk/cancereg/e_stat.asp [Accessed on 9th December 2014]

2. Morris LG, Sikora AG, Tosteson TD, Davies L. (2013) The increasing incidence of thyroid cancer: the influence of access to care. Thyroid. 23(7):885-91.

3. Lang BH, Chow SM, Lo CY, Law SC, Lam KY. (2007) Staging systems for papillary thyroid carcinoma: a study of 2 tertiary referral centers. Ann Surg. 246(1):114-21.

4. Lang BH, Lo CY, Chan WF, Lam KY, Wan KY. (2007) Staging systems for follicular thyroid carcinoma: application to 171 consecutive patients treated in a tertiary referral centre. Endocr Relat Cancer. 14(1):29-42. Review

5. Kattan MW, Scardino PT.(2007) Evidence for the usefulness of nomograms. Nat Clin Pract Urol. 4(12):638-9.

6. Touijer K, Scardino PT. (2009) Nomograms for staging, prognosis, and predicting treatment outcomes. Cancer. 115(13 Suppl):3107-11. doi: 10.1002/cncr.24352.

7. Pathak KA, Mazurat A, Lambert P, Klonisch T, Nason RW. (2013) Prognostic nomograms to predict oncological outcome of thyroid cancers. J Clin Endocrinol Metab. 98(12):4768-75.

8. Yang L, Shen W, Sakamoto N. (2013) Population-based study evaluating and predicting the probability of death resulting from thyroid cancer and other causes among patients with thyroid cancer. J Clin Oncol. 31(4):468-74.

9. Ho AS, Wang L, Palmer FL et al. (2014) Postoperative Nomogram for Predicting Cancer-Specific Mortality in Medullary Thyroid Cancer. Ann Surg Oncol. 2014 Nov 4. DOI: 10.1245/s10434-014-4208-2
10. Wang TS, Sosa JA. (2013) Thyroid gland: Can a nomogram predict death in patients with thyroid cancer? Nat Rev Endocrinol. 9(4):192-3.
11. Lang BH, Lo CY, Wong KP, Wan KY. (2014) Long-Term Outcomes for Older Patients with Papillary Thyroid Carcinoma: Should Another Age Cutoff Beyond 45 Years Be Added? Ann Surg Oncol. 2014 Sep 5. DOI: 10.1245/s10434-014-4055-1
12. Fine JP, Gray RJ. (1999) A proportional hazards model for the subdistribution of a competing risk. J Am Stat Assoc 94:496–509.
13. DeLong ER, DeLong DM, Clarke-Pearson DL. (1988) Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics. 44(3):837-45.
14. Grogan RH, Kaplan SP, Cao H, Weiss RE, Degroot LJ, Simon CA, Embia OM, Angelos P, Kaplan EL, Schechter RB. (2013) A study of recurrence and death from papillary thyroid cancer with 27 years of median follow-up. Surgery. 154(6):1436-46; discussion 1446-7.
15. Kim KJ, Kim SM, Lee YS, Chung WY, Chang HS, Park CS. (2014) Prognostic Significance of Tumor Multifocality in Papillary Thyroid Carcinoma and its Relationship with Primary Tumor Size: A Retrospective Study of 2,309 Consecutive Patients. Ann Surg Oncol. 2014 Aug 5. DOI: 10.1245/s10434-014-3899-8

16. Qu H, Sun GR, Liu Y, He QS. (2014) Clinical risk factors for central lymph node metastasis in papillary thyroid carcinoma: a systematic review and meta-analysis. Clin Endocrinol (Oxf). 2014 Aug 16. doi: 10.1111/cen.12583.
17. Vaisman F, Momesso D, Bulzico DA, Pessoa CH, Dias F, Corbo R, Vaisman M, Tuttle RM. (2012) Spontaneous remission in thyroid cancer patients after biochemical incomplete response to initial therapy. Clin Endocrinol (Oxf). 77(1):132-8.
18. Yu XM, Lo CY, Lam AK et al. (2008) The potential clinical relevance of serum vascular endothelial growth factor (VEGF) and VEGF-C in recurrent papillary thyroid carcinoma. Surgery. 144(6):934-40; discussion 940-1
19. Lang BH, Chai YJ, Cowling BJ et al. (2014) Is BRAFV600E mutation a marker for central nodal metastasis in small papillary thyroid carcinoma? Endocr Relat Cancer. 21(2):285-95

