Real-World Performance of the American Thyroid Association Risk Estimates in Predicting 1-Year Differentiated Thyroid Cancer Outcomes: A Prospective Multicenter Study of 2000 Patients

Thyroid

DOI: 10.1089/thy.2020.0272

Real-world performance of the American Thyroid Association risk estimates in predicting 1-year differentiated thyroid cancer outcomes: A prospective multicenter study of 2000 patients

Giorgio GRANI[1], Maria Chiara ZATELLI[2], Marco ALFÒ[3], Teresa MONTESANO[4], Massimo TORLONTANO[5], Silvia MORELLI[6], Maurilio DEANDREA[7], Alessandro ANTONELLI[8], Cecilia FRANCESE[9], Graziano CERESINI[10], Fabio ORLANDI[11], Carolina Adele MANIGLIA[12], Rocco BRUNO[13], Salvatore MONTI[14], Maria Giulia SANTAGUIDA[15], Andrea REPACI[16], Giovanni TALLINI[16], Laura FUGAZZOLA[17], Fabio MONZANI[18], Raffaele GIUBBINI[19], Ruth ROSSETTO[20], Caterina MIAN[21], Anna CRESCENZI[22], Dario TUMINO[23], Loredana PAGANO[20,24], Luciano PEZZULLO[25], Celestino Pio LOMBARDI[26], Emanuela ARVAT[27], Luisa PETRONE[28], Maria Grazia CASTAGNA[29], Giovanna SPIAZZI[30], Domenico SALVATORE[31], Domenico MERINGOLO[32], Erica SOLAROLI[33], Fabio MONARI[16], Flavia MAGRI[34], Vincenzo TRIGGIANI[35], Roberto CASTELLO[36], Cesare PIAZZA[37], Roberta ROSSI[38], Umberto FERRARO PETRILLO[3], Sebastiano FILETTI[1], Cosimo DURANTE[1]

1. Dept. of Translational and Precision Medicine, Sapienza University of Rome, Rome, Italy – giorgio.grani@uniroma1.it; sebastiano.filetti@uniroma1.it; cosimo.durante@uniroma1.it
2. Section of Endocrinology & Internal Medicine, Dept of Medical Sciences, University of Ferrara, Ferrara, Italy – ztlmch@unife.it
3. Dept. of Statistical Sciences, Sapienza University of Rome, Rome, Italy – marco.alfo@uniroma1.it; umberto.ferraro@uniroma1.it
4. Dept. of Radiological, Pathological and Oncological Sciences, Sapienza University of Rome, Rome, Italy – teresa.montesano@uniroma1.it
5. Dept. of Medical Sciences, Fondazione IRCCS Casa Sollievo della Sofferenza, San Giovanni Rotondo, Italy – m.torlontano@operapadrepio.it
6. Dept. of Medicine, University of Perugia, Perugia, Italy – silviettamorelli@gmail.com
7. Division of Endocrinology, Diabetology, and Metabolism, Mauriziano Umberto I Hospital, Turin, Italy – mdeandrea@mauriziano.it
8. Dept. of Clinical and Experimental Medicine, University of Pisa, Pisa, Italy – alessandro.antonelli@med.unipi.it

