

Staging systems for follicular thyroid carcinoma : application to 171 consecutive patients treated in a tertiary referral centre.

Author

Lang, Brian Hung-Hin, Lo, Chung-Yau, Chan, Wai-Fan, Lam, King-Yin, Wan, Koon-Yat

Published

2007

Journal Title

Endocrine-Related Cancer

Rights statement

Disclaimer. This is not the definitive version of record of this article. This manuscript has been accepted for publication in Endocrine-Related Cancer, but the version presented here has not yet been copy edited, formatted or proofed. Consequently, the Society for Endocrinology accepts no responsibility for any errors or omissions it may contain.

Downloaded from

<http://hdl.handle.net/10072/17369>

Link to published version

<http://erc.endocrinology-journals.org/>

Griffith Research Online

<https://research-repository.griffith.edu.au>

Staging systems for follicular thyroid carcinoma: application to 171 consecutive patients treated in a tertiary referral centre

Brian Hung-Hin Lang, Chung-Yau Lo, Wai-Fan Chan, King-Yin Lam¹ and Koon-Yat Wan²

Division of Endocrine Surgery, Department of Surgery, University of Hong Kong Medical Centre, Queen Mary Hospital, 102 Pokfulam Road, Hong Kong SAR, China

¹Discipline of Pathology, School of Medicine, Griffith University, Gold Coast, Australia

²Department of Clinical Oncology, University of Hong Kong Medical Centre, Queen Mary Hospital, Hong Kong SAR, China

(Requests for offprints should be addressed to C-Y Lo; Email: cylo@hkucc.hku.hk)

Abstract

A number of risk-group stratification or staging systems have been found useful at stratifying patients with differentiated thyroid carcinoma into risk groups. Those identified as high risk could be subjected to more aggressive treatment, while those at low risk could be spared of such treatment. However, the best stratification system in patients with follicular thyroid carcinoma (FTC) remains unclear. Through a comprehensive MEDLINE search from 1965 to 2005, a total of 18 different staging systems were identified in the literature and 14 of them were applicable to 171 patients, with FTC managed at our institution from 1961 to 2001. Cancer-specific survivals (CSS) were calculated by Kaplan–Meier method and were compared by log-rank test. Using Cox proportional hazards analysis, the relative importance of each staging system in determining CSS was calculated by the proportion of variation in survival time explained (PVE). CSS were predicted by 13 out of the 14 staging systems significantly ($P < 0.001$). The three highest ranked staging systems by PVE were the new American Joint Committee on Cancer/Union Internationale Centre le Cancer 6th edition, tumour, node, metastases (TNM; 22.4), followed by the Clinical Class (21.2) and the metastases, age, completeness of resection, invasion, size (MACIS; 20.4). In conclusion, 13 out of the 14 presently available staging systems predicted CSS significantly in FTC. When predictability was measured by PVE, the TNM system was found to have the best predictability and thus, should be the stratification system of choice for FTC in the future.

Endocrine-Related Cancer (2007) 14 29–42

Introduction

Follicular thyroid carcinoma (FTC) is the second commonest type of thyroid carcinoma behind papillary thyroid carcinoma (PTC). It accounts for 10–30% of all differentiated thyroid carcinoma (DTC; Jukkola *et al.* 2004, Machens *et al.* 2005, Lang *et al.* 2007a). Although, FTC is less common than PTC and unlike PTC, its incidence has remained constant in the last decade, it carries significantly a poorer survival outcome than PTC (Davies & Welch 2006, Lundgren *et al.* 2006). For this reason, the role of risk-group stratification or staging systems in FTC could be of greater clinical importance than PTC, as they could be used to identify those at high risk of cancer death for more aggressive surgical and adjuvant treatment (Shaha

2004). To date, there have been a number of staging systems being described in the literature. Although they are generally applicable to patients with PTC, FTC or DTC as a group, no specific staging system has been consistently demonstrated to be superior in FTC (Cooper *et al.* 2006). Furthermore to our knowledge, only one proposed staging system has derived solely from FTC patients (Brennan *et al.* 1991).

A number of studies have been attempted to compare the predictability of these staging systems by applying them to different populations. (Davis *et al.* 1995, Brierley *et al.* 1997, Sherman *et al.* 1998, Chaplin *et al.* 1999, Steinmuller *et al.* 2000, Passler *et al.* 2003, Yildirim 2005, Kjellman *et al.* 2006) However, most of these comparative studies evaluated

staging systems in patients with either DTC or PTC. There have been relatively few studies when comparing staging systems in patients with only FTC (D'Avanzo et al. 2004a,b, Lo et al. 2005) and to our knowledge, only one previous study comparing staging systems in FTC, based on a systematic search of the literature (Passler et al. 2003).

The objectives of the present study were to comprehensively review all of the present staging systems available for thyroid carcinoma in the literature by performing a computer-assisted MEDLINE search and to look for the most predictive staging system for cancer specific survivals (CSS) in a predominantly ethnic Chinese population with FTC.

Materials and methods

A comprehensive review of the literature was performed by searching MEDLINE for relevant articles in the English Language from 1965 to 2005, indexed under the key words: thyroid carcinoma/cancer, staging, risk stratification, multivariate analysis or risk factors. The abstracts of all captured articles were read and those describing staging systems or risk-group stratifications were identified and reviewed in detail. In addition, the bibliographies of captured articles were further searched to identify the potentially relevant articles not found in the original MEDLINE search.

Patients

From 1961 to 2001, a total of 171 consecutive patients with an histologic diagnosis of FTC, who all underwent a primary surgical treatment at our institution were included in the present analysis. Over the same period, 589 patients were diagnosed with PTC and were excluded from the analysis. The majority of the FTC cohort were female (83.6%) and ethnic Chinese (96.5%). The median age was 45.0 years (range: 13.0–87.0 years). All FTC patients were initially classified into the two classical histologic subtypes; according to the extent of invasiveness (i.e. minimally invasive (or encapsulated) and widely invasive; Hedginer et al. 1988). By definition, widely invasive FTC had to demonstrate extensive areas of invasion at the microscopic level (i.e. evidence of tumour extending beyond the tumour capsule and possibly infiltrating into the affected lobe or entire gland; Hedginer et al. 1988, Rosai et al. 1992). For those belonging to the minimally invasive category, they were further subdivided into those with capsular invasion only (i.e. truly minimal invasive; Baloch & Livolsi 2002) and those with vascular invasion (i.e. angioinvasive; Rosai et al. 1992, D'Avanzo et al. 2004a,b). Among them, 61 had minimally invasive FTC, 26 had angioinvasive FTC and

84 had widely invasive FTC. In this series, Hurthle cell carcinoma ($n=22$, 12.9%) were included. To ensure consistency and accuracy, each histologic diagnosis was reconfirmed after a careful review of the retrieved slides by a dedicated pathologist (KYL) unaware of the clinical data, according to the standardized criteria approved by World Health Organization (DeLellis et al. 2004). Table 1 shows the patient characteristics of the cohort.

Table 1 Patient characteristics

Clinicopathological factors	Patients no. (%)
Gender ($n=171$)	
Male	28 (16.4)
Female	143 (83.6)
Age ($n=171$)	
< 50 years	96 (56.1)
≥ 50 years	75 (43.9)
Final histology ($n=171$)	
Follicular thyroid carcinoma	
Minimally invasive	61 (35.7)
Angioinvasive	26 (15.2)
Widely invasive	84 (49.1)
Tumour size ($n=171$)	
≤ 2 cm	44 (25.7)
> 2–4 cm	79 (46.2)
> 4 cm	48 (28.1)
Metastatic lymph nodes (or N1) ^a ($n=171$)	
Yes	12 (7.0)
No	159 (93.0)
Distant metastases ($n=171$)	
Yes	19 (11.1)
No	152 (88.9)
Extrathyroidal involvement ($n=171$)	
Yes	16 (9.4)
No	155 (90.6)
Multifocality ($n=171$)	
Yes	35 (20.5)
No	136 (79.5)
Perineural infiltration ($n=171$)	
Yes	2 (1.2)
No	169 (98.8)
Associated thyroiditis ($n=171$)	
Yes	161 (94.2)
No	10 (5.8)
Completeness of resection ($n=171$)	
Yes	164 (95.9)
No	7 (4.1)
Surgery ($n=171$)	
Total thyroidectomy	121 (70.8)
Subtotal thyroidectomy	17 (9.9)
Lobectomy	33 (19.3)
Radioiodine ablation ($n=171$)	
Yes	56 (32.7)
No	115 (67.3)
External-beam irradiation to neck ($n=171$)	
Yes	9 (5.3)
No	162 (94.7)

^aBased on the AJCC/UICC 6th edition TNM staging system.