Table 1. A comparison of clinicopathological factors and patient status between the development and validation sets

	Development set (n=425)	**Validation set (n=424)**	**Entire cohort (n=849)**	***p*-value***
Age at diagnosis (years)				0.799
- < 25	42 (9.9)	45 (10.6)	87 (10.2)	
- ≥ 25 and < 45	183 (43.1)	172 (40.6)	355 (41.8)	
- ≥45 – 65	128 (30.1)	139 (32.8)	267 (31.4)	
- > 65	72 (16.9)	68 (16.0)	140 (16.5)	
Sex				0.564
- Male	91 (21.4)	84 (19.8)	175 (20.6)	
- Female	334 (78.6)	340 (80.2)	674 (79.4)	
Tumor size (cm)				0.308
- < 2.0	208 (48.9)	222 (52.3)	430 (50.6)	
- ≥ 2.0 and < 4.0	134 (31.5)	136 (32.1)	270 (31.8)	
- ≥ 4.0	83 (19.5)	66 (15.6)	149 (17.6)	
Tumor multifocality				**0.013**
- Absent	274 (64.5)	307 (72.4)	581 (68.4)	

- Present	151 (35.5)	117 (27.6)	268 (31.6)	
Extrathyroidal extension				0.138
- Absent	278 (65.4)	298 (70.3)	576 (67.8)	
- Present	147 (34.6)	126 (29.7)	273 (32.2)	
Nodal status#				0.176
- None (N0/Nx)	254 (59.8)	278 (65.6)	532 (62.7)	
- N1a	67 (15.8)	52 (12.3)	119 (14.0)	
- N1b	104 (24.5)	94 (22.2)	198 (23.3)	
Distant metastases				0.995
- Absent	412 (96.9)	411 (96.9)	823 (96.9)	
- Present	13 (3.1)	13 (3.1)	26 (3.1)	
Completeness of resection				0.644
- Complete	400 (94.1)	396 (93.4)	796 (93.8)	
- Incomplete	25 (5.9)	28 (6.6)	53 (6.2)	
Radioiodine ablation				0.166
- Not given	161 (37.9)	186 (43.9)	347 (40.9)	
- Given	264 (62.1)	238 (56.1)	502 (59.1)	

Patient status at analysis				0.714
- Alive without recurrence	312 (73.4)	323 (76.2)	571 (67.3)	
- Alive with recurrence	16 (3.8)	14 (3.3)	30 (3.5)	
- Alive with other malignancy	2 (0.5)	1 (0.2)	3 (0.4)	
- Died of PTC	34 (8.0)	27 (6.4)	61 (7.2)	
- Died of other malignancy	25 (5.9)	22 (5.2)	47 (5.5)	
- Died of medical disease	36 (8.5)	37 (8.7)	73 (8.6)	

Categorical variables are expressed as number (percentage)

PTC = papillary thyroid carcinoma

*comparison between development and validation sets

#by 7th edition *TNM* staging system

Table 2. Prognostic risk factors and their assigned score influencing 5- and 10-year disease-specific survivals in the proportional sub-distribution hazard competing risks model

	5-year disease-specific survival					**10-year disease-specific survival**				
	SHR	**95% CI**	***p*-value**	**ß-coeff.**	**Assigned Score**	**SHR**	**95% CI**	***p*-value**	**ß-coeff.**	**Assigned Score**
Age at diagnosis (yrs)										
- < 25	1	-	-	-	**0**	1	-	-	-	**0**
- ≥ 25 and < 45	1.244	0.200–7.724	0.815	0.218	**2**	1.284	0.213-7.754	0.785	0.249	**2**
- ≥45 – 65	5.116	0.995–26.320	0.051	1.632	**16**	5.617	1.103-28.597	0.038	1.726	**17**
- > 65	8.088	1.540-42.485	0.014	2.090	**21**	8.515	1.651-43.917	0.011	2.142	**21**
Tumor size (cm)										
- < 2.0	1	-	-	-	**0**	1	-	-	-	**0**
- ≥ 2.0 and < 4.0	1.798	0.693-4.663	0.227	0.587	**6**	1.730	0.657-4.555	0.267	0.548	**5**
- ≥ 4.0	2.488	0.999-6.193	0.050	0.911	**9**	2.303	0.856-6.196	0.099	0.834	**8**
Tumor multifocality										
- Absent	1	-	-	-	**0**	1	-	-	-	**0**
- Present	2.030	0.994-4.142	0.052	0.708	**7**	2.010	0.920-4.393	0.080	0.698	**7**
Nodal status#										
- None (N0/Nx)	1	-	-	-	**0**	1	-	-	-	**0**