9. Division of Endocrinology, Clinica Salus di Battipaglia, Salerno, Italy – studiomedico21@gmail.com
10. Dept. of Medicine and Surgery, University of Parma, Parma, Italy – graziano.ceresini@unipr.it
11. Dept. of Oncology, Division of Endocrinology and Metabolism, Humanitas-Gradenigo Hospital, University of Turin, Turin, Italy – fabio.orlandi@unito.it
12. Division of Endocrinology, Cervello Hospital, Palermo, Italy – adelemaniglia@libero.it
13. Unit of Endocrinology, Tinchi-Pisticci Hospital, Matera, Italy – roccobruno1@virgilio.it
14. Dept. of Endocrinology, AOU Sant'Andrea, Sapienza Università di Roma, Rome, Italy – salvatore.monti@uniroma1.it
15. Endocrine Unit, AUSL Latina, Italy – mariagiulia.santaguida@uniroma1.it
16. Endocrinology, Pathology and Radiotherapy Units, University of Bologna Medical Center, S. Orsola-Malpighi Hospital, Bologna, Italy – rep.rep@libero.it; giovanni.tallini@unibo.it, fabio.monari@aosp.bo.it
17. Dept. of Endocrine and Metabolic Diseases, IRCCS Istituto Auxologico Italiano, Milan and Dept. of Pathophysiology and Transplantation, University of Milan, Milan, Italy – laura.fugazzola@unimi.it
18. Geriatrics Unit, Dept. of Clinical and Experimental Medicine, University of Pisa, Pisa, Italy – fabio.monzani@med.unipi.it
19. Nuclear Medicine Unit, Spedali Civili Università degli Studi di Brescia, Brescia, Italy – giubbini@med.unibs.it
20. Division of Endocrinology, Diabetology and Metabolism, Dept. of Medical Sciences, University of Turin, Turin, Italy – ruth.rossetto@unito.it
21. Endocrinology Unit, Dept. of Medicine-DIMED, University Hospital of Padua, Padua, Italy – caterina.mian@unipd.it
22. Pathology Unit, University Hospital Campus Bio-Medico, Rome, Italy – a.crescenzi@unicampus.it
23. Dept. of Clinical and Experimental Medicine, University of Catania, Catania, Italy – dario.tumino@phd.unict.it

Thyroid
Real-world performance of ATA risk estimates in predicting 1-year DTC outcomes: A prospective multicenter study of 2000 patients (ITCO Study #3) (DOI: 10.1089/thy.2020.0272)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

24. Endocrinology, Dept. of Translational Medicine, Università del Piemonte Orientale, Novara, Italy – loredana.pagano@med.uniupo.it
25. Struttura Complessa Chirurgia Oncologica della Tiroide, Istituto Nazionale Tumori – IRCCS - Fondazione G.Pascale, Naples, Italy – l.pezzullo@istitutotumori.na.it
26. Division of Endocrine Surgery, Fondazione Policlinico Gemelli, Catholic University, Rome, Italy – celestinopio.lombardi@unicatt.it
27. Oncological Endocrinology Unit, Dept. of Medical Sciences, Molinette Hospital, A.O.U. Città della Salute e della Scienza di Torino, University of Turin, Turin, Italy – emanuela.arvat@unito.it
28. Endocrinology Unit, Careggi University Hospital, Florence, Italy – luisa.petrone@unifi.it
29. Dept. of Medical, Surgical and Neurological Sciences, University of Siena, Siena, Italy – mariagrazia.castagna@unisi.it
30. Section of Endocrinology, Diabetes, and Metabolism, Dept. of Medicine, University of Verona, Verona, Italy – giovanna.spiazzi@ospedaleuniverona.it
31. Dept. of Public Health, University of Naples "Federico II", Naples, Italy – domenico.salvatore@unina.it
32. Endocrinology Unit, Istituto Ramazzini, Bologna, Italy – meringolod@libero.it
33. Endocrinology Unit, Medical Department, AUSL Bologna Maggiore - Bellaria Hospital, Bologna, Italy – e.solaroli@ausl.bologna.it
34. Unit of Internal Medicine and Endocrinology, Laboratory for Endocrine Disruptors, Dept. of Internal Medicine, Istituti Clinici Scientifici Maugeri IRCCS, and Dept. of Internal Medicine and Therapeutics, University of Pavia, Pavia, Italy – flavia.magri@icsmaugeri.it
35. Section of Internal Medicine, Geriatrics, Endocrinology and Rare Disease, Interdisciplinary Dept. of Medicine, School of Medicine, University of Bari, Bari, Italy – vincenzo.triggiani@uniba.it
36. Division of General Medicine, University Hospital of Verona, Verona, Italy – roberto.castello@aovr.veneto.it

