

## CLINICAL STUDY

# Delayed risk stratification, to include the response to initial treatment (surgery and radioiodine ablation), has better outcome predictivity in differentiated thyroid cancer patients

Maria Grazia Castagna<sup>1</sup>, Fabio Maino<sup>1</sup>, Claudia Cipri<sup>1</sup>, Valentina Belardini<sup>1</sup>, Alexandra Theodoropoulou<sup>1</sup>, Gabriele Cevenini<sup>2</sup> and Furio Pacini<sup>1</sup>

<sup>1</sup>Department of Internal Medicine, Endocrinology and Metabolism and Biochemistry, Section of Endocrinology and Metabolism and <sup>2</sup>Department of Surgery and Bioengineering, University of Siena, Policlinico Santa Maria alle Scotte, Viale Bracci 1, 53100 Siena, Italy

(Correspondence should be addressed to F Pacini; Email: pacini8@unisi.it)

## Abstract

**Introduction:** After initial treatment, differentiated thyroid cancer (DTC) patients are stratified as low and high risk based on clinical/pathological features. Recently, a risk stratification based on additional clinical data accumulated during follow-up has been proposed.

**Objective:** To evaluate the predictive value of delayed risk stratification (DRS) obtained at the time of the first diagnostic control (8–12 months after initial treatment).

**Methods:** We reviewed 512 patients with DTC whose risk assessment was initially defined according to the American (ATA) and European Thyroid Association (ETA) guidelines. At the time of the first control, 8–12 months after initial treatment, patients were re-stratified according to their clinical status: DRS.

**Results:** Using DRS, about 50% of ATA/ETA intermediate/high-risk patients moved to DRS low-risk category, while about 10% of ATA/ETA low-risk patients moved to DRS high-risk category. The ability of the DRS to predict the final outcome was superior to that of ATA and ETA. Positive and negative predictive values for both ATA (39.2 and 90.6% respectively) and ETA (38.4 and 91.3% respectively) were significantly lower than that observed with the DRS (72.8 and 96.3% respectively,  $P < 0.05$ ). The observed variance in predicting final outcome was 25.4% for ATA, 19.1% for ETA, and 62.1% for DRS.

**Conclusions:** Delaying the risk stratification of DTC patients at a time when the response to surgery and radioiodine ablation is evident allows to better define individual risk and to better modulate the subsequent follow-up.

*European Journal of Endocrinology* 165 441–446

## Introduction

In the last several years, an increased emphasis has been posed on using individual estimates of risk to guide both initial therapy and follow-up in differentiated thyroid cancer (DTC) patients. Several different risk stratification systems have been published (1–6) and the most popular is the AJCC/UICC system (6). All of them have been developed to predict the risk of death but not of recurrence and, being based on clinicopathological factors available soon after diagnosis and initial surgical therapy, do not change over time. To overcome this limitation, the American Thyroid Association (ATA) in recently published guidelines (7) graded the risk of recurrence into three categories (low, intermediate, and high) based on tumor-related parameters (pTNM and histological variant) integrated with other clinical features, including the result of the first post-therapy radioiodine whole-body

scan (WBS) and serum thyroglobulin (Tg) measurement. Similarly, the European Thyroid Association (ETA) in its consensus report published in 2006 (8) stratified thyroid cancer patients into three risk levels (very low, low, and high) according to the pTNM, histological variant, age, and the results of the first post-therapy radioiodine WBS. When applied to large number of patients, the ATA and the ETA risk stratification systems gave quite similar figures. Recent reports (9–11) have proposed to revise the initial risk stratification incorporating additional clinical data that are available during the subsequent follow-up (after initial treatment): the so-called ‘ongoing risk stratification’.

The aim of our study was to evaluate the predictive value of a delayed risk stratification (DRS) system obtained at the time of the first diagnostic follow-up (8–12 months after radioiodine ablation) taking into account the effect of initial treatment.