- N1a	1.760	0.502-6.169	0.377	0.565	**6**	1.752	0.489-6.278	0.389	0.561	**6**
- N1b	3.114	1.404-6.904	0.005	1.136	**11**	3.328	1.483-7.468	0.004	1.202	**12**
Distant metastases										
- Absent	1	-	-	-	**0**	1	-	-	-	**0**
- Present	3.485	1.347-9.013	0.010	1.248	**12**	4.090	1.313-12.745	0.015	1.409	**14**

Abbreviation: SHR = Subhazard ratio; ß-coeff. = ß-coefficient

#based on 7th edition *TNM* staging system

Table 3. A comparison of discrimination (as measured by area under the curve (AUC)) and accuracy between nomograms using the development set and validation set

Risk algorithms	**Area under the curve (AUC) (95%CI)**	**Optimal cutoff value**	**Sensitivity (%)**	**Specificity (%)**	**Positive predictive value (%)**	**Negative predictive value (%)**
	Applying to the development set (n=425)					
Our model						
- 5-year DSS	0.900 (0.846-0.954)	30	100.0	79.14	7.77	100.0
- 10-year DSS	0.896 (0.827-0.965)	28	94.12	78.02	13.01	99.67
	Applying to the validation set (n=424)					
Our model						
- 5-year DSS	0.896 (0.822-0.971)	30	75.00	84.29	4.35	99.72
- 10-year DSS	0.919 (0.871-0.967)	28	88.89	80.34	9.09	99.72
Yang *et al.* [8]+						
- 10-year DSS	0.892 (0.822-0.962)	101.6	88.89	76.90	7.84	99.69
Pathak *et al.* [7]*						
- 10-year DSS	0.859 (0.757-0.961)	75	77.78	82.56	8.97	99.42

+the model was based on variables such as age, tumor size, extent of primary tumor (localized, regional or distant) and lymph node involvement (none, regional or distant)

*the model was based on variables such as age, sex, presence of distant metastases, tumor stage (T1 to T4) and post-treatment residual tumor