Thyroid
Real-world performance of ATA risk estimates in predicting 1-year DTC outcomes: A prospective multicenter study of 2000 patients (ITCO Study #3) (DOI: 10.1089/thy.2020.0272)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

37. Dept, of Otorhinolaryngology, Maxillofacial and Thyroid Surgery, Fondazione IRCCS, National Cancer Institute of Milan, University of Milan, Milan, Italy cesare.piazza@istitutotumori.mi.it
38. Endocrine Unit, Azienda Ospedaliero Universitaria S. Anna, Ferrara, Italy - roberta.rossi@unife.it

Running title: ATA risk estimate in real-world clinical practice

Keywords: differentiated thyroid cancer; evidence based guidelines; clinical practice; risk stratification

Thyroid
Real-world performance of ATA risk estimates in predicting 1-year DTC outcomes: A prospective multicenter study of 2000 patients (ITCO Study #3) (DOI: 10.1089/thy.2020.0272)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

Thyroid
Real-world performance of ATA risk estimates in predicting 1-year DTC outcomes: A prospective multicenter study of 2000 patients (ITCO Study #3) (DOI: 10.1089/thy.2020.0272)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

Abstract

Background. One of the most widely used risk stratification systems for estimating individual patients' risk of persistent or recurrent differentiated thyroid cancer (DTC) is the American Thyroid Association (ATA) guidelines. The 2015 ATA version, that has increased the number of patients considered at low or intermediate risk, has been validated in several retrospective, single center studies. The aims of this study were to evaluate the real-world performance of the 2015 ATA risk stratification system in predicting the response to treatment 12 months after the initial treatment and to determine the extent to which this performance is affected by the treatment center in which it is used.

Methods. A prospective cohort of DTC patients collected by the Italian Thyroid Cancer Observatory web-based database was analyzed. We reviewed all records present in the database and selected consecutive cases that satisfied inclusion criteria: 1) histological diagnosis of DTC, with the exclusion of noninvasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP); 2) complete data of the initial treatment and pathological features; 3) results of 1-year follow-up visit (6–18 months after the initial treatment), including all data needed to classify the estimated response to treatment.

Results. The final cohort was composed of 2071 patients from 40 centers. The ATA risk of persistent/recurrent disease was classified as *low* in 1109 patients (53.6%), *intermediate* in 796 (38.4%), and *high* in 166 (8.0%). Structural incomplete responses were documented in only 86 (4.2%) patients: 1.5% in the low-risk, 5.7% in the intermediate-risk, and 14.5% in the high-risk group. The baseline ATA risk class proved to be a significant predictor of structural persistent disease, both for intermediate- (OR 4.67; 95% CI 2.59-8.43) and high-risk groups (OR 16.48; 95% CI 7.87-34.5). Individual center did not significantly influence the prediction of the 1-year disease status.

Conclusions. The ATA risk stratification system is a reliable predictor of short-term outcomes in patients with DTC in real-world clinical settings characterized by center heterogeneity in terms of size, location, level of care, local management strategies, and resource availability.

Introduction

Most cases of differentiated thyroid cancer (DTC) currently diagnosed display indolent clinical behavior and are associated with very low mortality rates. Therefore, a more conservative approach in DTC management, with less extensive surgery, more selective use of radioiodine, and less intensive follow-up protocols, is being recommended (1). The aim is to avoid subjecting low-risk patients to unnecessary diagnostic procedures and overtreatment without reducing the chances of identifying those rare cancers that are likely to require more aggressive management (2). To facilitate this process, several scientific societies have developed tools for the prognostic stratification of patients with DTC. One of the most widely used is that developed by the American Thyroid Association (ATA), which aids clinicians in estimating individual patients' risk of persistent or recurrent disease (3). Its usefulness in these settings has been demonstrated by several retrospective, single-center studies (4).