Surgical treatment and adjuvant therapy

Details of surgical treatment, adjuvant therapy and follow-up protocol were similar to PTC and had been described previously (Lo *et al.* 2004, Lang *et al.* 2006). In brief, for those patients with a preoperative diagnosis of FTC, a total or near total thyroidectomy (defined as leaving less than one gram of thyroid tissue behind) had increasingly been the preferred procedure of choice. For those diagnosed after a lobectomy, the decision whether to perform a complete total thyroidectomy and/or to administer radioiodine (RAI) ablation afterwards was determined by known risk factors, such as the patient's age, tumour characteristics, as well as patient's preference. If patients underwent a complete total thyroidectomy within 6 months of their initial thyroid surgery, it was considered part of initial surgery. However, if reoperations were performed for residual or nodal disease 6 months after diagnosis, it was considered to be a locoregional recurrence. The routine palpation and the sampling of enlarged or suspicious lymph nodes in the central and lateral compartments was performed at the time of operation and was supplemented with the more frequent use of pre-operative ultrasonography of the neck \pm fine needle aspiration cytology. A selective neck dissection, defined as the clearance of cervical lymph nodes in levels II–V, while preserving internal jugular vein and accessory nerve was performed for cytologically- or histologically-proven lateral lymph node metastasis.

Patients with at least one or more of the following risk factors would be considered for RAI ablation 4–6 weeks after surgical treatment by T4 withdrawal: tumour size >1 cm, angio- or wide-invasion on histology, age older than 40 years, presence of extrathyroidal extension, macroscopic postoperative residual disease in the neck, and/or distant metastasis. Diagnostic whole-body I^{131} scans were performed at 8–12 weeks after RAI therapy. Three gigabecquerels (GBq) or 80 mCi I^{131} would be administered as standard ablative dose, while subsequent I^{131} therapy would be performed with 5.5 GBq (or 150 mCi). The additional 5.5 GBq I^{131} therapy would be administered periodically at 4–6 monthly intervals until uptake was no longer visible or disease progressed despite the treatment. External local radiotherapy would be given to patients with an extensive extrathyroidal tumour extension, incomplete resection and/or extra-capsular lymph node metastasis. Although the above protocol was strictly followed, individual patients' preference would be considered and respected.

Follow-up and surveillance of patients

Complete follow-up data were available for all patients. The median follow-up period for the cohort was 119 months (range: 0.5–488 months). All the patients after surgery were followed up within 4 weeks in a specialized combined surgical oncology clinic, where clinical oncologists and endocrine surgeons were present to discuss and decide on subsequent management. Follow-up visit was conducted at 3 monthly intervals in the first 2 years, 6 monthly for the subsequent 3 years and annually thereafter. Clinical examinations, chest X-ray, ultrasonography of neck and thyroglobulin levels (since 1989) were done during follow-up visits. Human recombinant thyroid-stimulating hormone was not available during the study period at our institution. Radioactive scans were done in the presence of elevated thyroglobulin level, documented nodal recurrence or radiological evidence of recurrence, or metastases. The diagnosis of distant metastases on presentation was based on the findings of histological, radiological or scintigraphic evidence and not based only on an elevated thyroglobulin level. Locoregional recurrences were frequently diagnosed by ultrasound, computed tomography (CT) or magnetic resonance imaging (MRI) and confirmed by fine needle aspiration cytology. The survival data including the cause of death were retrieved from the Hong Kong Hospital Authority territory-wide computerized medical system and from death certificates or postmortem examinations. The present study protocol was approved by the appropriate institutional review committee in accordance with the precepts established by the Helsinki Declaration.

Application of staging systems

There were a total of 18 staging systems identified from the computed-assisted MEDLINE search from 1965 to 2005. The 6th edition of the TNM (Greene *et al.* 2002) was chosen over the other editions because it was the present version. Out of the 18 systems, 14 were applicable to our FTC cohort. The four non-applicable systems were the age, grade, extent, size (AGES; Hay *et al.* 1987), the DNA ploidy, age, metastases, extent, size (DAMES; Pasiaka *et al.* 1992), the sex, age, grade (SAG; Akslen 1993) and Virgen de la Arrixaca University Hospital at Murcia staging system (Murcia; Ortiz Sebastian *et al.* 2000). The first three systems included parameters which were not available at our institution and therefore they were excluded from the present analysis. Parameters, such as the grading of PTC (in AGES), DNA ploidy (in DAMES) and nuclear atypia (in SAG) were not routinely reported by the pathologists. The Murcia

system was also not applicable because histologic variants could not be strictly applied to FTC. The 14 different staging systems were applied to all 171 FTC patients in accordance to the original description.

Statistical analysis

For each staging system, CSS were calculated by the Kaplan–Meier method and the difference between the stages and/or risk groups were compared by the log-rank test. Using Cox proportional hazards analysis, the relative importance of each staging system was determined by calculating the proportion of variation in survival time explained (PVE). PVE (%) ranges from 0 to 100 with large numbers suggesting better predictability. Therefore, the one with the largest PVE would suggest the best predictor on CSS. To determine PVE, a mathematical formula was used: $PVE = 1 - \exp(-G^2/n)$, where G^2 is the maximum likelihood ratio that is determined by the analysis of chi-square associated with the null hypothesis (i.e. that all predictor variables have coefficients of 0) and n is the total number of valid cases in the study. (Schemper & Stare 1996) $P < 0.05$ was considered to indicate statistical significance. Statistical analyses were performed using the SPSS for Windows 11.0 computer software (SPSS Inc., Chicago, IL, USA).

Results

European organization for research and treatment of cancer (EORTC; Byar et al. 1979)

The EORTC was published in 1979 and was the first ever attempt at staging all histologic types of thyroid carcinoma (including medullary and anaplastic thyroid carcinomas) under one system. This system was developed from a multivariate analysis of 507 patients from 23 European hospitals with a median follow-up of 40 months. Under this system, a prognostic score was derived and stratified into one out of the five risk groups/stages (score < 50 , 50–65, 66–83, 84–108 and > 108). Total score = (age of patient) + (12 if male) + (10 if poorly differentiated FTC) + (10 if invasion through thyroid capsule) + (15 if one distant metastasis) + (30 if two or more distant metastases). Their reported 5-year CSS for risk-group I, II, III, IV and V were 95, 80, 51, 33 and 5% respectively. Figure 1 shows CSS of 171 patients by the EORTC.

Lahey clinic (age, grade, extent, size or AMES; Cady & Rossi 1988)

The AMES staging system was developed in 1980s from a cohort of 814 DTC patients. The prognostic

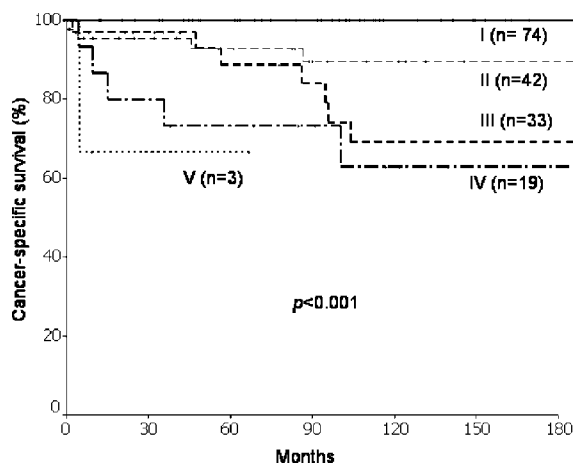


Figure 1 Cancer-specific survival of 171 follicular thyroid carcinoma by the EORTC staging system.

factors were age, distant metastases, extrathyroidal invasion and size. Both age and size were expressed as categorical variables. The cut-off point for size was 5 cm, but the cut-off point for age differed between the two sexes (41 years for men and 51 for women). Although it was developed from an expansion of previously defined risk groups, how the classification was derived remained unclear from the published reports. Low-risk = male < 41 years, or women < 51 years without distant metastases or those older patients but with intrathyroidal PTC, or minimally invasive FTC, tumour size < 5 cm and no distant metastases. Otherwise, the rest belonged to the high-risk group. Their reported mortality rate of DTC with a median follow-up of 13 years for low- and high-risk groups were 1.8 and 46% respectively. Figure 2 shows CSS of 171 FTC by the AMES.

