Performance of ATA Risk Stratification Systems, Response to Therapy, and Outcome in an Indian Cohort of Differentiated Thyroid Carcinoma Patients: A Retrospective Study

Rony Ruben Praveen V. Pavithran V. Usha Menon Vasantha Nair
Harish Kumar

Department of Endocrinology, Amrita Institute of Medical Sciences, Kochi, India

Keywords

ATA risk stratification systems · Differentiated thyroid carcinoma · Indian population · Kerala · Outcome · Validation

Abstract

Introduction: For better individualized management of differentiated thyroid carcinoma (DTC), ATA risk stratification systems (RSS) of 2009 and 2015 as well as a response to therapy re-classification (ATA RTR-2015) are used worldwide for assessing risk of recurrence. But there are no validation studies of these systems from the Indian subcontinent. **Objectives:** To compare ATA RSS-2009, ATA RSS-2015, and ATA RTR-2015 for their accuracy in predicting outcome in DTC patients. **Methods:** This was a retrospective review of 236 adult patients with DTC >1 cm attending the Thyroid Cancer Clinic at our Institute who had undergone total thyroidectomy and radioactive iodine ablation. Initial risk stratification using ATA RSS-2009 and RSS-2015, clinical response at 1 year and outcome at last follow-up measured by clinical end points were collected and analyzed. **Results:** ATA RSS-2015 could not be applied to this cohort due to lack of histopathology details. While 77.3% of low-risk ATA RSS-2009 had disease-free status (NED, no evidence of disease) on follow-

up, 96.1% of patients, in excellent response in ATA RTR-2015, showed NED. Whereas persistent structural disease was predicted by the high-risk group in ATA RSS-2009 (61.9%) and by the incomplete structural response group in ATA RTR-2015 (57.1%) equally well, the best predictor for NED at 1 year in this cohort was ATA RTR-2015 ($p < 0.001$). **Conclusion:** This study found that both ATA RSS-2009 and ATA RTR-2015 are reliable in predicting outcome in DTC patients after initial treatment. However, the response to initial therapy at 1 year predicted outcome more accurately than the initial risk status.

© 2019 European Thyroid Association
Published by S. Karger AG, Basel

Introduction

According to the National Cancer Registry Programme-Indian Council of Medical Research (ICMR) 2012–2014, thyroid cancer is among the top 10 cancers in women, after breast cancer. In Thiruvananthapuram and Kollam, cities in Kerala, a southern state of India, age-adjusted incidence rates are 13.3/100,000 females in Thiruvananthapuram and 12/100,000 females in Kollam. Though there is insufficient data on differentiated thyroid carcinoma (DTC) from Kerala except the regional

data mentioned above, increasing numbers of patients are seen at the Thyroid Carcinoma Clinic at our Institute in Kochi every year. This, however could be a selection bias, as ours is a tertiary care center. It is predicted that by 2019, papillary thyroid carcinoma will become the third most common cancer in women [1]. Hence, optimization of long-term health outcomes and education about potential prognosis for individuals with thyroid neoplasms are critically important.

In DTC, mortality is not a big issue, as the majority of patients have excellent survival [2]. More common is the persistence or recurrence of disease [3, 4] that leads to multiple retesting and interventions, which affect the quality of life. A system of evaluation is therefore necessary to predict the risk of persistence or recurrence of disease after initial therapy seen in about 12–29% [3, 4].

In 1996, the American Thyroid Association (ATA) published treatment guidelines for patients with thyroid nodules and DTC [5]. Because the American Joint Committee for Cancer (AJCC) staging system predicts only mortality from the disease [5–8], the 2009 version of the ATA thyroid cancer guidelines proposed a three-tiered clinicopathologic risk stratification system (RSS) that classified patients as having low, intermediate, or high risk of recurrence [9].

While the 2009 RSS (ATA RSS-2009) has proven to be a valuable tool for initial stratification in DTC, modifications were required to better incorporate the new understanding regarding the risks associated with the extent of lymph node involvement, mutational status, and specific DTC histologies. Hence, a modified RSS was brought about by ATA in 2015 (ATA RSS-2015) [10]. For the purpose of follow-up, the ATA 2015 guidelines also included a Response to Therapy Re-classification (ATA RTR-2015) which could be applied 1–2 years after initial therapy to assess the benefits and plan individualized management [10].