## Patients and methods

### Patients and treatment

Epidemiological and clinical features of the DTC patients are reported in Table 1. We retrospectively reviewed 512 patients with DTC treated at the Section of Endocrinology, University of Siena, Italy. There were 376 females (73.4%) and 136 males (26.6%) aged 4–86 years (mean  $\pm$  s.d. =  $46.4 \pm 16.1$ ) at the time of diagnosis. Four hundred and fifty-six (89%) patients had papillary and 56 (11%) had follicular thyroid cancer. Using the AJCC/UICC system (6), 348 (68%) patients were stage I, 43 (8.4%) were stage II, 96 (18.6%) were stage III, and 25 (5%) were stage IV. All patients had been treated with near-total thyroidectomy with (7.4%) or without lymph node dissection of the central neck compartment. Unilateral or bilateral lymph node dissection was performed in 14.8% of patients with preoperatively known lymph node metastases. After surgery, all patients were treated with

radioiodine therapy with a mean activity of  $2960 \pm 1295$  MBq of  $^{131}\text{I}$  (range 555–7400 MBq). Radioiodine therapy was given after L-thyroxine (L-T<sub>4</sub>) withdrawal in 280/512 (54.6%) patients and after recombinant human TSH (rhTSH) in the remaining 232/512 (45.4%) patients.

As shown in Table 1, during follow-up ( $6.8 \pm 4.8$  years, median 5.6 years, range 1.08–52 years), 8/512 (1.6%) patients died of thyroid cancer after a mean interval of 5.4 years (range 1.9–14.3 years). The overall survival was 98.4% at 5 years and 97.6% at 10 years. Of the 512 patients, ten (1.9%) developed proven recurrence (local disease in five patients and distant metastases in five patients) after a mean interval of 5.6 years (range 1.6–9.1 years). Four additional patients developed elevation of serum Tg without any structural evidence of disease. The disease-free survival at 5 and 10 years was 98.3 and 91.6% respectively. At the last follow-up, 385/512 (75.2%) patients were in clinical remission whereas persistent disease was observed in 119/512 (23.2%) patients (biochemical disease in 49 patients and evidence of disease in 70 patients).

**Table 1** Epidemiological and clinical features of differentiated thyroid cancer patients (n=512).

Characteristics	Values
Age (years)	
Mean $\pm$ s.d.	46.4 $\pm$ 16.1
Median (range)	46 (4–86)
Gender (n (%))	
Male	136 (26.6%)
Female	376 (73.4%)
Histology (n (%))	
Papillary	456 (89%)
Follicular	56 (11%)
TNM staging <sup>a</sup> (n (%))	
T1–2 N0 M0	234 (45.7%)
T3–4 N0 M0	87 (17%)
AnyT N1 M0	146 (28.5%)
AnyT Any N M1	45 (8.8)
Stage (n (%))	
Stage 1	348 (68.0%)
Stage 2	43 (8.4%)
Stage 3	96 (18.6%)
Stage 4	25 (5.0%)
Surgery (n (%))	
Total Tx	398 (77.8%)
Total Tx + central lymph node dissection	38 (7.4%)
Total Tx + lateral lymph node dissection	76 (14.8%)
$^{131}\text{I}$ activity for ablation (MBq)	
Mean $\pm$ s.d.	2960 $\pm$ 1295
Median (range)	3700 (555–11 100)
Preparation for thyroid remnant ablation (n (%))	
Hypothyroidism	280 (54.6%)
rhTSH	232 (45.4%)
Follow-up (years)	
Mean $\pm$ s.d.	6.8 $\pm$ 4.8
Median (range)	5.6 (1.08–52)
Clinical status at the end of follow-up (n (%))	
Remission	385 (75.2%)
Persistent disease	119 (23.2%)
Death of disease	8 (1.6%)

<sup>a</sup>AJCC 6th edition, 2002.

