Modified dynamic risk stratification for predicting recurrence using the response to initial therapy in patients with differentiated thyroid carcinoma

Min Ji Jeon¹, Won Gu Kim¹, Woo Ri Park¹, Ji Min Han¹, Tae Yong Kim¹, Dong Eun Song³, Ki-Wook Chung², Jin-Sook Ryu⁴, Suck Joon Hong², Young Kee Shong¹ and Won Bae Kim¹

Departments of ¹Internal Medicine, ²Surgery, ³Pathology and ⁴Nuclear Medicine, Asan Medical Center, University of Ulsan College of Medicine, 88, Olympic-ro 43-gil, Songpa-gu, Seoul 138-736, Korea

Correspondence should be addressed to W B Kim
Email
kimwb@amc.seoul.kr

Abstract

Objective: A new risk stratification system was proposed to estimate the risk of recurrence in patients with differentiated thyroid carcinoma (DTC) using the response to initial therapy. Here, we describe the modified dynamic risk stratification system, which takes into consideration the status of serum anti-Tg antibody (TgAb), and validate this system for assessing the risk of recurrence in patients with DTC.

Patients and methods: Patients who underwent total thyroidectomy with radioiodine remnant ablation due to DTC between 2000 and 2005 were included. We classified patients into four groups based on the response to the initial therapy ('excellent', 'acceptable', 'biochemical incomplete', and 'structural incomplete' response).

Results: The median follow-up period of 715 patients with DTC was 8 years. The response to initial therapy was an important risk predictor for recurrent/persistent DTC. The relative risks (95% CI) of recurrence were 16.5 (6.3–43.0) in the 'acceptable response' group, 41.3 (15.4–110.8) in the 'biochemical incomplete response' group, and 281.2 (112.9–700.5) in the 'structural incomplete response' group compared with the 'excellent response' group ($P < 0.001$, $P < 0.001$, and $P < 0.001$ respectively). The disease-free survival rate of the 'excellent response' group to initial therapy was 98.3% whereas that of the 'structural incomplete response' group was only 6.8%.

Conclusions: Our study validates the usefulness of the modified dynamic risk stratification system including the status of serum TgAb for predicting recurrent/persistent disease in patients with DTC. Personalized risk assessment using the response to initial therapy could be useful for the follow-up and management of patients with DTC.

European Journal of
Endocrinology
(2014) 170, 23–30

Introduction

The practical and effective risk stratification system is essential for the individualized assessment and management of patients with differentiated thyroid carcinoma (DTC). Over the last decades, a number of studies have identified various clinicopathological risk factors and suggested staging systems to distinguish low-risk patients from high-risk patients (1, 2, 3, 4, 5, 6). The survival rate for patients with DTC is usually excellent and only a small

percentage of patients die due to DTC (7, 8). A recent study estimated the 10-year probability of death resulting from thyroid cancer to be 3% (9). Therefore, the risk stratification system predicting the recurrence of DTC could be more practical for clinicians to identify high-risk patients and to provide more personalized management of the disease.

The American Joint Committee on Cancer (AJCC)/International Union against Cancer (UICC) Tumor Node

Metastasis (TNM) staging system is commonly used in clinical practice and to predict the risk of disease-related death (6). However, the TNM staging system is not appropriate for assessing the risk of DTC recurrence (1, 10). To improve prognostication and follow-up management for patients with DTC, the American Thyroid Association (ATA) recommended the use of a new post-operative clinicopathological staging system (1). The ATA risk stratification system is a comprehensive approach that includes postoperative pathology analysis, radioiodine therapy, and serum thyroglobulinemia. However, the ATA risk stratification system neither defines the specific thyroglobulin (Tg) level nor describes serial ultrasonography (US) results for assessment of the risk of recurrence. Moreover, the risk of recurrence and death may change over time depending on the response to therapy and the clinical course of the disease.

Recently, the dynamic risk stratification system, which reassesses the risk of recurrence using the level of serum Tg and results of follow-up diagnostic images as markers of the response to initial surgery and radioactive iodine (RAI) remnant ablation, was described (11, 12). This system classified patients into four categories by disease status during the first 6–24 months after initial treatment. This practical approach effectively estimates the risk of recurrent disease in DTC patients. This system has been validated in three large cohort studies in the United States and Brazil (11, 12, 13). However, this dynamic risk assessment did not take into consideration the status of serum anti-Tg antibody (TgAb) even though 10–25% of patients with DTC have a positive serum TgAb level (14). Delayed risk stratification, proposed by Castagna *et al.* (15), also did not include patients with a positive serum TgAb level.