A revised version of this system was included in the ATA guidelines published in 2016 (1). The criteria for classifying the risk of recurrence as high in this version are substantially more restrictive as compared to the previous ones and, as a result, the number of patients who will be considered at low- or intermediate-risk has increased. The updated stratification system has also been validated in several retrospective cohort studies in different parts of the world. Most of these studies, however, were conducted at single healthcare facilities that served as a referral center for patients with thyroid cancer (5-7), and it is unclear whether their findings reflect the performance of the system in real-world, heterogeneous clinical settings. New data, ideally from prospective, multicenter studies, are needed to better define the impact of the recent revisions on the system's ability to predict post-treatment DTC evolution, fundamental information for developing cost-effective follow-up strategies.

In this prospective cohort study, we analyzed data of over 2000 cases of DTCs managed in 40 diverse healthcare settings in Italy. Our aims were: 1) to evaluate the performance of the 2015 ATA risk stratification system in predicting the response to treatment documented approximately 12 months after the initial disease treatment; and 2) to determine the extent to which this performance is affected by the treatment center itself.

Thyroid
Real-world performance of ATA risk estimates in predicting 1-year DTC outcomes: A prospective multicenter study of 2000 patients (ITCO Study #3) (DOI: 10.1089/thy.2020.0272)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

Methods

The Italian Thyroid Cancer Observatory (ITCO) web-based database was opened in 2013 at the Thyroid Cancer Center of the Sapienza University of Rome (the network's Coordinating Center). Since then, it expanded to 49 other thyroid cancer centers in the country that joined the network (8). The database now includes prospectively collected data on nearly 7000 patients with histologically confirmed diagnoses of differentiated, medullary, poorly differentiated, or anaplastic thyroid cancer. Cases are inputted in the database at the time of the initial treatment in the reporting ITCO center, or when the patient begins follow-up in the reporting center within 12 months after undergoing initial treatment in a non-ITCO center. Each case record contains information on patient demographics and biometrics, circumstances of the diagnosis, tumor pathology, surgical and radioactive iodine treatments, as well as the results of periodic follow-up examinations. The ITCO provides no guidance or restrictions in terms of patient management to the participating centers, since the database is designed to provide a picture of real-world practices. Sensitive patient data are encrypted, and the database is anonymously managed for statistical analysis.

For the purposes of the present study, we reviewed all records present in the database and selected consecutive cases that satisfied the following criteria: 1) histological diagnosis of DTC—papillary (PTC), follicular (FTC), and poorly differentiated thyroid cancers and their variants (with the exclusion of non-invasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP); 2) availability of all information on the initial treatment and pathological characteristics of the tumor required for the ATA recurrence risk assessment; 3) results of the 1-year follow-up visit (carried out 6–18 months after the initial treatment), including all data needed to classify the estimated response to treatment.

For each case, we recorded the following information:

Initial treatment: Treatment of the primary tumor was classified as *thyroid lobectomy* or *total thyroidectomy*. The latter category also included patients who had had a completion thyroidectomy following thyroid lobectomy. For all patients who had total thyroidectomy, we also recorded the use of radioiodine remnant ablation (RRA) as

performed or not performed. Cervical lymph node dissection, when performed, was described as central compartment dissection, lateral compartment dissection, or both.

Risk of persistent or recurrent disease. The estimated level of risk was determined by the study team in accordance with the 2009 ATA Guidelines (3) and relevant modifications in the 2015 update (1). Classification was based on the data available immediately after the initial treatment **(Figure 1)**. When treatment consisted of lobectomy followed by completion thyroidectomy, we considered pathologic data on tissues collected during both surgical procedures. PTCs regarded as "aggressive" were those with 1) solid, insular, tall-cell, columnar-cell, hobnail-cell, sclerosing, or poorly differentiated histologic subtypes, or 2) evidence of intratumoral blood vessel invasion (regardless of histologic subtype).