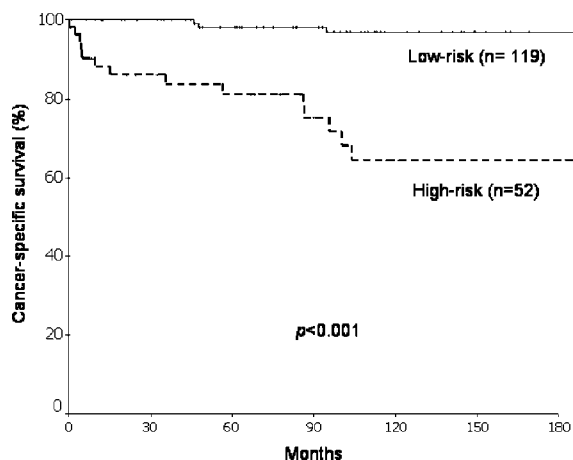


Figure 2 Cancer-specific survival of 171 follicular thyroid carcinoma by the AMES staging system.

University of Chicago (Clinical Class; DeGroot *et al.* 1990)

The Clinical Class system was developed in the 1980s from a cohort of 269 PTC patients and was initially intended for PTC only. Its applicability to FTC was validated in a separate study with a cohort of 49 patients (DeGroot *et al.* 1995). Patients were categorized into four classes based on the anatomical extent of the primary tumour. It remained unclear, how this system was derived and why age was not included in the system, despite being a significant factor in the multivariate analysis. Class I=disease limited to the thyroid gland; class II=locoregional lymph node involvement; class III=extrathyroidal tumour invasion; class IV=distant metastases. Their reported mortality rate of FTC over an average follow-up of 10.7 years (range: <1–36 years) for Clinical Class I, III, and IV were 13.9, 25 and 50% respectively. Figure 3 shows CSS of 171 FTC by the Clinical Class system.

Mayo clinic (age, invasion to blood vessels, metastases or AIM; Brennan *et al.* 1991)

This system was derived from 100 histologically confirmed FTC patients treated at the Mayo Clinic from 1946 to 1970. Based on the univariate and multivariate analyses of 100 FTC patients with a mean follow-up of 17.4 years, the authors were able to identify three independent prognostic factors and proposed a novel risk stratification system for FTC. These factors were age ≥ 50 years, vascular invasion and distant metastases. Low-risk group was referred to those with none or only one of the three independent factors, whereas high-risk group was those with two or

more independent factors. Their reported 5- and 20-year CSS in FTC for low-risk group were 99 and 86% respectively and for high-risk group were 47 and 8% respectively. Figure 4 shows CSS of 171 FTC by the AIM system.

Mayo clinic (metastases, age, completeness of resection, invasion, size or MACIS; Hay *et al.* 1993)

This system was developed as an alternative to the AGES system because grade for PTC was not available in most centers. It was based solely on PTC patients. The system was derived using a cohort of 1779 patients divided into two equal-sized groups based on the date of diagnosis: 1957–1972 and 1973–1988. A Cox model analysis using a stepwise variable selection led to a prognostic model derived from the former cohort. The prognostic scoring system was validated with the latter cohort. Under this system, a prognostic score was derived and the patients were stratified into four risk groups (<6.00, 6.00–6.99, 7.00–7.99, >8.00). Total score = 3.1 (if aged ≤ 39 years) or $0.08 \times$ age (if aged ≥ 40 years) + $(0.3 \times$ tumour size in cm) + 1 (if not completely resected) + 1 (if locally invasive) + 3 (if distant metastases). Their reported 20-year mortality rate for risk-group I, II, III and IV were 0.9, 11.3, 44.4 and 76.5% respectively. Figure 5 shows CSS of 171 FTC patients by the MACIS.

Ohio State University (OSU; Mazzaferri & Jhiang 1994)

This system was developed in 1994 from a multivariate analysis of 1355 DTC patients. Patients were divided into four stages based on the criteria of tumour size, lymph node involvement, multifocality, local tumour

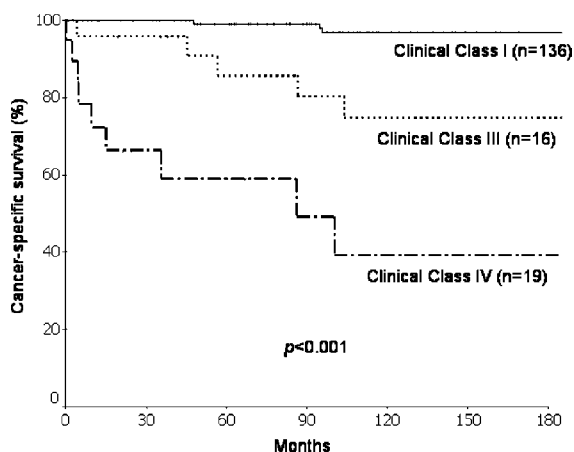


Figure 3 Cancer-specific survival of 171 follicular thyroid carcinoma by the Clinical Class system (no patient in Class II).

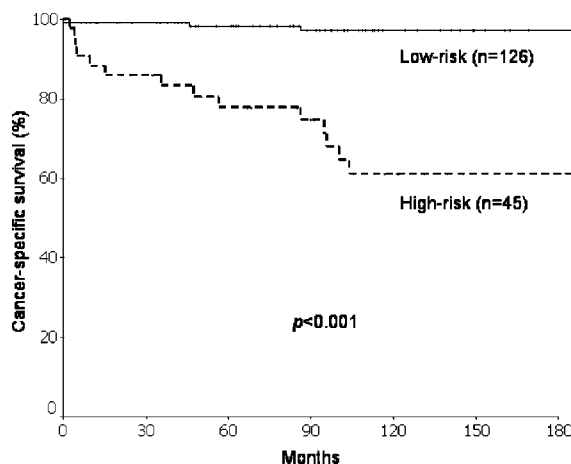


Figure 4 Cancer-specific survival of 171 follicular thyroid carcinoma by the AIM staging system.

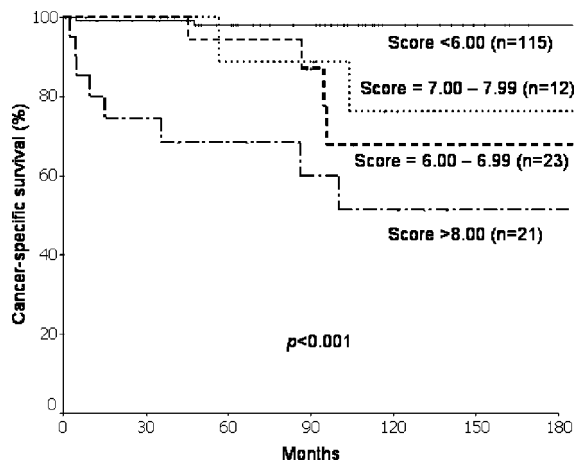


Figure 5 Cancer-specific survival of 171 follicular thyroid carcinoma by the MACIS staging system.

invasion and distant metastases. Staging criteria were as follows: Stage I – tumour smaller than 1.5 cm; Stage II – tumour size between 1.5 and 4.4 cm or presence of cervical lymph node metastases, or more than three intrathyroidal foci of tumour; Stage III – tumour at least 4.5 cm or presence of extrathyroidal invasion; Stage IV – distant metastases. Unlike most other systems, it uses multifocality as a prognostic factor. Their reported mortality in DTC over a median follow-up of 15.7 years for risk-group I, II, III and IV were 0, 6, 14 and 65% respectively. Figure 6 shows CSS of 171 FTC by the OSU.

Noguchi thyroid clinic staging system (Noguchi et al. 1994)

This system was published in 1994 after a multivariate analysis of 2192 PTC patients over a 24-year period.