In this study, we describe the modified dynamic risk stratification system, which includes the level of serum TgAb, and analyze this system for assessing the risk of recurrence/persistence in patients with DTC at a single tertiary referral hospital in Korea. We also evaluated the relative importance of this system for predicting prognosis compared with TNM staging and ATA risk classification.

Subjects and methods

Patients

From 2000 to 2005, 1296 patients with DTC who underwent first surgery at the Asan Medical Center (Seoul, Korea) were retrieved from a historical cohort. Patients were treated by total or near total thyroidectomy with subsequent RAI ablation therapy according to a protocol

established by the Endocrinology Division of the Asan Medical Center (Seoul, Korea) (16). Patients with anaplastic, poorly differentiated, or medullary thyroid carcinoma were excluded. Patients with a tumor size of <1 cm ($n=366$) or without adequate follow-up data for determining dynamic risk stratification ($n=215$) were also excluded. A total of 715 patients were eligible for this study. This study was approved by the Institutional Review Board of the Asan Medical Center.

Follow-up protocol to detect recurrence

After remnant ablation, all patients were prescribed levothyroxine for TSH suppression and were regularly followed-up with physical examination. The level of serum Tg and serum TgAb was measured every 6–12 months. The level of serum-stimulated Tg after thyroid hormone withdrawal was measured within 2 years after remnant ablation. A TgAb value exceeding 60 IU/ml was considered to be positive for interfering with the Tg measurement. Neck US and diagnostic radioiodine whole-body scan (WBS) were done during the first 6–24 months after initial therapy and repeated at 12- to 24-month intervals. Additional diagnostic imaging studies, such as computed tomography, magnetic resonance scan, or whole-body fluorodeoxyglucose (FDG)-positron emission tomography (PET) scan, were also performed in some patients as needed.

Definitions of clinical outcomes

Patients with an undetectable level of serum-stimulated Tg (<1 ng/ml), negative TgAb, and no suspicious structural disease at the end of the follow-up period were considered to have no clinical evidence of disease (NED). Patients were considered to have persistent disease if the suppressed or stimulated Tg level was ≥ 1 ng/ml or if suspicious structural disease was seen on images. The recurrence was defined as a newly developed biochemical (suppressed or stimulated Tg level ≥ 1 ng/ml) or structural disease on images after the NED period. Patients with persistent disease were classified as having either biochemical persistent disease or structural persistent disease. Biochemical persistent disease was defined as a persistently elevated suppressed or stimulated Tg level ≥ 1 ng/ml without structural evidence of disease. Structural persistent disease was defined as pathologically or cytologically proven malignant tissue and/or highly suspicious metastatic lesions by imaging studies such as neck US, WBS, or FDG-PET scan.

Figure 1

Disease-free survival (DFS) curves according to the serum Tg level and the status of serum TgAb. These preliminary data included 930 DTC patients who underwent first thyroid surgery between 2000 and 2005 at the Asan Medical Center and whose tumor size was ≥ 1 cm. Patients with positive TgAb (≥ 60 IU/ml) show DFS rates similar to those with serum-stimulated Tg level 1–10 ng/ml.

Risk stratification by each classification

We categorized eligible patients according to the sixth edition of the AJCC TNM staging system (stages I, II, III, and IV) and the ATA risk classification (low, intermediate, and high risk of recurrence) (1). We also classified patients by the response to initial therapy using the levels of serum-stimulated Tg and TgAb and the results of neck US and diagnostic WBS 6–24 months after initial therapy.