The validation of ATA RSS 2009 and 2015 was based on clinical response and clinical end points in various studies [11–13]. However, to our knowledge, there have been no Indian Guidelines incorporating risk status and response to therapy or any studies in DTC patients on the Indian subcontinent validating the ATA RSS-2009, ATA RSS-2015, or ATA RTR-2015. Therefore, this study was undertaken to try and validate the ATA guidelines in our population of DTC patients before applying these in our day-to-day practice. We also attempted to look at which risk stratification, 2009 versus 2015, is better at predicting outcome in the Indian population.

Objectives

The primary objective of the present study was to compare ATA 2009 (ATA RSS-2009) and 2015 RSS (ATA RSS-2015) in a cohort of DTC patients who underwent total thyroidectomy and radioactive iodine ablation at the Thyroid Cancer Clinic of our Institute and to assess which of these most accurately predicted outcomes using clinical end points.

The secondary objective was to compare the performance of the better of the above two systems of risk stratification with ATA RTR-2015 to assess which best predicted outcome using clinical end points at last follow-up.

Materials and Methods

Study Participants

The study participants consisted of DTC patients aged >18 years, who had undergone total thyroidectomy and radioiodine ablation, attending the Thyroid Cancer Clinic in The Department of Endocrinology at Amrita Institute of Medical Sciences Kochi (AIMS). Patients who completed minimum 3 years of follow-up from 2010 to 2018 were included in the study. Patients with micro-papillary thyroid carcinoma and patients with insufficient data for risk stratification/follow-up were excluded. A total of 236 patients formed the study cohort. Institutional ethical committee approval was obtained for retrospective data collection.

The Thyroid Cancer Clinic was set up at our Institute in January 2010. From the outset, the management was based on the ATA Guidelines published in 2009. The follow-up used the response to therapy categories outlined and validated by Tuttle et al. [7]. The high-risk patients were followed closely at 3- to 6-month intervals and the low- and intermediate-risk groups at 6–12 months. Whenever there was suspicion for residual or recurrent disease (RD), further evaluation was performed with diagnostic whole body iodine scanning, neck ultrasonography, chest X-ray, computed tomography (CT), FDG PET/CT, etc. Neck reexploration was performed for suitable cases, and repeat radioiodine therapy was also given in selected patients. To evaluate outcome at last follow-up, clinical end points were defined as five categories: no evidence of disease (NED), persistent biochemical disease (PBD), persistent structural disease (PSD), recurrent disease (RD), and indeterminate disease (ID) [11–13].

NED was defined as stimulated/suppressed thyroglobulin (Tg) <1 ng/mL or undetectable, Tg antibody (Tg-Ab) negative, and no functional or structural evidence of disease.

PBD was defined as suppressed Tg >1 ng/mL or stimulated Tg >10 ng/mL, and if Tg-Ab was present, rising with no structural evidence on imaging.

PSD was defined as local, regional, or distant disease with any Tg or Tg-Ab levels.

RD was diagnosed if development of any disease activity, biochemical or structural any time after NED state was evidenced.

ID was defined as Tg and Tg-Ab still detectable but not rising significantly (<20%), suppressed Tg level 0.2–1 ng/mL, stimulated Tg >1 but <10 ng/mL, thyroid bed lesions stable and showing no features of malignancy (microcalcification), and <1-cm neck

Fig. 1. AJCC staging categories in DTC patients ($n = 232$).

Fig. 2. Categories of ATA 2009 risk stratification in our study cohort ($n = 236$).

lymph nodes stable, no features suggestive of malignancy or persistence of faint/negligible uptake in the neck on a nuclear scan.