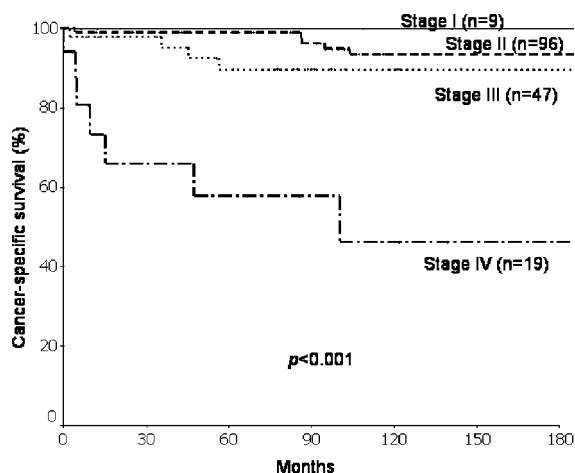


Figure 6 Cancer-specific survival of 171 follicular thyroid carcinoma by the OSU staging system.

Patients were stratified into three risk groups (excellent, intermediate or poor) based on gender, age, tumour size, extrathyroidal extension and gross lymph node metastases. It is unclear how these risk groups were derived from the original publication. Excellent-risk group = all male aged ≤ 45 , male aged ≤ 60 and without gross lymph node metastases, all female aged ≤ 50 or female aged 51–55 and without lymph node metastases; intermediate-risk group = male aged > 60 and without gross lymph node metastases, male aged 46–55 and with gross lymph node metastases, female aged 56–65 and no gross lymph node metastases, female aged > 65 and tumour size $< 30\text{mm}$ or female aged 50–55 with gross lymph node metastases; poor-risk group = male aged > 55 with gross lymph node metastases, female not included in the above two risk groups. Their reported 10-year CSS in male PTC for excellent-, intermediate- and poor-risk groups were 98.4, 90.1 and 74.4% respectively. Their reported 10-year CSS in female PTC for excellent-, intermediate- and poor-risk groups were 99.3, 96.4 and 88.8% respectively. Figure 7 shows CSS of 171 FTC patients by the Noguchi system.

Memorial sloan kettering (grade, age, metastases, extent, size or GAMES; Shaha et al. 1994)

It was published in 1994 after an analysis of 1038 DTC patients. Under this system, patients were stratified into low-, intermediate- and high-risk categories. Age of 45 and size of 4 cm were the cut-off points for continuous variables. From the original article, it was not clear how the actual stage groupings were determined. Low risk = patients aged < 45 , with no distant metastases, with tumour size $< 4\text{ cm}$ and PTC on histology; high

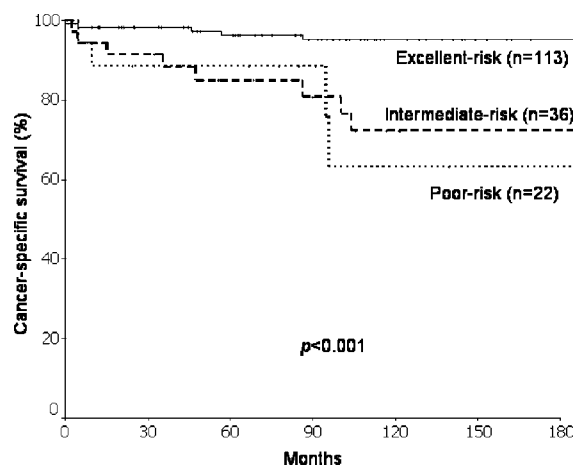


Figure 7 Cancer-specific survival of 171 follicular thyroid carcinoma by the Noguchi staging system.

risk=patients aged ≥ 45 , with distant metastases, tumour size > 4 cm, FTC on histology; the rest not belonging to the two risk groups were assigned as intermediate-risk group. Their reported 5-year CSS of DTC patients in low-, intermediate- and high-risk groups were 100, 96 and 72% respectively. Figure 8 shows CSS of 171 patients by the GAMES system.

University of Münster (Münster; Lerch *et al.* 1997)

This system was published in 1995 after an analysis of 500 DTC patients. Of note, this series comprised a high proportion of FTC (almost 40%). Patients were stratified into low- and high-risk groups. Those tumours with extrathyroidal invasion and/or distant metastases were classified as the high-risk group. The rest would be classified as the low-risk group. Their reported mortality rate in DTC over a median follow-up of 6 years for low- and high-risk groups were 0 and 11.2% respectively. Figure 9 shows CSS of 171 FTC by the Münster system.

National Thyroid Cancer Treatment Cooperative Study (NTCTCS; Sherman *et al.* 1998)

The NTCTCS registry was a multicenter thyroid cancer registry established in 1986 with an aim of creating a broadly applicable staging classification in predicting the outcome for all the histological types of thyroid carcinoma. Parameters selected in this system were age, tumour size, tumour type, extrathyroidal invasion, lymph node and distant metastases. Several factors tend to predominate in the assignment of tumour stage for each histologic type. For PTC, significant factors include age, size, extrathyroidal

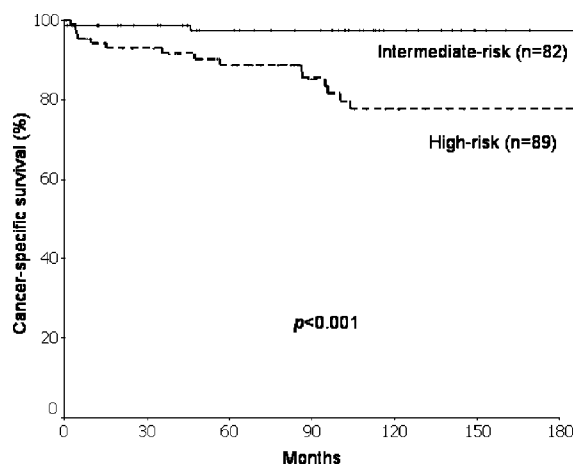


Figure 8 Cancer-specific survival of 171 follicular thyroid carcinoma by the GAMES staging system (no patient in low-risk group).

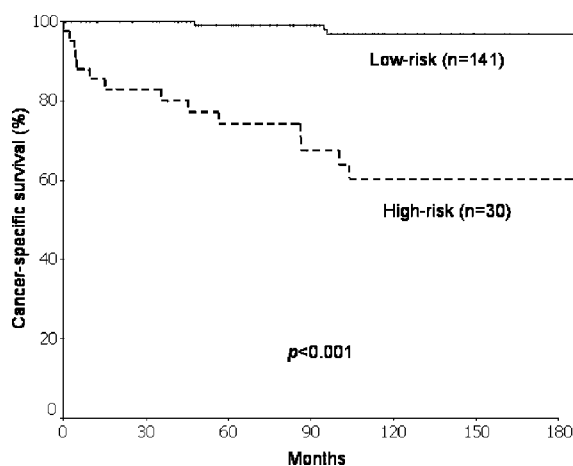


Figure 9 Cancer-specific survival of 171 follicular thyroid carcinoma by the Münster staging system.

invasion and metastases. For FTC, they include age, size, distant metastases and poor differentiation. Although it remains unclear how the actual staging system was derived, it was validated prospectively with 1607 patients recruited from 14 different U.S. Institutions. Their reported 5-year CSS in FTC for Stage I, II, III and IV were 100, 100, 81.9 and 37.1% respectively. Figure 10 shows CSS of 171 FTC by the NTCTCS system.

University of Alabama and M D Anderson (UAB&MDA; Beeken *et al.* 2000)

This system stratified PTC patients into low-, intermediate- and high-risk groups. Although three prognostic factors were found to be significant, only two prognostic factors (age and distant metastases)

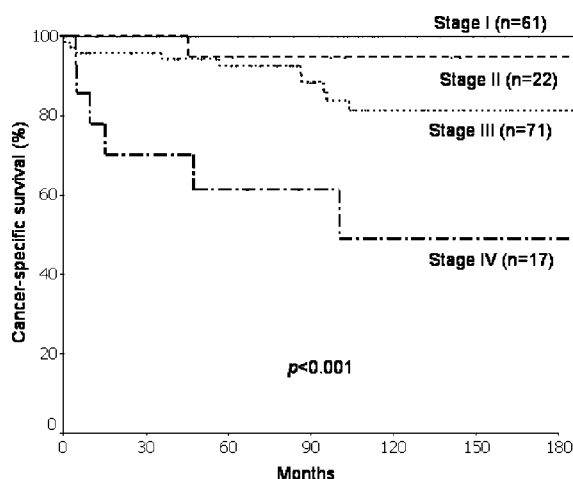


Figure 10 Cancer-specific survival of 171 follicular thyroid carcinoma by the NTCTCS staging system.

actually affected the final tumour stage. Low-risk = those aged ≤ 50 years and without distant metastases; intermediate-risk = those aged > 50 years without metastases; high-risk = those with distant metastases irrespective of age. Their reported 5-year CSS in DTC for low-, intermediate- and high-risk groups were 100, 90 and 40% respectively. Figure 11 shows CSS of 171 FTC by the UAB&MDA system.