An ‘excellent response’ was defined as a negative serum-stimulated Tg level (<1 ng/ml) with the absence of serum TgAb (<60 IU/ml) and no suspicious metastatic lesions by imaging. An ‘acceptable response’ was defined

as negative findings by imaging with a low serum-stimulated Tg level (1–10 ng/ml) or the presence of serum TgAb (≥ 60 IU/ml) because patients with positive serum TgAb showed a similar disease-free survival (DFS) rate as patients with a low serum-stimulated Tg level (1–10 ng/ml) (Fig. 1). A ‘biochemical incomplete response’ was defined as a high serum-stimulated Tg level (≥ 10 ng/ml) without any evidence of structural metastatic lesions by imaging. A ‘structural incomplete response’ was defined as the presence of a known metastatic lesion or a newly developed suspicious metastatic lesion by imaging. We adopted the definition of the dynamic risk stratification system by Tuttle *et al.* (11, 12). However, we also took into consideration the levels of serum TgAb during our analysis. We did not include the level of suppressed Tg or changes in the Tg level as a risk classifier. This modified definition of dynamic risk stratification system is shown in Table 1.

Statistical analysis

We used R version 2.13 and R libraries survival, car and Cairo to analyze data and draw graphs (R Foundation for Statistical Computing, Vienna, Austria, <http://www.R-project.org>). Continuous variables were presented as medians with interquartile range (IQR) and categorical variables were presented as numbers with percentages.

The primary endpoint of this study was the structural recurrence or persistence after initial therapy, pathologically proven malignant tissue and/or metastatic lesions in other distant organs by imaging studies.

A Cox proportional hazard model was used to evaluate the risk of structural recurrence/persistence of disease. DFS curves were constructed using the Kaplan–Meier method, and the log rank test was used to evaluate differences in DFS between groups. The relative importance of each risk stratification was determined by calculating the proportion of variation in survival time explained (PVE) by

Table 1 Definition of the modified dynamic risk stratification by the response to the initial therapy.

Excellent response	Acceptable response	Biochemical incomplete response	Structural incomplete response
Stimulated Tg <1 ng/ml	Stimulated Tg 1–10 ng/ml or positive TgAb ^a	Stimulated Tg ≥ 10 ng/ml	Suspicious recurrent or persistent disease on images
Negative TgAb ^a		No evidence of recurrent or persistent disease on images	
No evidence of recurrent or persistent disease on images ^b	No evidence of recurrent or persistent disease on images		

^aThe cutoff value of TgAb was 60 IU/ml. A TgAb value ≥ 60 IU/ml was considered to be positive.

^bImages: diagnostic whole-body scan or neck ultrasonography or any cross-sectional images during first 2 years after remnant ablation.

Table 2 Description of study cohort.

Characteristics	Median (IQR)	n (%)
Age (years)	46.2 (37.1–54.7)	
Sex		
Female		622 (87)
Male		93 (13)
¹³¹ I dose for remnant ablation (GBq)	5.6 (3.7–7.4)	
Follow-up duration (years)	8 (6.3–9.3)	
Histology		
Papillary		643 (90)
Variants of papillary		38 (5)
Follicular		18 (2)
Hürthle cell		4 (1)
Mixed		12 (2)
TgAb status		
Positive TgAb at the time of ablation		110 (15)
Positive TgAb at 6–24 months after initial treatment		35 (5)
Extent of surgery		
Total thyroidectomy alone		99 (14)
Total thyroidectomy with central neck dissection		499 (70)
Total thyroidectomy with lateral neck dissection		117 (16)
Clinical status at the end of period		
No evidence of disease		552 (88)
Biochemical persistent disease		29 (4)
Structural persistent disease		127 (18)
Recurrence		7 (1)

IQR, interquartile range.

the maximum likelihood ratio (G^2) from the Cox model ($PVE = 1 - \exp(-G^2/n)$) (4, 17). The PVE (%) ranges from 0 to 100 and the higher PVE value suggests the more accurate predictability of risk of recurrence. DFS curves were also constructed using the Kaplan–Meier method. All P values were two sided with $P < 0.05$ considered statistically significant.

Results

The median follow-up period of 715 patients was 8 years (IQR 6.3–9.3). The median age of the patients was 46.2 years (IQR 37.1–54.7) and 622 patients (87%) were female. The median dose of administered radioiodine was 5.6 GBq. The majority of patients had papillary thyroid carcinoma (PTC, 95%) including variant types of papillary carcinoma, while a small percentage of patients had follicular thyroid carcinoma (2%), Hürthle cell carcinoma (1%), or a mixed type of thyroid carcinoma (2%). At the time of last follow-up, 552 patients (88%) were NED, 156 patients (22%) had persistent disease, and seven patients

(1%) showed recurrent disease (Table 2). The primary endpoint, structural recurrent/persistent disease, was found in 132 patients (18%) during the follow-up period. Most of the structural recurrent/persistent disease was confined to the cervical lymph nodes in these patients ($n=100$; 76%). Seven patients (1%) with structural persistent disease died due to the progression of DTC. The median period from initial therapy to the detection of structural recurrent/persistent disease was 2.8 years (IQR 1.4–4.8).