### Follow-up and diagnostic tests

At 8–12 months after initial therapy (total thyroidectomy and radioiodine remnant ablation), the follow-up was based on physical examination, neck ultrasound (US), and TSH-stimulated serum Tg measurement with or without diagnostic WBS. According to the results of this evaluation and regardless of the ETA/ATA risk category, the subsequent follow-up was also established: patients in clinical remission (undetectable basal and stimulated serum Tg, negative neck US) were followed annually with additional basal and stimulated serum Tg and neck US; patients with persistent disease (evidence of disease by any imaging modality or cytology or histology) were followed with imaging techniques and subjected to additional therapies whenever indicated; patients with only evidence of biochemical disease (detectable basal and/or stimulated serum Tg, but no positive imaging) were followed with periodical assessment of serum Tg and imaging.

### Methods

Serum Tg, TSH, and Tg antibodies (TgAb) were measured by solid-phase chemiluminescent assays (Immulate 2000, DPC Diagnostic Products Corporation, Los Angeles, CA, USA). Tg assays had a functional sensitivity of 0.9 ng/ml.

Neck US was performed with a color Doppler apparatus (AU 590 Asynchronous, Esaote Biomedica, Firenze, Italy) and a 7.5 MHz linear transducer by experienced endocrinologists (members of our staff) trained in neck US. Whenever neck lymph nodes with US criteria of malignancy (cystic appearance, hyper-echoic punctuations, loss of hilum, and peripheral

vascularization) were present, they were subjected to fine needle aspiration cytology (FNAC) and measurement of Tg in the FNAC washout.

$^{131}\text{I}$  WBS was obtained using a one-head gamma camera (Apex SPX 4000, Elscint Italia, Milano, Italy) with a high-energy collimator and a sensitivity of 160 c.p.m./ $\mu\text{Ci}$ . The scan speed was 6 cm/min for diagnostic WBS and 8 cm/min for post-therapeutic WBS, with total counts of at least 100 000 c.p.m.

### Risk stratification

**Post-surgical risk assessment** Soon after initial treatment (total thyroidectomy and radioiodine remnant ablation), all patients were stratified using the ATA (7) and ETA (8) guidelines. According to ETA (8), patients were classified at high risk when the primary tumor was  $\geq 4$  cm in diameter or extending beyond the thyroid capsule, had an aggressive histology, or in the presence of lymph node or distant metastases. All other patients were considered at low risk of recurrence.

According to the ATA guidelines (7), we classified low-risk patients with the following characteristics: no local or distant metastases, all macroscopic tumor has been resected, no tumor invasion of locoregional tissues or structures, no aggressive histology or vascular invasion, and, if  $^{131}\text{I}$  was given, no  $^{131}\text{I}$  uptake outside the thyroid bed on the post-therapeutic WBS. Patients with microscopic invasion of tumor into the perithyroidal soft tissues at initial surgery, cervical lymph node metastases, or  $^{131}\text{I}$  uptake outside the thyroid bed on the post-therapeutic WBS or tumor with aggressive histology or vascular invasion were classified as at intermediate risk. Finally, we classified at high-risk patients with macroscopic tumor invasion, incomplete tumor resection, distant metastases, and possibly thyroglobulinemia out of proportion to what is seen on the post-ablative scan. For statistical purposes, patients in the intermediate and high-risk groups were pooled together.

**Delayed risk stratification** At the time of the first control, performed 8–12 months after initial therapy, we re-stratified our patients according to the clinical status. Patients with undetectable basal and stimulated serum Tg, negative anti-TgAb, and no evidence of disease (at clinical examination, neck US and diagnostic  $^{131}\text{I}$  WBS when performed) were defined in ‘clinical remission’ and were shifted to the long-term follow-up based on annual serum Tg measurement on replacement doses of L-T<sub>4</sub> therapy and neck US. Patients with any evidence of disease at clinical and neck US examination, imaging (chest X-ray,  $^{131}\text{I}$  WBS, 18-fluorodeoxyglucose-positron emission tomography, computed tomography, magnetic resonance imaging, and bone scan), and/or detectable basal/stimulated serum Tg were defined as ‘persistent disease’ and were

subjected to appropriate treatments. For the aim of our study, patients in clinical remission composed the DRS low-risk group and patients with persistent disease composed the DRS high-risk category.