AJCC/UICC 6th edition TNM staging system (TNM; Greene et al. 2002)

The TNM staging system was first described in the 1940s and the 6th edition came into use in January 2003. Like the previous editions, TNM is a system that describes the anatomical extent of the primary tumour (T), the involvement of regional lymph nodes (N) and distant metastasis (M). Although the system is applicable to all histologies of thyroid carcinoma, the stage grouping varies with different histologic types. PTC and FTC are being staged in the same way. It is the only staging system that regularly undergoes revision in order to keep up with prevailing changes in the field of thyroid carcinoma. Their reported 5-year CSS in FTC for Stage I, II, III and combined stage IV were 100, 100, 79.4 and 47.1% respectively. Figure 12 shows CSS of 171 FTC by the TNM.

Cancer Institute Hospital in Tokyo (CIH; Sugitani et al. 2004)

This system was derived from a multivariate analysis of 604 PTC patients and there were four prognostic parameters incorporated into the system, namely age (< 50 or ≥ 50), distant metastases, extrathyroidal extension and large nodal metastases (≥ 3 cm).

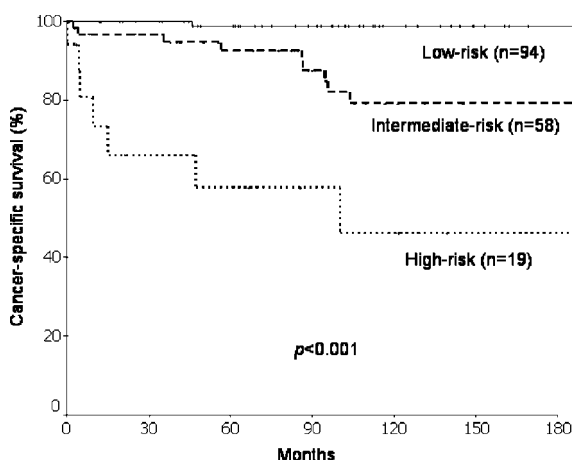


Figure 11 Cancer-specific survival of 171 follicular thyroid carcinoma by the UAB&MDA staging system.

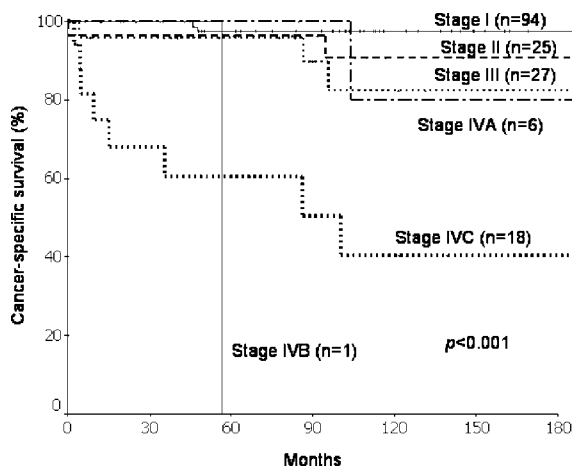


Figure 12 Cancer-specific survival of 171 follicular thyroid carcinoma by the AJCC/UICC 6th edition TNM staging system.

Under this system, patients were stratified into two risk groups (low- and high-risk). High-risk = those aged < 50 but with distant metastases or those aged ≥ 50 and with at least one unfavorable prognostic factor such as large (≥ 3 cm) nodal metastases, extrathyroidal extension or distant metastases. The rest not belonging to the high-risk group would be in the low-risk group. Their reported 10-year CSS in PTC for low- and high-risk groups were 99 and 69% respectively. Figure 13 shows CSS of FTC by the CIH system.

Ankara Oncology Training and Research Hospital (Ankara; Yildirim 2005)

This system was developed from a cohort of 347 DTC patients. Univariate and multivariate prognostic factor analyses were carried out and four risk groups

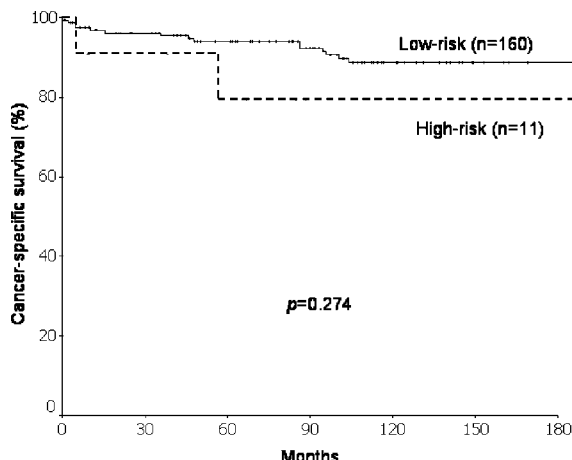


Figure 13 Cancer-specific survival of 171 follicular thyroid carcinoma by the CIH staging system.

(very low, low, high and very high) were identified by the logistic regression equation. The proposed model was validated using the split sample method. The authors came up with two equations; the first included clinicopathological factors only (the pretreatment formula) and the other included clinicopathological as well as treatment factors (the post-treatment formula). However, the authors did not clearly define the criteria of risk groups in the post-treatment formula. The treatment factors included whether a patient has undergone a total thyroidectomy as opposed to subtotal and has received adjuvant RAI. It is the only staging system, which incorporated treatment-related factors. Pretreatment score = $\exp [(0.2 \times \text{tumour size}) + (1 \text{ if age} > 45 \text{ years}) + (0.7 \text{ if angioinvasion}) + (1 \text{ if distant metastasis})]$ and pretreatment probability (P) = $(\text{score}) / (1 + \text{score})$. Patients were then stratified into four risk groups, according to the value of P ($\leq 55\%$ for very-low risk; from 56 to 85% for low risk; from 86 to 95% for high risk; $\geq 96\%$ for very high risk). Their reported 10-year overall survival in DTC for very-low, low-, high- and very high-risk groups were 100, 88, 30 and 5% respectively. Figure 14 shows CSS of 171 FTC by the Ankara system.

Evaluation and comparison of staging systems

In terms of CSS, 13 out of the 14 staging systems demonstrated highly significant differences between the different stages and/or risk groups ($P < 0.001$). The CIH staging system was the only system, which failed to reproduce significant differences in CSS between low- and high-risk groups ($P = 0.274$). Table 2 shows the 5-, 10- and 15-year CSS of patients when being

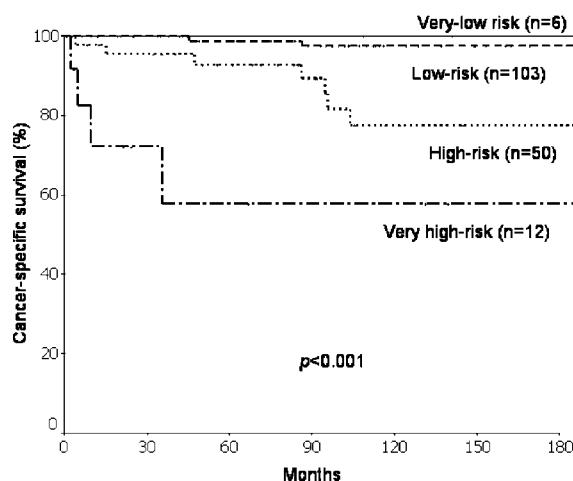


Figure 14 Cancer-specific survival of 171 follicular thyroid carcinoma by the Ankara staging system.

staged by each of the 14 staging systems. Table 3 shows the PVE and the relative ranking of the 14 applicable staging systems for FTC. In terms of PVE, the TNM had the highest PVE value of 22.4% and this was followed by the Clinical Class (21.2%) and MACIS (20.4%).

Discussion

Cancer staging is an essential and integral part of cancer management and a predictive staging system not only provides an accurate prognostic information to clinicians and their patients, but also helps to facilitate exchange of cancer information between different medical centers (Greene 2005, Cooper *et al.* 2006). However, it remains unclear, which of the presently available staging systems is most predictive for CSS in patients with FTC.