As shown in Table 3, the TNM staging system classified patients into stage I ($n=348$; 49%), stage II ($n=32$; 4%), stage III ($n=271$; 38%), and stage IV ($n=64$; 9%). DFS rates were 83.0% in stage I, 58.8% in stage II, 85.6% in stage III, and 22.3% in stage IV at the end of the follow-up period (Fig. 2A, $\chi^2=10.06$, $P=0.002$). Structural recurrent/persistent disease was more prevalent in stage II patients compared with stage III patients because 17 patients (53%) in stage II had distant metastasis. The PVE value of the TNM staging system was 8.68%, and it was the lowest value among the three risk stratification systems (Table 3).

When we classified patients according to the ATA risk classification, 88 patients (12%) belonged to the low-risk group, 578 patients (81%) belonged to the intermediate-risk group, and 49 patients (7%) belonged to the high-risk group. The risk of recurrence/persistence was significantly higher in intermediate-risk ($HR=15.34$, $P=0.007$) and high-risk groups ($HR=106.16$, $P<0.001$) than in the low-risk group. The ATA risk classification clearly predicted the risk of DTC recurrence/persistence in our study subjects, especially in the low-risk group. DFS rates were 98.7% in the low-risk group, 78.7% in the intermediate-risk group, and 33.8% in the high-risk group at the end of the follow-up period as shown in Kaplan–Meier curves for DFS (Fig. 2B, $\chi^2=74.96$, $P<0.001$). The PVE value of the ATA risk classification was 12.14% (Table 3).

The modified dynamic risk stratification, including the level of serum TgAb, was also applied to our study subjects. Table 4 shows the clinical outcome at the end of the follow-up period according to the modified dynamic risk stratification. The number of patients in the four groups was as follows: an excellent response was obtained in 435 patients (61%), an acceptable response was obtained in 143 patients (20%), a biochemical incomplete response was obtained in 49 patients (7%), and a structural incomplete response was obtained in 88 patients (12%). DFS rates were 98.3% in patients with an excellent response, 71.6% in patients with an acceptable response, 54.2% in patients with a biochemical incomplete

Clinical Study	M J Jeon and others	Modified dynamic risk stratification of DTC	170:1	27
----------------	---------------------	---	-------	----

Table 3 Risk stratification and the risk of recurrence/persistence in patients with differentiated thyroid carcinoma.

Classification	n (%)	Hazard ratio (95% CI)	P value	G ²	PVE (%)
TNM staging system (AJCC 6th)				64.89	8.68
I	348 (49)	Ref			
II	32 (4)	3.58 (1.95–6.58)	<0.001		
III	271 (38)	0.79 (0.51–1.23)	0.297		
IV	64 (9)	4.95 (3.22–7.61)	<0.001		
ATA risk classification				92.53	12.14
Low	88 (12)	Ref			
Intermediate	578 (81)	15.34 (2.14–110.00)	0.007		
High	49 (7)	106.16 (14.51–776.50)	<0.001		
Modified dynamic risk stratification				422.10	44.59
Excellent response	435 (61)	Ref			
Acceptable response	143 (20)	16.51 (6.34–43.01)	<0.001		
Biochemical incomplete response	49 (7)	41.33 (15.42–110.78)	<0.001		
Structural incomplete response	88 (12)	281.18 (112.87–700.47)	<0.001		
Dynamic risk stratification in patients with negative TgAb				385.30	43.26
Excellent response	435 (64)	Ref			
Acceptable response	117 (17)	13.57 (5.04–36.57)	<0.001		
Biochemical incomplete response	49 (7)	41.45 (15.46–111.15)	<0.001		
Structural incomplete response	79 (12)	267.79 (106.99–670.27)	<0.001		
Modified dynamic risk stratification in ATA intermediate-risk group					
Excellent response	352 (61)	Ref			
Acceptable response	127 (22)	13.52 (5.14–35.57)	<0.001		
Biochemical incomplete response	45 (8)	34.34 (12.65–93.18)	<0.001		
Structural incomplete response	54 (9)	300.10 (117.55–766.18)	<0.001		