### Statistical analysis

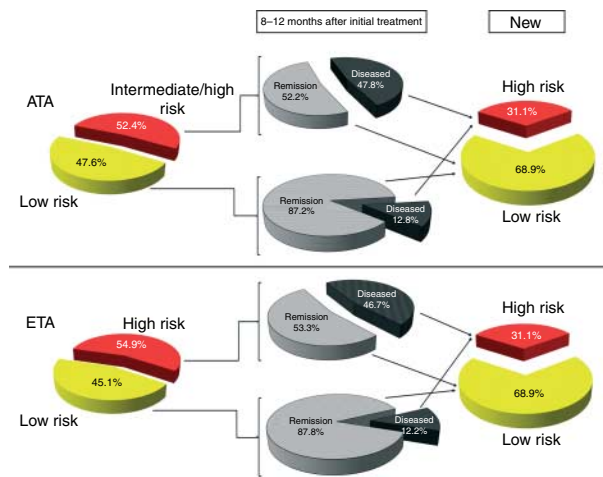
Epidemiological data are presented as the mean  $\pm$  s.d., with median when appropriate. The Kaplan–Meier test was performed to analyze time-dependent variables. To evaluate significant differences in data frequency, we analyzed  $2 \times 2$  contingency tables by the Fisher exact test or  $2 \times 3$  contingency table by the  $\chi^2$  test. Statistical analysis was performed using the software StatView for Windows version 5.0.1 (SAS Institute, Cary, NC, USA). Diagnostic accuracy was calculated according to Galen (12) and was based on true positive (TP), true negative (TN), false positive (FP), and false negative (FN) results. The positive predictive value (PPV) was  $\text{TP}/(\text{TP} + \text{FP})$  and the negative predictive value (NPV) was  $\text{TN}/(\text{FN} + \text{TN})$ . The 95% confidence interval (CI) of all estimates was also evaluated.

The agreement between different risk stratification systems was calculated using Cohen’s  $\kappa$  coefficient. A value of 1 implies perfect agreement and values  $< 1$  implies less than perfect agreement. It was evaluated using the Landis and Koch semi-quantitative scale (poor agreement =  $< 0.20$ , fair agreement =  $0.20$ – $0.40$ , moderate agreement =  $0.40$ – $0.60$ , good agreement =  $0.60$ – $0.80$ , and very good agreement =  $0.80$ – $1.0$ ). Negative values of  $\kappa$  express disagreement. Nagelkerke’s estimation of proportion of variance explained (PVE %) was evaluated from logistic regression analysis computed using the SPSS statistical software (SPSS version 10.0, Inc., Chicago, IL, USA). We considered a  $P < 0.05$  to be statistically significant for all analysis.

## Results

### ATA, ETA, and DRS

At the time of initial treatment, according to ETA consensus (8), 231/512 (45.1%) patients resulted at low risk and 281/512 (54.9%) at high risk of recurrence (Fig. 1, left side). According to the ATA guidelines (7), 244/512 (47.6%) patients resulted at low risk and 268/512 (52.4%) at intermediate/high risk (Fig. 1, left side). At the first control, performed 8–12 months after surgery and thyroid ablation, we assessed how many patients in the low- and high-risk groups were in complete remission or had persistent disease. As shown in Fig. 1 (center), according to the definition given in the section ‘Delayed risk stratification’, complete remission was observed in 87.8% ETA and 87.2% ATA low-risk patients and in 53.3% ETA and 52.2% ATA high-risk patients. As expected, complete remissions were more frequent ( $P < 0.0001$ ) among