From the MEDLINE search, there were a total of 18 different staging systems described for thyroid carcinoma. Eight of these were derived solely from one histologic type, namely PTC, and they were the AGES, Clinical Class, DAMES, MACIS, SAG, Noguchi, Murcia and CIH (Hay *et al.* 1987, 1993, DeGroot *et al.* 1990, Pasiaka *et al.* 1992, Akslen 1993, Noguchi *et al.* 1994, Ortiz Sebastian *et al.* 2000, Sugitani *et al.* 2004). The other nine systems were either derived from DTC (i.e. PTC and FTC together; Cady & Rossi 1988, Mazzaferri & Jhiang 1994, Shaha *et al.* 1994, Lerch *et al.* 1997, Beeken *et al.* 2000, Greene *et al.* 2002, Yildirim 2005) or derived from all histologic types of thyroid carcinoma, including medullary and anaplastic thyroid carcinoma (Byar *et al.* 1979, Sherman *et al.* 1998). Interestingly, there was only one of the 18 systems (i.e. AIM) solely derived from and specific for FTC (Brennan *et al.* 1991). Perhaps, this may be because the number of FTC accumulated within a single institution is relatively fewer than that of PTC. This may also be the reason why despite the distinct differences between PTC and FTC as shown in previous studies (Mazzaferri & Jhiang 1994, Holzer *et al.* 2000, Hundahl *et al.* 2000, Chow *et al.* 2002, Passler *et al.* 2004, Besic *et al.* 2005, Machens *et al.* 2005, Lang *et al.* 2007a), the majority of the presently available staging systems including the 6th edition TNM system still make no distinction in staging between PTC and FTC.

From the results, there were many similarities and differences noted in allocation of patients into different risk groups and CSS between those reported in the original article and in the present study. However, these results should be interpreted cautiously as most of the original articles did not report the survival data

Table 2 Cancer-specific survival of follicular thyroid carcinoma in 14 different staging systems

Staging system	No. of patients	No. of deaths	Cancer-specific survival (CSS)		
			5-year (%)	10-year (%)	15-year (%)
EORTC	171	20			
I	74	0	100.0	100.0	100.0
II	42	4	92.7	89.6	89.6
III	33	7	88.7	69.2	69.2
IV	19	8	73.3	62.9	62.9
V	3	1	66.7	66.7	66.7
AMES	171	20			
Low-risk	119	3	98.1	97.0	97.0
High-risk	52	17	81.2	64.4	64.4
Clinical Class	171	20			
I	136	5	99.1	96.9	96.9
II	0	0	n/a	n/a	n/a
III	16	4	85.8	74.6	74.6
IV	19	11	58.8	39.3	39.3
AIM	171	20			
Low-risk	126	3	98.4	97.3	97.3
High-risk	45	17	77.9	61.2	61.2
MACIS	171	20			
I	115	2	98.2	98.2	98.2
II	23	4	94.4	67.8	67.8
III	12	2	88.9	76.2	76.2
IV	21	12	68.4	51.3	51.3
OSU	171	20			
I	9	0	100.0	100.0	100.0
II	96	4	99.0	93.5	93.5
III	47	5	89.7	89.7	89.7
IV	19	11	57.8	46.2	46.2
Noguchi	171	20			
Excellent	113	5	96.4	95.3	95.3
Intermediate	36	8	85.0	72.4	72.4
Poor	22	7	88.5	63.2	63.2
GAMES	171	20			
Low-risk	0	0	n/a	n/a	n/a
Intermediate	82	2	97.6	97.6	97.6
High-risk	89	18	88.9	77.7	77.7
Münster	171	20			
Low-risk	141	6	99.1	96.9	96.9
High-risk	30	14	74.1	60.3	60.3
NTCTCS	171	20			
I	61	0	100.0	100.0	100.0
II	22	1	94.7	94.7	94.7
III	71	11	92.6	81.4	71.4
IV	17	8	61.4	49.1	49.1
UAB&MDA	171	20			
Low-risk	94	0	98.9	98.9	98.9
Intermediate	58	9	92.6	79.2	79.2
High-risk	19	11	61.4	49.1	49.1
TNM	171	20			
I	94	2	97.5	97.5	97.5
II	25	2	96.4	90.8	90.8
III	27	3	95.8	82.4	82.4
IVA	6	1	100.0	80.0	80.0
IVB	1	1	0.00	0.00	0.00
IVC	18	11	60.4	40.3	40.3

(continued)

Table 2 continued

Staging system	No. of patients	No. of deaths	Cancer-specific survival (CSS)		
			5-year (%)	10-year (%)	15-year (%)
CIH ^a	171	20			
Low-risk	160	17	94.2	88.9	88.9
High-risk	11	3	79.5	79.5	79.5
Ankara	171	20			
Very-low	6	0	100.0	100.0	100.0
Low-risk	103	3	98.8	97.5	97.5
High-risk	50	11	92.9	77.6	77.6
Very-high	12	6	57.8	57.8	57.8

Abbreviations: EORTC, European Organization for Research and Treatment of Cancer; AMES, age, metastases, extent, size; AIM, age, invasion to blood vessel, metastasis; MACIS, metastases, age, completeness of surgery, invasion, size; OSU, Ohio State University; Noguchi, Noguchi Thyroid Clinic; GAMES, Grade, Age, Metastases, Extent, Size; Münster, University of Munster; NTCTCS, National Thyroid Cancer Treatment Cooperative Study; UAB&MDA, University of Alabama and M D Anderson; TNM, tumour, node, metastasis; CIH, Cancer Institute Hospital (Tokyo); Ankara, Ankara Oncology Training and Research Hospital (Turkey); n/a= not applicable.

^aNot statistically significant by log-rank test ($P=0.274$).

separately for PTC and FTC. As shown in previous studies, there were potential biases if PTC and FTC were not reported separately because of the relative dominance of PTC in DTC (Mazzaferri & Jhiang 1994, Holzer *et al.* 2000, Hundahl *et al.* 2000, Chow *et al.* 2002, Passler *et al.* 2004, Machens *et al.* 2005, Lang *et al.* 2007a). Only the AIM, Clinical Class, NTCTCS and 6th edition TNM systems reported 5-year actuarial survival data separately for FTC. In EORTC, the CSS of all the histologic types were reported collectively as a group. The CSS reported by Byar *et al.* (1979) were significantly worse than those of our cohort in all risk groups because they had included patients with other less differentiated thyroid carcinoma, which are known to have a poorer outcome. In AMES, OSU, GAMES, Munster, UAB&MDA and Ankara, the survival outcome of DTC were reported collectively as a group. Again, this made comparison difficult and confusing because PTC and FTC exhibited different clinical behaviour and outcome (Mazzaferri & Jhiang 1994, Holzer *et al.* 2000, Hundahl *et al.* 2000, Chow *et al.* 2002, Passler *et al.* 2004, Machens *et al.* 2005, Lang *et al.* 2007a) Apart from this, some systems such as AMES, Clinical Class, OSU and Munster only reported cancer-related mortality rate instead of the preferred actuarial CSS. This made it difficult for direct comparison because mortality rate depends on the length of follow-up. The length of follow-up in these studies were significantly different from that of the present study. In Ankara, only actuarial overall survival rates were reported instead of the preferred actuarial CSS rates.