G², maximal likelihood ratio; PVE, proportion of variance explained; Ref, reference; TNM, Tumor Node Metastasis; AJCC, American Joint Cancer Committee; ATA, American Thyroid Association.

response, and 6.8% in patients with a structural incomplete response (Fig. 2C, $\chi^2=609.9$, $P<0.001$). The risk of structural recurrent/persistent disease was significantly higher in acceptable responders (HR=16.51, $P<0.001$), biochemical incomplete responders (HR=41.33, $P<0.001$), and structural incomplete responders (HR=281.18, $P<0.001$) than in excellent responders. This risk stratification system had the highest PVE value (44.59%) among the three risk stratification systems (Table 3). Furthermore, this PVE value was higher than the PVE value of dynamic risk stratification for patients that were negative for TgAb (Table 3).

We also applied this modified dynamic risk stratification system in each subgroup of patients according to the ATA risk classification. This dynamic risk stratification predicted recurrence/persistence in patients with intermediate-risk ($P<0.001$, Fig. 2E) and high-risk groups ($P<0.001$, Fig. 2F). As shown in Table 3, patients belonging to the intermediate-risk group by the ATA classification were reclassified as an excellent response ($n=352$, 61%), an acceptable response ($n=127$, 22%), a biochemical incomplete response ($n=45$, 8%), and a structural incomplete response ($n=54$, 9%). The DFS rate was 97.9% in excellent responders, 71.9%

in acceptable responders, 54.9% in biochemical incomplete responders, and 1.9% in structural incomplete responders (Fig. 2E).

Discussion

This study demonstrated that the modified dynamic risk stratification system, which takes into consideration the level of serum TgAb, effectively estimated the risk of recurrence/persistence in patients with DTC. Most excellent responders did not experience structural recurrent/persistent disease, whereas it was confirmed in 93.2% of patients with structural incomplete response at the end of the follow-up period. The ATA risk classification also effectively predicted the low risk of recurrent/persistent disease. There was only one recurrence in 88 patients belonging to the low-risk classification. However, our modified dynamic risk stratification reclassified the risk of recurrent/persistent disease in intermediate- and high-risk groups according to the ATA risk classification. The TNM staging system was not appropriate for estimating the risk of recurrence/persistence in our study and this finding was consistent with previous reports (1, 18). We also calculated PVE values, and the modified dynamic

Figure 2

Disease-free survival (DFS) curves according to the risk of recurrence categories for each staging system. The time point corresponds to the time of initial treatment. (A) AJCC/TNM staging ($P=0.015$). (B) ATA risk classification system ($P<0.001$). (C) Modified dynamic risk stratification ($P<0.001$).

(D) DFS curves using the modified dynamic risk stratification in patients belonging to the ATA low-risk group ($P=0.703$). (E) The ATA intermediate-risk group ($P<0.001$) and (F) the ATA high-risk group ($P<0.001$).

risk stratification had the highest PVE value compared with the TNM staging system and the ATA risk classification system.

This new and practical dynamic risk stratification system was proposed by Tuttle *et al.* in 2010 (11). This system included the clinical status of the disease, the level of serum Tg, and the results of cross-sectional images (11, 12). The level of serum Tg is a well-known tumor marker for the patients with DTC, and we showed that only about 1% of DTC patients who had a serum-stimulated Tg level <1 ng/ml at 1 year after initial treatment developed recurrent disease (19, 20). However, the prognostic value of serum Tg measurement was limited in the presence of TgAb because of false-negative results owing to TgAb interference with immunometric assay (14, 21). We also suggested that the level of serum TgAb and a change in the serum TgAb concentration could be a prognostic indicator during the early postoperative period. Therefore, we included the level of serum TgAb in

our modified dynamic risk stratification and validated this system in our cohort (22).