**Figure 1** (Left side) Post-surgical risk stratification based on the ETA and ATA guidelines; (center) clinical status at the 8- to 12-month control after initial treatment; (right side) composition of the DRS system.

low-risk patients than among high-risk patients (both ETA and ATA classification). At this point, patients in complete remission composed the DRS low-risk category and patients with persistent disease composed the DRS high-risk category. As shown in Fig. 1 (right side), the ATA and ETA high-risk patients resulting in complete remission (nearly 50%) moved to the DRS low-risk category and the ATA and ETA low-risk patients displaying persistent disease (nearly 12%) composed the DRS high-risk category. Thus, using DRS in 353/512 (68.9%) patients may now be considered at low risk and 159/512 (31.1%) at high risk.

Using the Cohen's  $\kappa$  coefficient, the agreement between ATA and ETA risk stratification was 'very good' ( $\kappa=0.925$ ), but no agreement was found between ATA or ETA and DRS stratification ( $\kappa=-0.360$  and  $-0.359$  respectively), demonstrating that the risk stratification is similar between ATA and ETA but significantly different between ATA or ETA and DRS.

### Correlation between risk stratification systems (ATA, ETA, and DRS) and final outcome

To assess which risk stratification system had the better predictive value, the ATA, ETA, and DRS systems were correlated with the final outcome (Table 2). As expected, over 90% (91.4% of the patients in the ETA stratification, 90.8% in the ATA stratification, and 96.6% in the DRS stratification) of low-risk patients were in clinical remission at final outcome, regardless of the risk stratification system used. However, the rate of clinical remission was significantly higher in the DRS low-risk category ( $P=0.005$ ). In the group of patients at intermediate/high risk, we found a larger proportion of patients in clinical remission at the last follow-up in the ATA and ETA groups (60.8 and 61.6% respectively), compared with that observed in the group of the DRS high-risk category (27.1%;  $P<0.0001$ ). On the contrary, the rate of patients with persistent disease in the intermediate/high-risk groups was significantly higher in the DRS high-risk category (66.6%) than in the ATA and ETA intermediate/high-risk groups (33.9 and 32.8% respectively;  $P<0.001$ ).

In our study, the rate of recurrent disease, after a period of complete remission, was very low (2.7%) and not different between low- and intermediate/high-risk patients, regardless of the risk stratification system (ATA low 3.2%, intermediate/high 2.3%,  $P=0.59$ ; ETA low 2.5%, high 2.8%,  $P>0.99$ ; and DRS low 3.4%, high 1.2%,  $P=0.24$ ), suggesting that patients in clinical remission have a good prognosis, regardless of their initial staging. The mortality for thyroid cancer was observed only in a few intermediate/high-risk patients.

### PPV, NPV, and PVE for the different risk stratification systems in predicting final outcome

We evaluated the ability of different risk stratification systems to predict the final outcome by determining the PPV and the NPV. As shown in Table 3, the PPV

**Table 2** Clinical outcome at the end of follow-up according to ETA, ATA, and DRS.

	Final outcome			
	Remission	Persistent disease	Recurrent	Mortality
<b>ETA</b>				
Low risk ( <i>n</i> 231)	211* (91.4%)	14 (6.1%)	6 (2.5%)	0 (0%)
High risk ( <i>n</i> 281)	173† (61.6%)	92‡ (32.8%)	8 (2.8%)	8 (2.8%)
<b>ATA</b>				
Low risk ( <i>n</i> 244)	221* (90.8%)	15 (6.0%)	8 (3.2%)	0 (0%)
Intermediate/high risk ( <i>n</i> 268)	163† (60.8%)	91‡ (33.9%)	6 (2.3%)	8 (3.1%)
<b>DRS</b>				
Low risk ( <i>n</i> 353)	341* (96.6%)	0 (0%)	12 (3.4%)	0 (0%)
High risk ( <i>n</i> 159)	43† (27.1%)	106‡ (66.6%)	2 (4.0%) <sup>a</sup>	8 (5.1%)

\* $P=0.005$ ; † $P<0.0001$ ; ‡ $P<0.0001$ .