Responses to the initial treatment. These were classified as *excellent*, *biochemical incomplete*, *structural incomplete*, or *indeterminate* on the basis of data collected during the clinical evaluation carried out at the 1-year follow-up visit. These data included imaging findings (cervical ultrasound in all patients, and radioactive iodine [RAI] scintigraphy in selected individuals), basal or stimulated serum thyroglobulin (Tg) levels, and anti-Tg antibody (TgAb) levels. Additional imaging studies were performed at the clinicians' discretion. The results were classified as specified in the ATA Guidelines (1) for patients who had undergone thyroidectomy followed by RRA, and as advocated by the European Society for Medical Oncology (9) for those whose initial treatment consisted of surgery alone (thyroidectomy or lobectomy) (**Supplemental Table 1**). Cervical lymph nodes with highly suspicious features on ultrasonography, as defined by the European Thyroid Association guidelines (10), were considered imaging evidence of persistent disease; those displaying low-suspicion features were classified as non-specific imaging findings (11). Suspicious findings of other imaging studies were classified by the treating physicians. The presence of structural disease at the 1-year evaluation was considered as persistent disease (1, 12).

Statistical analyses

In our descriptive analysis, continuous variables were expressed as medians with interquartile ranges (IQR) and nominal variables in terms of frequency counts and corresponding percentages. To model the response to treatment, we chose a cumulative

Thyroid
Real-world performance of ATA risk estimates in predicting 1-year DTC outcomes: A prospective multicenter study of 2000 patients (ITCO Study #3) (DOI: 10.1089/thy.2020.0272)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

Thyroid
Real-world performance of ATA risk estimates in predicting 1-year DTC outcomes: A prospective multicenter study of 2000 patients (ITCO Study #3) (DOI: 10.1089/thy.2020.0272)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

link model, which is routinely used to describe the distribution of ordinal categorical response data. The ordinal response was categorized in descending order of desirability as excellent, indeterminate, biochemical incomplete, or structural incomplete. In this framework, the link used was the log of a probability ratio, calculated as the probability of not exceeding a given category versus the probability of exceeding it (in a logit model perspective), and predictors were inserted on a linear scale. Given the hierarchical structure of the data, with patients nested within treatment centers, we used a mixed-effect model specification, with a center-specific intercept summarizing unobserved center-specific characteristics. To account for potential overlap or association between the center-specific features that were observed, and the unobserved features described by the random intercepts, we inserted in the model the average for each covariate on patients from a given center. We approximated the integral defining the likelihood via the Laplace approximation, using the R library ordinal (13). We also evaluated a binary response to treatment consisting of excellent vs. structural incomplete responses only, in a mixed logit model. This was estimated by the R library lme4 (14). All statistical analyses were performed with the R statistical software package, R Core Team (2017) (15).

Results

Out of 6867 case records in the database at data lock (2019), 1452 (21.1%) were excluded because the histological diagnosis was incomplete or failed to meet the inclusion criteria (i.e., tumors diagnosed as medullary thyroid cancer, anaplastic thyroid cancer, NIFTP, or tumors with unknown malignant potential), 148 additional cases were excluded because one or more items essential to estimate the risk of recurrence were missing. Subsequent exclusions consisted of 3158 cases in which the patient had not yet undergone a 1-year follow-up assessment, and 38 others in which 1-year follow-up data needed to classify the treatment response were lacking. Thus, the final cohort consisted of 2071 patients who have been followed in 40 of the ITCO centers (**Table 1)**.