The prevalence of positive serum TgAb in patients with DTC is 10–25% and is about twofold higher than in the general population (14, 21). The timing of TgAb measurement is likely to affect the prevalence of positive serum TgAb level because TgAb decreases and eventually disappears as it removes the antigenic stimulus from Tg after initial therapy (22). A previous study reported that TgAb continuously decreased in most patients after surgery and the prevalence of positive serum TgAb decreased to $<10\%$ after 3 years (23). In this study, the prevalence of positive serum TgAb was 15% at the time of the remnant ablation and decreased to 5% at the time of restratification, 6–24 months after initial treatment. Thus, the rate of TgAb positivity in this study may be underestimated because of a delay in TgAb measurement.

The proportion of patients in low- and high-risk groups according to the ATA risk classification was

Table 4 The clinical outcome according to the modified dynamic risk stratification.

n (%)	Excellent response	Acceptable response	Biochemical incomplete response	Structural incomplete response
NED	428 (98)	106 (74)	13 (27)	5 (6)
Biochemical, persistent disease	0	11 (8)	17 (35)	1 (1)
Structural, persistent disease	0	26 (18)	19 (39)	82 (93)
Recurrence	7 (2)	0	0	0

NED, no evidence of disease.

12 and 7% respectively in this study. These values are much lower than previous reports using the dynamic risk stratification system (11, 12, 13). We excluded patients with microcarcinomas in this study because of the controversy in using remnant RAI ablation for patients with microcarcinoma (24, 25, 26). Moreover, we only included patients who underwent surgery between 2000 and 2005 because routine neck US was performed during follow-up after 2000. This could be the reason for the low number of patients in low- and high-risk groups. More intensive examination of pathological specimens could also be another possible reason for the low number of patients in the low-risk group. Identification of small lymph node metastases, single microscopic extrathyroidal extensions, or vascular invasions can change the risk of recurrence from low to intermediate unnecessarily (18). In this regard, the dynamic risk stratification system is a simple and practical system that could overcome weak points of initial pathological staging by reflecting clinical disease status with highly sensitive neck US and the level of serum Tg and TgAb. This approach may help to tailor management of the disease.

This study has several limitations. First, we could not validate the dynamic risk stratification system for patients who did not receive RAI ablation therapy or for those who did lobectomy because we only included patients who underwent total or near total thyroidectomy followed by RAI ablation, which made relatively homogenous characteristics and less confounding factors for analysis. A recent study suggested that the serum Tg level may also be useful during the follow-up period and it can predict outcomes for patients who did not receive RAI ablation therapy (27). In this regard, further studies are required to validate whether the dynamic risk stratification system could be useful for patients who did not receive RAI ablation therapy. Secondly, most of the patients in this study had

PTC (95%) and the proportion of PTC was much higher than in prior studies (11, 12). This is an epidemiological characteristic of Korean patients with thyroid carcinoma (28), but we suggest that this dynamic system could be applicable to general DTC patients. The relatively shorter follow-up period (median 8 years) is also a limitation of the current study.

The strength of this study is that all patients received their initial therapies, including surgery and RAI therapy, at our center. Therefore, we can minimize the effects of many variable confounding factors, such as surgical skills, extent of surgery, strategies of RAI ablation, and so on.

In conclusion, we propose the modified dynamic risk stratification by including an additional variable of serum TgAb level to evaluate the risk of recurrent/persistent DTC. We have shown that this modified dynamic risk stratification effectively predicts the recurrence/persistence of DTC in patients who received total thyroidectomy followed by RAI ablation. This personalized risk assessment using the response to initial therapy could be useful for the follow-up and management of patients with DTC.

Declaration of interest

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

Funding

This study was supported by the National Research Foundation (NRF) of Korea (Research grant number NRF-2012R1A1A2038383).