<sup>a</sup>Calculated on the number of high-risk patients who achieved remission ( $n=45$ ) at some point during the follow-up.



**Table 3** PPV, NPV, and PVE according to the different risk stratification.

	PPV	95% CI	NPV	95% CI	PVE (%)
ETA	<b>0.384</b> <sup>*,†</sup>	0.355– 0.407	<b>0.913</b> <sup>*,†</sup>	0.878– 0.941	19.1
ATA	<b>0.392</b> <sup>*,†</sup>	0.360– 0.417	<b>0.906</b> <sup>*,†</sup>	0.871– 0.934	25.4
DRS	<b>0.728</b> <sup>†</sup>	0.685– 0.759	<b>0.963</b> <sup>†</sup>	0.944– 0.977	62.1

\* $P > 0.05$ ; † $P < 0.05$ .

was very low for both ATA and ETA, 39.2% (95% CI = 36.0–41.7%) and 38.4% (95% CI = 35.5–40.7%), respectively, whereas the NPV was good for both systems, 90.6% (95% CI = 87.1–93.4%) and 91.3% (95% CI = 87.8–94.1%), respectively, without difference between the two ( $P > 0.05$ ). The NPV and the PPV of DRS were good, 96.3% (95% CI = 94.4–97.7%) and 72.8% (95% CI = 68.5–75.9%), respectively, and significantly better than that observed for ATA and ETA ( $P < 0.05$ ).

The ability of different risk stratification systems to fit the actual outcome (remission/persistent disease) was determined using the PVE, which was 25.4% for ATA, 19.1% for ETA, and 62.1% for DRS (Table 3).

## Discussion

The concept of 'ongoing risk stratification' has been proposed and validated by Tuttle *et al.* (9–11) as a new tool to assess the individual risk of recurrent disease in DTC patients treated with total thyroidectomy and radioiodine remnant ablation. This concept is based on the continuous integration of the initial risk stratification (at the time of diagnosis) with the clinical and morphological data becoming available during follow-up. Our results are in agreement with and reinforce these observations. Although we found that the risk stratification proposed by ATA (7) and ETA (8) is a good starting point for initial decision making, they are less accurate in predicting the long-term outcome in DTC patients. Indeed, both systems have a very low PPV due to the fact that a large number of patients (about 60%) classified at intermediate/high risk are in complete remission at the end of follow-up.

This drawback is probably due to the lack of consideration of the effects of the initial therapy. When we re-stratified our patients according to the results of the 8–12 months control after initial treatment, we were able to classify as low risk a significant number of patients that were initially considered (misleadingly) as high risk. Interestingly, almost all of these patients continued to be in apparent remission up to the end of the follow-up. The clinical relevance is that our DRS has significantly better PPV and NPV in predicting the final outcome than that obtained using ATA and ETA risk stratification. It is apparent that the favorable clinical

outcome of the DRS low-risk patients is due to the radical effect of total thyroidectomy and radioiodine remnant ablation. Shifting these previously high-risk patients in the DRS low-risk category, just 8–12 months after initial treatment, allows applying a less aggressive diagnostic follow-up in the subsequent years in a considerable number of patients.

Similar results have been reported by Tuttle *et al.* (11) in 588 DTC patients stratified according to the response to therapy after 2 years of follow-up. Our protocol has the advantage of reaching the same conclusions earlier in the follow-up (8–12 months versus 2 years).

Another important finding of our study is the observation that the rate of recurrent disease was very low (2.7%) and was not different between low- and high-risk patients. These results confirm previous report by Verburg *et al.* (13) and Tuttle *et al.* (11) and suggest that the clinical outcome of DTC patients in clinical remission after initial treatment is not correlated with the initial risk stratification.