Laboratory Assays

Tg level was determined by the electrochemiluminescence assay (Elecsys 2010; Roche, Switzerland). Tg- Ab was assessed using the ARCHITECT anti-Tg assay (Abbott, USA) from 2010 to mid-2017. From July 2017, Tg-Ab testing used ELECSYS anti-Tg assay (Roche Laboratories). TSH was assayed using the ARCHITECT TSH assay from 2010 to mid-2017. From then on, the testing used the electrochemiluminescence assay Cobas system (Roche). The reference range is 0.005–100 $\mu\text{IU/mL}$ with a functional sensitivity of 0.014 $\mu\text{IU/mL}$. The lowest level exhibiting a 20% CV was 0.31 IU/mL .

Statistical Analysis

Statistical analysis was performed using IBM SPSS Statistics 20 for Windows (SPSS Inc., Chicago, IL, USA). Continuous quantitative data was expressed using mean and standard deviation and categorical, qualitative data as numbers and percentages. The χ^2 test was used for categorical comparisons. A p value of <0.05 was considered statistically significant. The positive predictive value and accuracy of the 2 systems ATA RTR-2015 and ATA RSS-2009 were calculated using appropriate diagnostic tests.

Results

Two hundred and thirty-six patients were included in the study. Their baseline characteristics are shown in Table 1.

The study participants were risk stratified using the 8th edition of the AJCC/UICC staging system and the ATA RSS-2009. As we compiled the data, we realized that information on micrometastasis to lymph nodes, which is needed for ATA RSS 2015, was not consistently reported in the histopathology reports of these patients. Hence,

Table 1. Basic characteristics of study population ($n = 236$)

Females	172 (72.9%)
Age at diagnosis, mean $\pm$ SD, years	42.09 $\pm$ 13.09
Duration of follow-up, months	
Mean $\pm$ SD	65.5 $\pm$ 28.72
Median	58
Type of thyroidectomy	
Total thyroidectomy alone	136 (57.6%)
Hemi- and completion thyroidectomy	13 (5.5%)
Total thyroidectomy with neck dissection	87 (36.9%)
Histology	
Classical papillary carcinoma	129 (54.7%)
Follicular variant of papillary carcinoma	75 (31.8%)
Follicular thyroid carcinoma	12 (5.1%)
Papillary carcinoma variants	13 (5.4%)
Hurthle cell carcinoma	3 (1.3%)
Mixed types	5 (2.1%)

although nodes larger than 3 cm helped to categorize high risk, full characterization of low- and intermediate-risk groups was not possible. Therefore, the ATA RSS-2015 could not be applied to this cohort, and ATA RSS-2009 was used throughout the study. Response to therapy re-stratification was performed at 1 year of follow-up according to ATA 2015 guidelines (ATA RTR-2015).

When AJCC staging (8th edition) was applied to our cohort (Fig. 1), most patients were in Stage I ($n = 207$, 87.7%) followed by Stage II ($n = 11$, 4.7%). Only 6 (2.5%) fell in stage III, 5 in IVa (2.1%), and 3 in IVb (1.3%). Data was not available in 4 patients. According to ATA 2009 risk stratification (Fig. 2), the maximum number of patients were in the intermediate risk category ($n = 118$, 50%) followed by low risk ($n = 97$, 41.1%). When clinical

Fig. 3. Categories of clinical response at 1 year in our study cohort ($n = 236$).

response was assessed (Fig. 3), the majority of patients were in the excellent category ($n = 129$, 54.6%). When outcome was assessed as clinical end points (Table 2), the largest number had NED ($n = 157$, 66.5%), followed by indeterminate ($n = 38$, 16.1%). Recurrence was seen in 2 patients only (0.8%).

Comparison of AJCC Staging, ATA RSS-2009, and ATA RTR-2015 in Predicting Outcome

AJCC Staging versus Clinical End Points

72% of patients in AJCC Stage I had NED on follow-up (Table 3). The rest of the groups contained too small numbers to obtain any meaningful interpretation.

ATA RSS-2009 Categories versus Clinical End Points

In the low risk group of ATA RSS-2009, 77.3% of patients had NED, while only 19% of patients in the high-risk group had NED. 61.9% of patients in the high-risk group had PSD versus only 5.1% in the low-risk group. PBD was most frequent in the high-risk group (9.5%) followed by the intermediate risk category (3.4%). ID was seen most frequently in the intermediate group (18.7%).