Although, there had been numerous studies aimed at comparing the predictability of staging systems for DTC, the majority of these studies had not been analysed and reported FTC specifically (Sherman *et al.* 1998, Chaplin *et al.* 1999, Steinmuller *et al.* 2000, Yildirim 2005, Kjellman *et al.* 2006). To date, only four studies had specifically compared the staging systems in patients with FTC. The first comparative study was reported by Davis *et al.* (1995), comparing three commonly-used staging systems (the EORTC, AGES and AMES; Davis *et al.* 1995). The authors concluded that the AGES and EORTC were the two best systems at stratifying FTC patients into low- and

Table 3 Proportion of variation explained (PVE) and ranking of all 14 applicable staging systems for follicular thyroid carcinoma

Staging system	Follicular thyroid carcinoma (n=171)	
	PVE (%)	Ranking
EORTC	18.9	5
AMES	18.4	7
Clinical Class	21.2	2
AIM	20.0	4
MACIS	20.4	3
OSU	12.7	11
Noguchi	11.0	13
GAMES	11.2	12
Munster	17.8	9
NTCTCS	18.4	8
UAB&MDA	18.7	6
TNM	22.4	1
CIH	0.7	14
Ankara	14.3	10

high-risk groups. Their conclusion was based on the magnitude of cancer-specific mortality ratio between low- and high-risk groups. The second study was reported by D'Avanzo *et al.* (2004a,b) and it utilized the same statistical technique as the present study, namely the PVE (D'Avanzo *et al.* 2004a,b). In that study, the authors found that of the five commonly used staging systems (i.e the 5th edition TNM, MACIS, EORTC, AGES and AMES), the MACIS had the highest PVE value and therefore was the best predictive staging system for FTC. The third study was reported by Passler *et al.* (2003) and was carried out at the University of Vienna in Austria where the majority of patients came from an endemic goiter region (Passler *et al.* 2003). Passler *et al.* (2003) were able to retrieve 14 staging systems in a 30-year MEDLINE search, but could only manage to apply nine staging systems to their 440 DTC patients. Like the present study, Passler *et al.* (2003) were not only unable to apply the AGES, DAMES and SAG systems but also the NTCTCS and Murcia systems, because of the unavailability of required variables at their center. They found that the EORTC and the 5th edition TNM were the two highest ranked staging systems for FTC. Although the findings by Passler *et al.* (2003) appeared to concur with those of the present study, there was a slight difference in the edition of TNM system. In the present study, the present or the 6th edition TNM was used instead of the 5th edition and also the EORTC was found to be ranked below the Clinical Class and MACIS. The fourth study was reported from our institution (Lo *et al.* 2005). However, the study only evaluated four commonly-used staging systems, namely the 5th edition TNM, MACIS, AMES and Clinical Class, by calculating the PVE for each system. Although TNM was reported to be the most predictive system as in the present study, the 5th edition was used instead of the 6th edition TNM. The potential changes in predictability between the 5th vs 6th edition TNM for FTC remain undefined and need further evaluation (Lang *et al.* 2007b).

From the relatively low PVE values obtained in the present study, it is clear that all 14 staging systems had a lower than expected predictability for CSS. However, as compared to the two previous comparisons, these PVE values were within the range reported by previous studies (Brierley *et al.* 1997, Passler *et al.* 2003, D'Avanzo *et al.* 2004a,b). In fact, a large multi-center study using PVE as a measurement had reported even lower values (Sherman *et al.* 1998). Undoubtedly, there is room for improvement as none of the examined staging systems was able to account for the small number of cancer-related death in the so-called the

low-risk group (see Table 16). Perhaps, more powerful prognostic biological factors and molecular markers could be added to existing staging systems in the future for improving survival prediction (Sherman *et al.* 1998, D'Avanzo *et al.* 2004a,b, Kebebew *et al.* 2006).

However, when choosing the most appropriate staging system for clinical application, apart from evaluating its predictability, practicality, reproducibility and applicability are other important qualities, which require further evaluation. As seen in the present and in other previous studies (Brierley *et al.* 1997, Passler *et al.* 2003), staging systems, which incorporated less widely accepted variables such as PTC grading, DNA analysis and nuclear atypia were generally not applicable. Also staging systems such as the EORTC, MACIS, Noguchi, NTCTCS and Ankara would appear over complicated for daily clinical usage and so their practicality had been questioned (Brierley *et al.* 1997, Passler *et al.* 2003). Although the present or 6th edition TNM may also appear complicated at first due to the increased number of pT and pN categories when compared with the 5th edition TNM, the new stage groupings have enhanced its usability and practicality (Wittekind *et al.* 2003).

One of the limitations with the present study is that there is no universally accepted and objective measurement for predictability. Apart from PVE, a number of other statistical methods had been put forward but none had been shown to be superior to PVE (Brierley *et al.* 1997). To date, the PVE remained the most accepted measurement of predictability and had been used extensively in many recent comparative studies. (D'Avanzo *et al.* 2004a,b, Barbet *et al.* 2005, Lo *et al.* 2005, Yildirim 2005, Boulos *et al.* 2006, Lang *et al.* 2007b)

In conclusion, out of the 18 staging systems presently available in the systematic review of the literature, 14 were applied successfully to the FTC cohort and 13 out of the 14 systems were able to predict CSS significantly in patients with FTC ($P < 0.001$). The three highest ranked staging systems by PVE were the 6th edition TNM, Clinical Class and MACIS systems. Based on our findings, the 6th TNM should be the stratification system of choice for FTC in the future.

Funding

The authors declare that there is no conflict of interest that would prejudice the impartiality of this scientific work.

References

- Akslen LA 1993 Prognostic importance of histologic grading in papillary thyroid carcinoma. *Cancer* **72** 2680–2685.
- Baloch ZW & Livolsi VA 2002 Pathology of the thyroid gland. In *Endocrine Pathology*, pp 61–88. Eds VA Livolsi & SL Asa. New York: Churchill Livingstone.
- Barbet J, Campion L, Kraeber-Bodere F, Chatal JF & GTE Study Group 2005 Prognostic impact of serum calcitonin and carcinoembryonic antigen doubling times in patients with medullary thyroid carcinoma. *Journal of Clinical Endocrinology and Metabolism* **90** 6077–6084.
- Beeken S, Roye D, Weiss H, Sellers M, Urist M & Diethelm A 2000 Extent of surgery for intermediate-risk well-differentiated thyroid cancer. *American Journal of Surgery* **179** 51–56.
- Besic N, Zgajnar J, Hocevar M & Frkovic-Grazio S 2005 Is patient's age a prognostic factor for follicular thyroid carcinoma in the TNM classification system? *Thyroid* **15** 439–448.
- Boulos DL, Groome PA, Brundage MD, Siemens DR, Mackillop WJ, Meaton JP, Schulze KM & Rohland SL 2006 Predictive validity of five comorbidity indices in prostate carcinoma patients with curative intent. *Cancer* **106** 1804–1814.
- Brennan MD, Bergstralh EJ, van Heerden JA & McConahey WM 1991 Follicular thyroid cancer treated at the Mayo Clinic, 1946 through 1970: initial manifestations, pathologic findings, and outcome. *Mayo Clinic Proceedings* **66** 11–22.
- Brierley JD, Panzarella T, Tsang RW, Gospodarowicz MK & O'Sullivan B 1997 A comparison of different staging systems, predictability of patient outcome. Thyroid carcinoma as an example. *Cancer* **79** 2414–2423.
- Byar DP, Green SB, Dor P, Williams ED, Colon J, van Gilse HA, Mayer M, Sylvester RJ & van Glabbeke M 1979 A prognostic index for thyroid carcinoma: a study of the EORTC Thyroid Cancer Cooperative Group. *European Journal of Cancer* **15** 1033–1041.
- Cady B & Rossi R 1988 An expanded view of risk-group definition in differentiated thyroid carcinoma. *Surgery* **104** 947–953.
- Chaplin JM, O'Brien CJ, McNeil EB & Haghghi K 1999 Application of prognostic scoring systems in differentiated thyroid carcinoma. *Australian and New Zealand Journal of Surgery* **69** 625–628.
- Chow SM, Law SC, Au SK, Leung TW, Chan PT, Mendenhall WM & Lau WH 2002 Differentiated thyroid carcinoma: comparison between papillary and follicular carcinoma in a single institute. *Head and Neck* **24** 670–677.
- Cooper DS, Doherty GM, Haugen BR, Kloos RT, Lee SL, Mandel SJ, Mazzaferri EL, McIver B, Sherman SI & Tuttle RM 2006 Management guidelines for patients with thyroid nodules and differentiated thyroid carcinoma. *Thyroid* **16** 109–142.
- D'Avanzo A, Ituarte P, Treseler P, Kebebew E, Wu J, Wong M, Duh QY, Siperstein AE & Clark OH 2004a Prognostic scoring systems in patients with follicular thyroid cancer: a comparison of different staging systems in predicting the patient outcome. *Thyroid* **14** 453–458.
- D'Avanzo A, Treseler P, Ituarte PH, Wong M, Streja L, Greenspan FS, Siperstein AE, Duh QY & Clark OH 2004b Follicular thyroid carcinoma: histology and prognosis. *Cancer* **15** 1123–1129.
- Davies L & Welch HG 2006 Increasing incidence of thyroid cancer in the United States, 1973–2002. *Journal of the American Medical Association* **295** 2164–2167.
- Davis NL, Bugis SP, McGregor GI & Germann E 1995 An evaluation of prognostic scoring systems in patients with follicular thyroid cancer. *American Journal of Surgery* **170** 476–480.
- DeGroot LJ, Kaplan EL, McCormick M & Straus FH 1990 Natural history, treatment, and course of papillary thyroid carcinoma. *Journal of Clinical Endocrinology and Metabolism* **71** 414–424.
- DeGroot LJ, Kaplan EL, Shukla MS, Salti G & Straus FH 1995 Morbidity and mortality in follicular thyroid cancer. *Journal of Clinical Endocrinology and Metabolism* **80** 2946–2953.
- DeLellis RA, Lloyd RV, Heitz PU & Eng C 2004 *World Health Organization Classification of Tumours: Pathology and Genetics of Tumours of Endocrine Organs*. France: IARC (International Agency for Research on Cancer) Press.
- Greene FL 2005 Cancer staging, prognostic factors, and our surgical challenges. *American Surgeon* **71** 615–620.
- Greene FL, Page DL, Fleming ID, Fritz AG, Balch CM, Haller DG & Morrow M 2002 *AJCC Cancer Staging Handbook: TNM Classification of Malignant Tumors*. 6. New York: Springer-Verlag, Inc.
- Hay ID, Grant CS, Taylor WF & McConahey WM 1987 Ipsilateral lobectomy versus bilateral lobar resection in papillary thyroid carcinoma: a retrospective analysis of surgical outcome using a novel prognostic scoring system. *Surgery* **102** 1088–1095.
- Hay ID, Bergstralh EJ, Goellner JR, Ebersold JR & Grant CS 1993 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* **114** 1050–1058.
- Hedginer CE, Williams ED & Sobin LH 1988 In *Histological Typing of Thyroid Tumours, International Histological Classification of Tumors*, vol 11, pp 7–68. Ed CE Hedginer., Berlin: Springer-Verlag.
- Holzer S, Reiners C, Mann K, Bamberg M, Rothmund M, Dudeck J, Stewart AK & Hundahl SA 2000 Patterns of care for patients with primary differentiated carcinoma of the thyroid gland treated in Germany during 1996. U.S. and German Thyroid Cancer Group. *Cancer* **89** 192–201.
- Hundahl SA, Cady B, Cunningham MP, Mazzaferri E, McKee RF, Rosai J, Shah JP, Fremgen AM, Stewart AK & Holzer S 2000 Initial results from a prospective cohort study of 5583 cases of thyroid carcinoma treated in the United States during 1996. U.S. and German Thyroid