In conclusion, our results indicate the importance of re-classifying DTC patients on the basis of the results obtained at the time of the first control after initial therapy (total thyroidectomy and radioiodine ablation), particularly in the intermediate/high-risk patients. This DRS allows to better modulate the subsequent follow-up excluding a significant number of intermediate/high-risk patients from unnecessary intensive work-up. The initial risk stratification at the time of initial treatment may still be used to dictate the need for thyroid remnant ablation but should no longer dictate the strategy of the subsequent diagnostic follow-up.

## 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 work was supported in part by grants from Ministero Italiano dell'Università e Ricerca (MIUR, 2008).

## References

- Hay ID, Bergstralh EJ, Goellner JR, Ebersold JR & Grant CS. Predicting outcome in papillary thyroid carcinoma: development of a reliable prognostic scoring system in a cohort of 1779 patients surgically treated at one institution during 1940 through 1989. *Surgery* 1993 **114** 1050–1057.
- Byar DP, Green SB, Dor P, Williams ED, Colon J, van Gilse HA, Mayer M, Sylvester RJ & van Glabbeke M. A prognostic index for thyroid carcinoma. A study of the E.O.R.T.C. Thyroid Cancer Cooperative Group. *European Journal of Cancer* 1979 **15** 1033–1041.
- Tubiana M, Schlumberger M, Rougier P, Laplanche A, Benhamou E, Gardet P, Caillou B, Travagli JP & Parmentier C. Long-term results and prognostic factors in patients with differentiated thyroid carcinoma. *Cancer* 1985 **55** 794–804. (doi:10.1002/1097-0142(19850215)55:4<794::AID-CNCR2820550418>3.0.CO;2-Z)

- 4 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)
  - 5 Mazzaferri EL & Jhiang SM. Long-term impact of initial surgical and medical therapy on papillary and follicular thyroid cancer. *American Journal of Medicine* 1994 **97** 418–428. (doi:10.1016/0002-9343(94)90321-2)
  - 6 Greene FL, Page DL, Fleming ID, Fritz A, Balch CM, Haller DG & Morrow M. *AJCC Cancer Staging Handbook: TNM Classification of Malignant Tumors*, 6th edn. DRS York: Springer-Verlag, 2002.
  - 7 Cooper DS, Doherty GM, Haugen BR, Kloos RT, Lee SL, Mandel SJ, Mazzaferri EL, McIver B, Pacini F, Schlumberger M, Sherman SI, Steward DL & Tuttle RM. 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)
  - 8 Pacini F, Schlumberger M, Dralle H, Elisei R, Smit JW & Wiersinga W & European Thyroid Cancer Taskforce. European consensus for the management of patients with differentiated thyroid carcinoma of the follicular epithelium. *European Journal of Endocrinology* 2006 **154** 787–803. (doi:10.1530/eje.1.02158)
  - 9 Tuttle RM. Risk-adapted management of thyroid cancer. *Endocrine Practice* 2008 **14** 764–774.
  - 10 Tuttle RM & Leboeuf R. Follow up approaches in thyroid cancer: a risk adapted paradigm. *Endocrinology and Metabolism Clinics of North America* 2008 **37** 419–435. (doi:10.1016/j.ecl.2008.02.008)
  - 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 DRS American Thyroid Association staging system. *Thyroid* 2010 **20** 1341–1349. (doi:10.1089/thy.2010.0178)
  - 12 Galen RS. Application of the predictive value model in the analysis of test effectiveness. *Clinics in Laboratory Medicine* 1982 **2** 685–699.
  - 13 Verburg FA, Stokkel MP, Düren C, Verkooijen RB, Mäder U, van Isselt JW, Marlowe RJ, Smit JW, Reiners C & Luster M. No survival difference after successful (131)I ablation between patients with initially low-risk and high-risk differentiated thyroid cancer. *European Journal of Nuclear Medicine and Molecular Imaging* 2010 **37** 276–283. (doi:10.1007/s00259-009-1315-6)
- 

Received 25 May 2011

Revised version received 7 July 2011

Accepted 12 July 2011