Clinical Response at 1 Year versus Clinical End Points

Most (96.1%) of the excellent response patients had NED at last follow-up. 39.4% of patients who had biochemical incomplete response initially, became NED on follow-up, whereas only 20.4% of the patients who had

Table 2. Clinical end points of patients at last follow-up

Clinical end points	<i>n</i>	%
No evidence of disease	157	66.5
Persistent biochemical disease	8	3.4
Persistent structural disease	31	13.1
Recurrent disease	2	0.9
Indeterminate disease	38	16.1

structural incomplete response initially did so. On follow-up, 40.1% of the patients, who were in the indeterminate category initially also had NED. Among the patients with structural incomplete response, 57.1% remained as PSD. Only 9.1% of patients, who had biochemical incomplete response initially, had PSD on follow-up. Almost 60% of the patients in the indeterminate group remained so on the last follow-up. Only 1 patient in the excellent group had recurrence. There were no deaths. Both high-risk group of ATA RSS-2009 and incomplete structural response in ATA RTR-2015 were able to equally predict PSD in 61.9 and 57.1%, respectively, a group that is likely to progress to advanced disease.

Although ATA RSS 2009 as well as clinical response to therapy at 1 year had good positive predictive values for predicting the NED status when compared to rest of the outcome groups (Table 4), clinical response to therapy at 1 year was more accurate (PPV 96.1%, accuracy 83.9%, and specificity 93.7%) in predicting disease-free status than the ATA risk category. Hence, ATA RTR-2015 appears to be better in predicting disease-free status in our patients.

Discussion

In the follow-up of DTC, from the outset, one should be able to predict the possibility of disease persistence/recurrence as these patients can be followed up more closely and recurrent/persistent disease dealt with as and when needed for a good outcome.

Our cohort consists of 236 patients who were followed and managed under supervision of a single senior endocrinologist. Since 2010, we implemented the clinical response system as proposed by Tuttle et al. [7] along with the ATA 2009 RSS in our thyroid cancer clinic. Systematic use of these 2 systems for clinical decision making as well as the meticulous follow-up of these patients could explain the excellent response to therapy noted in the vast

Table 3. Comparison of AJCC staging, ATA RSS-2009, and ATA RTR-2015 for outcome at last follow-up using clinical end points

Staging systems	Clinical end points				
	no evidence of disease	persistent disease biochemical evidence	persistent disease structural identifiable	recurrent disease	indeterminate disease
<i>AJCC stage</i>					
I	149 (72%)	8 (3.9%)	17 (8.1%)	1 (0.5%)	32 (15.5%)
II	6 (54.5%)	0 (0%)	2 (18.2%)	1 (9.1%)	2 (18.2%)
III	0 (0%)	0 (0%)	5 (83.3%)	0 (0%)	1 (16.7%)
IVa	0 (0%)	0 (0%)	3 (60%)	0 (0%)	2 (40%)
IVb	0 (0%)	0 (0%)	3 (100%)	0 (0%)	0 (0%)
<i>ATA RSS-2009 risk categories</i>					
Low	75 (77.3%)	2 (2.1%)	5 (5.1%)	0 (0%)	15 (15.5%)
Intermediate	78 (68.1%)	4 (3.4%)	13 (11%)	1 (0.8%)	22 (18.7%)
High	4 (19%)	2 (9.5%)	13 (61.9%)	1 (4.8%)	1 (4.8%)
<i>ATA RTR-2015 categories</i>					
Excellent	124 (96.1%)	0 (0%)	0 (0%)	1 (0.8%)	4 (3.1%)
Biochemical incomplete response	13 (39.4%)	6 (18.2%)	3 (9.1%)	1 (3%)	10 (30.3%)
Structural incomplete response	10 (20.4%)	2 (4.1%)	28 (57.1%)	0 (0%)	9 (18.4%)
Indeterminate	10 (40%)	0 (0%)	0 (0%)	0 (0%)	15 (60%)

Table 4. Sensitivity and accuracy of the ATA risk stratification system 2009 (ATA RSS-2009) and clinical response to therapy (ATA RTR 2015) in predicting NED

	PPV, %	NPV, %	Sensitivity, %	Specificity, %	Accuracy, %
<i>ATA RSS-2009</i>					
Low risk (<i>n</i> = 75) vs. others (<i>n</i> = 82)	77.3	41.0	47.8	72.1	55.9
<i>ATA RTR-2015</i>					
Excellent response (<i>n</i> = 124) vs. others (<i>n</i> = 33)	96.1	69.2	79.0	93.7	83.9

majority of our patients, which is comparable to the findings in Western and other studies [11–13].