References

- Cooper DS, Doherty GM, Haugen BR, Hauger BR, Kloos RT, Lee SL, Mandel SJ, Mazzaferri EL, McIver B, Pacini F *et al.* Revised American Thyroid Association management guidelines for patients with thyroid nodules and differentiated thyroid cancer. *Thyroid* 2009 **19** 1167–1214. (doi:10.1089/thy.2009.0110)
- DeGroot LJ, Kaplan EL, McCormick M & Straus FH. Natural history, treatment, and course of papillary thyroid carcinoma. *Journal of Clinical Endocrinology and Metabolism* 1990 **71** 414–424. (doi:10.1210/jcem-71-2-414)
- Cady B & Rossi R. An expanded view of risk-group definition in differentiated thyroid carcinoma. *Surgery* 1988 **104** 947–953.
- Brierley JD, Panzarella T, Tsang RW, Gospodarowicz MK & O'Sullivan B. A comparison of different staging systems predictability of patient outcome. Thyroid carcinoma as an example. *Cancer* 1997 **79** 2414–2423. (doi:10.1002/(SICI)1097-0142(19970615)79:12<2414::AID-CNCR18>3.0.CO;2-U)
- Lang BH, Lo CY, Chan WF, Lam KY & Wan KY. Staging systems for papillary thyroid carcinoma: a review and comparison. *Annals of Surgery* 2007 **245** 366–378. (doi:10.1097/01.sla.0000250445.92336.2a)
- Greene FL, Page DL, Fleming ID, Balch CM & Fritz AG. *AJCC Cancer Staging Handbook: from the AJCC Cancer Staging Manual*. Springer Verlag, 2002.