- Cancer Study Group. An American College of Surgeons Commission on Cancer Patient Care Evaluation Study. *Cancer* **89** 202–217.
- Jukkola A, Bloigu R, Ebeling T & Salmela P 2004 Prognostic factors in differentiated thyroid carcinomas and their implications for current staging classifications. *Endocrine-Related Cancer* **11** 571–579.
- Kebebew E, Peng M, Reiff E & McMillan A 2006 Diagnostic and extent of disease multigene assay for malignant thyroid neoplasms. *Cancer* **106** 2592–2597.
- Kjellman P, Zedenius J, Lundell G, Backdahl M, Farnebo LO, Hambeger B, Larsson C & Wallin G 2006 Predictors of outcome in patients with papillary thyroid carcinoma. *European Journal of Surgical Oncology* **32** 345–352.
- Lang BH, Lo CY, Chan WF, Lam KY & Wan KY 2006 Classical and follicular variant of papillary thyroid carcinoma: a comparative study on clinicopathological features and long-term outcome. *World Journal of Surgery* **30** 752–758.
- Lang BH, Lo CY, Chan WF, Lam KY & Wan KY 2007a Prognostic factors in papillary and follicular thyroid carcinoma: their implications for cancer staging. *Annals of Surgical Oncology* (In Press).
- Lang BH, Lo CY, Chan WF, Lam KY & Wan KY 2007b Restaging of differentiated thyroid carcinoma by the 6th edition AJCC/UICC TNM staging system: stage migration and predictability. *Annals of Surgical Oncology* (In Press).
- Lerch H, Schober O, Kuwert T & Saur HB 1997 Survival of differentiated thyroid carcinoma studied in 500 patients. *Journal of Clinical Oncology* **15** 2067–2075.
- Lo CY, Chan WF, Lam KY & Wan KY 2004 Optimizing the treatment of AMES high-risk papillary thyroid carcinoma. *World Journal of Surgery* **28** 1103–1109.
- Lo CY, Chan WF, Kam KY & Wan KY 2005 Follicular thyroid carcinoma: the role of histology and staging systems in predicting survival. *Annals of Surgery* **242** 708–715.
- Lundgren CI, Hall P, Dickman PW & Zedenius J 2006 Clinically significant prognostic factors for differentiated thyroid carcinoma: a population-based, nested case-control study. *Cancer* **106** 524–531.
- Machens A, Holzhausen HJ & Dralle H 2005 The prognostic value of primary tumor size in papillary and follicular thyroid carcinoma. *Cancer* **103** 2269–2273.
- Mazzaferri EL & Jhiang SM 1994 Long-term impact of initial surgical and medical therapy on papillary and follicular thyroid cancer. *American Journal of Medicine* **97** 418–428.
- Noguchi S, Murakami N & Kawamoto H 1994 Classification of papillary cancer of the thyroid based on prognosis. *World Journal of Surgery* **18** 552–558.
- Ortiz Sebastian S, Rodriguez Gonzalez JM, Parilla Paricio P, Sola Perez J, Perez Flores D, Pinero Madrona A, Ramirez Romero P & Tebar FJ 2000 Papillary thyroid carcinoma. Prognostic index for survival including the histological variety. *Archives of Surgery* **135** 272–277.
- Pasieka JL, Zedenius J, Auer G, Gremelius L, Hoog A, Lundell G, Wallin G & Backdahl M 1992 Addition of nuclear DNA content to the AMES risk-group classification for papillary thyroid carcinoma. *Surgery* **112** 1154–1160.
- Passler C, Prager G, Scheuba C, Prager G, Kaczirek K, Kaserer K, Zettinig G & Niederle B 2003 Application of staging systems for differentiated thyroid carcinoma in an endemic goiter region with iodine substitution. *Annals of Surgery* **207** 227–234.
- Passler C, Scheuba C, Prager G, Kaczirek K, Kaserer K, Zettinig G & Niederle B 2004 Prognostic factors of papillary and follicular thyroid cancer: differences in an iodine-replete endemic goiter region. *Endocrine-Related Cancer* **11** 131–139.
- Rosai J, Carcangiu ML & DeLellis RA 1992 Tumours of the thyroid gland. In *Atlas of Tumour Pathology*, 3rd series, fascicle 5, pp 49–63. Eds J Rosai & LH Sobin. Washington, DC: Armed Forces Institute of Pathology.
- Schemper M & Stare J 1996 Explained variation in survival analysis. *Statistics in Medicine* **15** 1999–2012.
- Shaha AR 2004 Implications of prognostic factors and risk groups in the management of differentiated thyroid cancer. *Laryngoscope* **114** 393–402.
- Shaha AR, Loree TR & Shah JP 1994 Intermediate-risk group for differentiated carcinoma of the thyroid. *Surgery* **116** 1036–1041.
- Sherman SI, Brierley JD, Sperling M, Ain KB, Bigos ST, Cooper DS, Haugen BR, Ho M, Klein I, Ladenson PW et al. 1998 Prospective multicenter study of thyroid carcinoma treatment: initial analysis of staging and outcome. *Cancer* **83** 1012–1021.
- Steinmuller T, Klupp J, Rayes N, Ulrich F, Jonas S, Graf KJ & Neuhaus P 2000 Prognostic factors in patients with differentiated thyroid carcinoma. *European Journal of Surgery* **166** 29–33.
- Sugitani I, Kasai N, Fujimoto Y & Yanagisawa A 2004 A novel classification system for patients with PTC: addition of the new variables of large (3 cm or greater) nodal metastases and reclassification during the follow-up period. *Surgery* **135** 139–148.
- Wittekind CH, Greene FL, Henson DE, Hutter RVP & Sobin LH 2003 *TNM Supplement: a Commentary on Uniform Use*. edn 3. New York: Wiley-Liss, Inc.
- Yildirim E 2005 A model for predicting outcomes in patients with differentiated thyroid cancer and model performance in comparison with other classification systems. *Journal of the American College of Surgeons* **200** 378–392.