The AJCC staging system accounts for only mortality from thyroid carcinoma, a disease where mortality is rare [2]. The very low percentage of patients in the AJCC Stage III (2.5%) in our study can be attributed to the downstaging in the 8th edition that we used compared with the use of the 7th edition in other studies. There were no deaths in our cohort, and as such, the AJCC staging system was not found to be very useful in predicting recurrence or persistence of disease in this cohort.

The inability to apply the ATA 2015 RSS was an important finding in this study due to the lack of histopathology details. This may be the situation in many oth-

er developing countries with resource-limited settings. Hence, ATA RSS-2009 and ATA RTR-2015 together could be used as a useful alternative for planning follow-up in such clinics.

Although both ATA RSS-2009 and ATA RTR-2015 were able to predict incomplete structural response equally well, disease-free status was better predicted by ATA RTR-2015 in our patient cohort. Though the sample size was not large enough to make any firm conclusions that can be generalized to all thyroid cancer patients, this finding could suggest that, irrespective of the initial risk category, if the patients attain the NED status by first year by surgery and radioiodine therapy, such patients have a very high chance of remaining disease free. This may also

emphasize the fact that initial aggressive management of DTC and making patients disease free may be a more suitable approach in countries like India where long-term follow-up and close monitoring are not feasible or possible due to financial constraints and lack of insurance cover.

Comparison between our and other validation studies was not possible as there were differences, though minor, in patient selection. It is interesting to note that recurrences, although rare, did occur despite excellent response to therapy in our patients, and this finding is similar to other validation studies [11–13]. This would suggest that all such patients require some form of follow-up at perhaps yearly intervals to detect recurrence early.

To the best of our knowledge, the present study is the first study from the Indian subcontinent, to date, validating ATA RSS-2009 and ATA RTR-2015 to predict outcome in DTC patients. From our study, we conclude that both ATA RSS-2009 and ATA RTR-2015 are capable of predicting outcome. However, ATA RTR-2015 could independently predict outcome with greater accuracy and hence can be used as the primary tool in place of ATA 2009 for our patients. This study has convincingly shown that the application of ATA guidelines, largely used in clinical practice all over the world in the management of DTC, were able to predict the disease status on follow-up in our patient population. In the resource-limited setting in which we practise, this would serve to streamline follow-up using intensive investigations only in those expected to have incomplete response to initial therapy, at the same time, avoiding unnecessary expensive and exhaustive tests in the large majority who are likely to have NED on follow-up.

Acknowledgement

The authors would like to thank the Thyroid Cancer Care Team at Amrita Institute of Medical Sciences comprising of the Departments of Head and Neck Surgery, Nuclear Medicine, Medical Oncology, Surgical Oncology, Radiation Oncology, General Surgery, and Endocrinology. We would also like to thank Dr. Sundaram, Head of Biostatistics, and Dr. Gangadharan, Department of Cancer Registry, at Amrita Institute of Medical Sciences.

Statement of Ethics

The study protocol has been approved by the Research Institute's Committee on Human Research.

Disclosure Statement

The authors have no conflicts of interest to declare.

Funding Sources

No funding has been received for the study.

Author Contributions

Dr. Vasantha Nair is in charge of Thyroid Cancer Clinic and has been following up the patients. Dr. V. Usha Menon is also associated with the thyroid cancer clinic and was involved in statistical analysis. Dr. Praveen V. Pavithran was responsible for overseeing the project and statistical analysis. Dr. Rony Ruben was in charge of collection of data, analysis, and writing up. Dr. Harish Kumar was the overall coordinator of the entire study.