- 7 Davies L & Welch HG. Increasing incidence of thyroid cancer in the United States, 1973–2002. *Journal of the American Medical Association* 2006 **295** 2164–2167. (doi:10.1001/jama.295.18.2164)
- 8 Howlader N, Noone A, Krapcho M, Neyman N, Aminou R, Waldron W, Altekruse S, Kosary C, Ruhl J & Tatalovich Z. *SEER Cancer Statistics Review, 1975–2008*. Bethesda, MD: National Cancer Institute, 2011.
- 9 Yang L, Shen W & Sakamoto N. Population-based study evaluating and predicting the probability of death resulting from thyroid cancer and other causes among patients with thyroid cancer. *Journal of Clinical Oncology* 2013 **31** 468–474. (doi:10.1200/JCO.2012.42.4457)
- 10 Eustatia-Rutten CF, Corssmit EP, Biermasz NR, Pereira AM, Romijn JA & Smit JW. Survival and death causes in differentiated thyroid carcinoma. *Journal of Clinical Endocrinology and Metabolism* 2006 **91** 313–319. (doi:10.1210/jc.2005-1322)
- 11 Tuttle RM, Tala H, Shah J, Leboeuf R, Ghossein R, Gonen M, Brokhin M, Omry G, Fagin JA & Shaha A. 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** 1341–1349. (doi:10.1089/thy.2010.0178)
- 12 Vaisman F, Tala H, Grewal R & Tuttle RM. In differentiated thyroid cancer, an incomplete structural response to therapy is associated with significantly worse clinical outcomes than only an incomplete thyroglobulin response. *Thyroid* 2011 **21** 1317–1322. (doi:10.1089/thy.2011.0232)
- 13 Vaisman F, Momesso D, Bulzico DA, Pessoa CH, Dias F, Corbo R, Vaisman M & Tuttle RM. Spontaneous remission in thyroid cancer patients after biochemical incomplete response to initial therapy. *Clinical Endocrinology* 2012 **77** 132–138. (doi:10.1111/j.1365-2265.2012.04342.x)
- 14 Spencer CA. Clinical review: clinical utility of thyroglobulin antibody (TgAb) measurements for patients with differentiated thyroid cancers (DTC). *Journal of Clinical Endocrinology and Metabolism* 2011 **96** 3615–3627. (doi:10.1210/jc.2011-1740)
- 15 Castagna MG, Maino F, Cipri C, Belardini V, Theodoropoulou A, Cevenini G & Pacini F. Delayed risk stratification, to include the response to initial treatment (surgery and radioiodine ablation), has better outcome predictivity in differentiated thyroid cancer patients. *European Journal of Endocrinology* 2011 **165** 441–446. (doi:10.1530/EJE-11-0466)
- 16 Yim JH, Kim WB, Kim EY, Kim WG, Kim TY, Ryu JS, Gong G, Hong SJ & Shong YK. The outcomes of first reoperation for locoregionally recurrent/persistent papillary thyroid carcinoma in patients who initially underwent total thyroidectomy and remnant ablation. *Journal of Clinical Endocrinology and Metabolism* 2011 **96** 2049–2056. (doi:10.1210/jc.2010-2298)
- 17 Schemper M. The relative importance of prognostic factors in studies of survival. *Statistics in Medicine* 1993 **12** 2377–2382. (doi:10.1002/sim.4780122413)
- 18 Randolph GW, Duh QY, Heller KS, LiVolsi VA, Mandel SJ, Steward DL, Tufano RP & Tuttle RM. 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** 1144–1152. (doi:10.1089/thy.2012.0043)
- 19 Han JM, Kim WB, Yim JH, Kim WG, Kim TY, Ryu JS, Gong G, Sung TY, Yoon JH, Hong SJ *et al.* Long-term clinical outcome of differentiated thyroid cancer patients with undetectable stimulated thyroglobulin level one year after initial treatment. *Thyroid* 2012 **22** 784–790. (doi:10.1089/thy.2011.0322)
- 20 Mazzaferri EL, Robbins RJ, Spencer CA, Braverman LE, Pacini F, Wartofsky L, Haugen BR, Sherman SI, Cooper DS, Braunstein GD *et al.* A consensus report of the role of serum thyroglobulin as a monitoring method for low-risk patients with papillary thyroid carcinoma. *Journal of Clinical Endocrinology and Metabolism* 2003 **88** 1433–1441. (doi:10.1210/jc.2002-021702)
- 21 Spencer CA. Challenges of serum thyroglobulin (Tg) measurement in the presence of Tg autoantibodies. *Journal of Clinical Endocrinology and Metabolism* 2004 **89** 3702–3704. (doi:10.1210/jc.2004-0986)
- 22 Kim WG, Yoon JH, Kim WB, Kim TY, Kim EY, Kim JM, Ryu JS, Gong G, Hong SJ & Shong YK. Change of serum antithyroglobulin antibody levels is useful for prediction of clinical recurrence in thyroglobulin-negative patients with differentiated thyroid carcinoma. *Journal of Clinical Endocrinology and Metabolism* 2008 **93** 4683–4689. (doi:10.1210/jc.2008-0962)
- 23 Gorges R, Maniecki M, Jentzen W, Sheu SN, Mann K, Bockisch A & Janssen OE. Development and clinical impact of thyroglobulin antibodies in patients with differentiated thyroid carcinoma during the first 3 years after thyroidectomy. *European Journal of Endocrinology* 2005 **153** 49–55. (doi:10.1530/eje.1.01940)
- 24 Gallicchio R, Giacomobono S, Capacchione D, Nardelli A, Barbato F, Nappi A, Pellegrino T & Storto G. Should patients with remnants from thyroid microcarcinoma really not be treated with iodine-131 ablation? *Endocrine* 2013 **44** 426–433. (doi:10.1007/s12020-013-9935-9)
- 25 Hay ID, Hutchinson ME, Gonzalez-Losada T, McIver B, Reinalda ME, Grant CS, Thompson GB, Sebo TJ & Goellner JR. Papillary thyroid microcarcinoma: a study of 900 cases observed in a 60-year period. *Surgery* 2008 **144** 980–987 (discussion 987–988). (doi:10.1016/j.surg.2008.08.035)
- 26 Mihailovic J, Stefanovic L & Stankovic R. Influence of initial treatment on the survival and recurrence in patients with differentiated thyroid microcarcinoma. *Clinical Nuclear Medicine* 2013 **38** 332–338. (doi:10.1097/RLU.0b013e3182872ed2)
- 27 Durante C, Montesano T, Attard M, Torlontano M, Monzani F, Costante G, Meringolo D, Ferdeghini M, Tumino S, Lamartina L *et al.* Long-term surveillance of papillary thyroid cancer patients who do not undergo postoperative radioiodine remnant ablation: is there a role for serum thyroglobulin measurement? *Journal of Clinical Endocrinology and Metabolism* 2012 **97** 2748–2753. (doi:10.1210/jc.2012-1123)
- 28 Cho BY, Choi HS, Park YJ, Lim JA, Ahn HY, Lee EK, Kim KW, Yi KH, Chung JK, Youn YK *et al.* Changes in the clinicopathological characteristics and outcomes of thyroid cancer in Korea over the past four decades. *Thyroid* 2013 **23** 797–804. (doi:10.1089/thy.2012.0329)

Received 26 June 2013

Revised version received 4 September 2013

Accepted 2 October 2013