LEGENDS

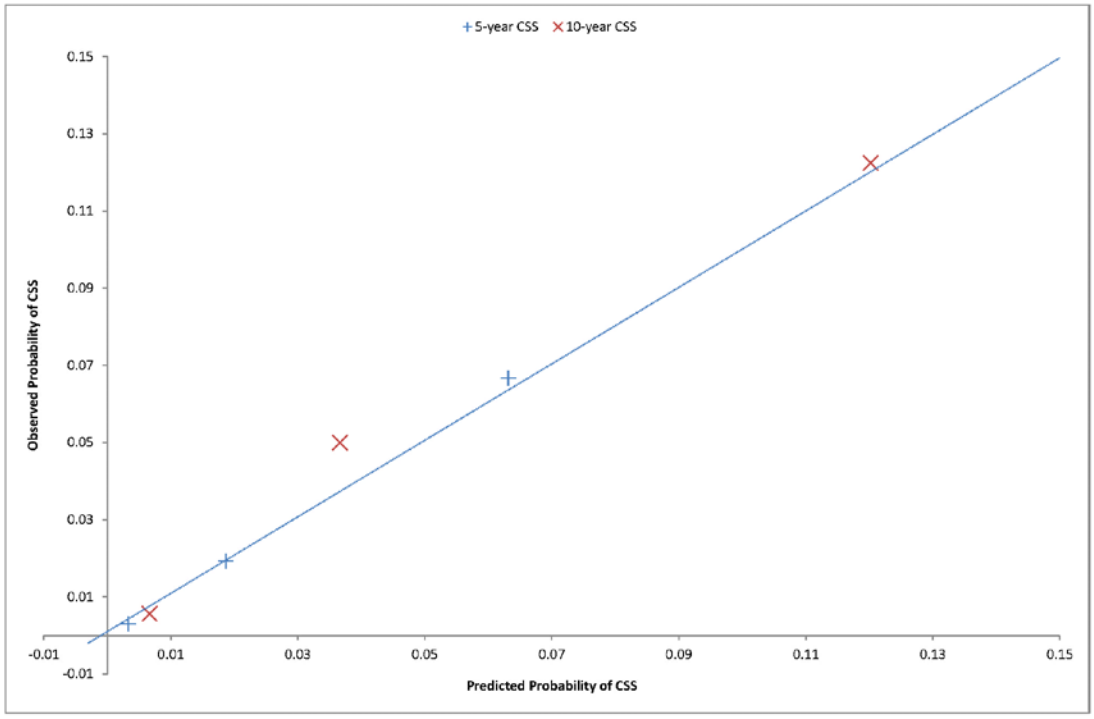

Figure 1. The calibration curve for predicting 5-year and 10-year disease-specific survival.

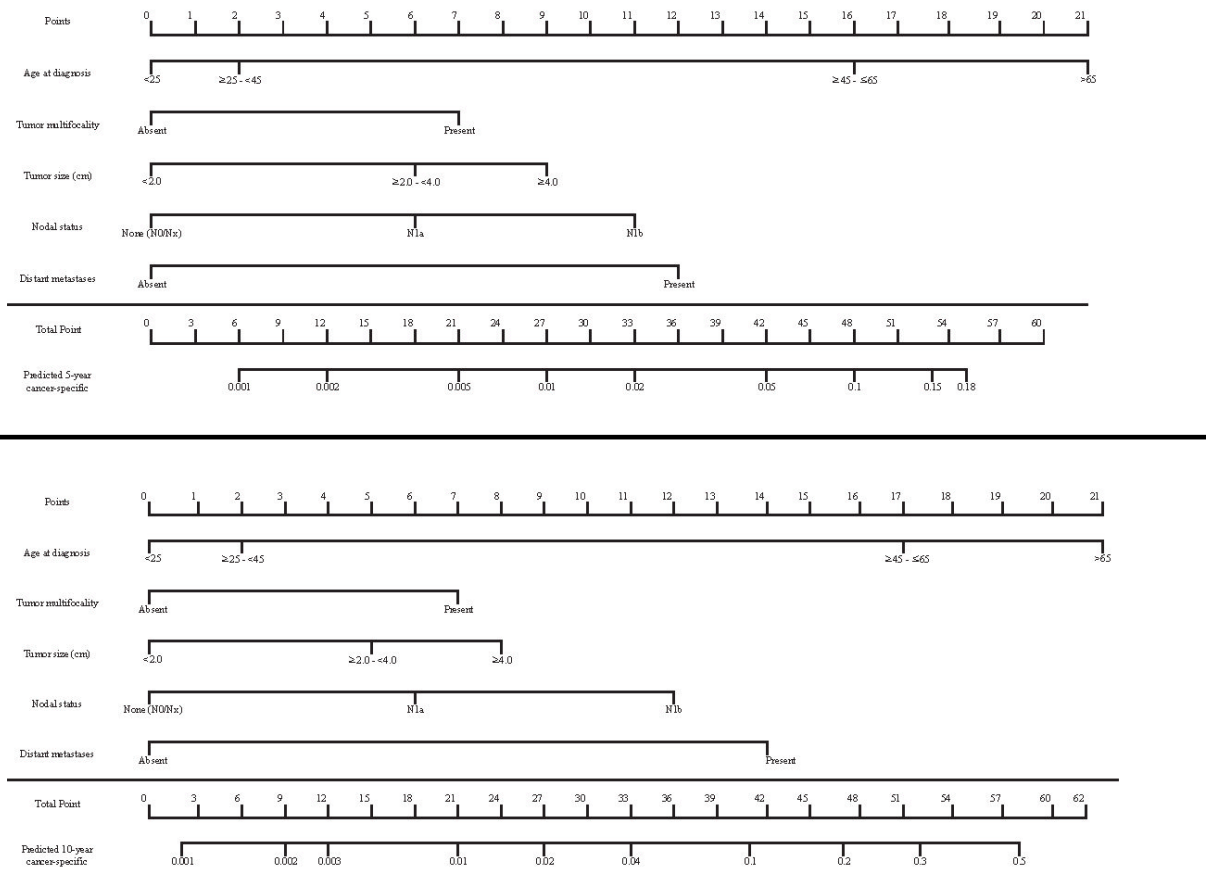

Figure 2. The two new prognostic nomogram for predicting 5-year and 10-year risk of dying from papillary thyroid carcinoma