References

- 1 Aschebrook-Kilfoy B, Schechter RB, Shih YC, Kaplan EL, Chiu BC, Angelos P, et al. The clinical and economic burden of a sustained increase in thyroid cancer incidence. *Cancer Epidemiol Biomarkers Prev*. 2013 Jul;22(7):1252–9.
- 2 Xu B, Ibrahimasic T, Wang L, Sabra MM, Migliacci JC, Tuttle RM, et al. Clinicopathologic Features of Fatal Non-Anaplastic Follicular Cell-Derived Thyroid Carcinomas. *Thyroid*. 2016 Nov;26(11):1588–97.
- 3 de Castro TP, Waissmann W, Simões TC, de Mello RC, Carvalho DP. Predictors for papillary thyroid cancer persistence and recurrence: a retrospective analysis with a 10-year follow-up cohort study. *Clin Endocrinol (Oxf)*. 2016 Sep;85(3):466–74.
- 4 Leung AM, Dave S, Lee SL, Campion FX, Garber JR, Pearce EN. Factors determining the persistence or recurrence of well-differentiated thyroid cancer treated by thyroidectomy and/or radioiodine in the Boston, Massachusetts area: A retrospective chart review. *Thyroid Res*. 2011 Apr;4(1):9.
- 5 Singer PA, Cooper DS, Daniels GH, Ladenson PW, Greenspan FS, Levy EG, et al.; American Thyroid Association. Treatment guidelines for patients with thyroid nodules and well-differentiated thyroid cancer. *Arch Intern Med*. 1996 Oct;156(19):2165–72.
- 6 Baek SK, Jung KY, Kang SM, Kwon SY, Woo JS, Cho SH, et al. Clinical risk factors associated with cervical lymph node recurrence in papillary thyroid carcinoma. *Thyroid*. 2010 Feb;20(2):147–52.

- 7 Tuttle RM, Tala H, Shah J, Leboeuf R, Ghossein R, Gonen M, et al. Estimating risk of recurrence in differentiated thyroid cancer after total thyroidectomy and radioactive iodine remnant ablation: using response to therapy variables to modify the initial risk estimates predicted by the new American Thyroid Association staging system. *Thyroid*. 2010 Dec; 20(12):1341–9.
- 8 Vaisman F, Momesso D, Bulzico DA, Pessoa CH, Dias F, Corbo R, et al. Spontaneous remission in thyroid cancer patients after biochemical incomplete response to initial therapy. *Clin Endocrinol (Oxf)*. 2012 Jul;77(1): 132–8.
- 9 Cooper DS, Doherty GM, Haugen BR, Kloos RT, Lee SL, Mandel SJ, et al.; American Thyroid Association (ATA) Guidelines Taskforce on Thyroid Nodules and Differentiated Thyroid Cancer. Revised American Thyroid Association management guidelines for patients with thyroid nodules and differentiated thyroid cancer. *Thyroid*. 2009 Nov;19(11):1167–214.
- 10 Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, et al. 2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer: The American Thyroid Association Guidelines Task Force on Thyroid Nodules and Differentiated Thyroid Cancer. *Thyroid*. 2016 Jan;26(1):1–133.
- 11 Castagna MG, Maino F, Cipri C, Belardini V, Theodoropoulou A, Cevenini G, et al. Delayed risk stratification, to include the response to initial treatment (surgery and radioiodine ablation), has better outcome predictivity in differentiated thyroid cancer patients. *Eur J Endocrinol*. 2011 Sep;165(3): 441–6.
- 12 Pitoia F, Bueno F, Urciuoli C, Abelleira E, Cross G, Tuttle RM. Outcomes of patients with differentiated thyroid cancer risk-stratified according to the American thyroid association and Latin American thyroid society risk of recurrence classification systems. *Thyroid*. 2013 Nov;23(11):1401–7.
- 13 Shen FC, Hsieh CJ, Huang IC, Chang YH, Wang PW. Dynamic Risk Estimates of Outcome in Chinese Patients with Well-Differentiated Thyroid Cancer After Total Thyroidectomy and Radioactive Iodine Remnant Ablation. *Thyroid*. 2017 Apr;27(4):531–6.