The ATA risk of persistent/recurrent disease was classified as *low* in 1109 patients (53.6%), *intermediate* in 796 (38.4%), and *high* in 166 (8.0%). Treatment responses observed at the 1-year follow-up visit are summarized in **Table 2.** Overall, structural incomplete responses were documented in 86 (4.2%) patients. The frequency of structural

incomplete responses increased progressively with the level of risk estimated at baseline from 1.5% in the low-risk group, to 5.7% in those at intermediate-risk, and 14.5% in the small subset of patients considered at high-risk for persistent/recurrent disease (**Figure 2**).

As shown in **Table 3**, the ATA risk class assigned at baseline proved to be a significant predictor of the response to treatment observed at the 1-year follow-up visit. It was able to predict the presence of structural disease as opposed to an excellent response. Furthermore, in cases classified as intermediate- or high-risk, there was a significantly higher probability of a "less-than-excellent response" (i.e., indeterminate or biochemical incomplete or structural incomplete, in decreasing order) (**Table 4**).

We also assessed whether the performance of the initial persistent disease risk estimate was significantly influenced by the practices of individual reporting centers, which included both academic and non-academic healthcare facilities distributed throughout Italy (**Table 1**). Some potential biases (such as the case-mix of patients treated, surgical volumes, and different tools used) are difficult to document but can potentially influence both the initial risk estimation and the subsequent assessment of the response to treatment. The mixed effect model we used took these into account with a center-specific intercept summarizing un-observed center-specific features. The practices of individual recruiting centers did not influence the prediction of the 1-year status by the ATA risk stratification system (coefficient -0.88±1.53, p=0.57, in intermediate-risk patients, and -0.77±2.19, p=0.72, in high-risk patients, for the prediction of structural disease).

Discussion

A reliable estimate of the post-treatment risk of persistent or recurrent disease in a patient with DTC on the basis of clinical, histopathological, and perioperative data provides valuable prognostic information. Importantly, it supports clinicians' efforts to develop personalized treatment and follow-up strategies (4, 16). Most cases can be safely managed with less extensive surgery, more selective use of RAI therapy, and relatively relaxed follow-up schedules. The expected benefits are substantial and include reduced healthcare costs, lower treatment-related morbidity rates and improved quality of life for patients. These expected benefits must, however, be weighed against the risk of missing those

thyroid cancers that warrant intensive therapeutic efforts and close post-treatment surveillance due to their intrinsic biologic aggressiveness.

The ATA risk-stratification system was validated in different cohorts around the world (17-20). Additional features were published in 2016 to include evaluation of the number of vascular invasion foci, number and size of involved lymph nodes, presence of extranodal extension and (if available) *BRAF* and *TERT* promoter mutational status. This updated version has been already validated (5-7, 21-23). However, all of these validation studies were based on retrospective review and conducted in a few high-volume thyroid cancer referral centers. The current study is based on a large, contemporary cohort of patients with prospectively collected data in many thyroid centers across Italy, including academic and non-academic institutions, to validate the ATA risk stratification system in predicting persistent disease at the 1-year follow-up visit.

Inter-institutional and inter-observer variabilities have been reported in the diagnosis of histological subtypes (24), the detection and quantification of extrathyroidal extension (25), neck ultrasonographic examination (26), and various aspects of RAI administration, including indications, the amount of activity administered, and the method used to ensure appropriately elevated TSH levels at the time of RAI therapy (withdrawal of thyroid hormone replacement therapy vs. rhTSH injections) (27). Treatment centers also vary widely in the number of thyroidectomies performed by their staff each year and this factor is a well-established predictor of outcome (28, 29). Furthermore, different assays for Tg and TgAb are used in different institutions. All of these confounding factors could not be systematically documented on the basis of the information available in our database, but it was reasonable to suspect that some of them could potentially influence both the initial risk estimates in our cohort as well as the subsequent assessment of the response to treatment. By the use of mixed effect model, we took into account unobserved center-specific features, documenting that the performance of ATA risk stratification system is not affected by the center in which it is applied.

Our findings demonstrate that the ATA risk stratification system for recurrent/persistent disease is indeed a reliable predictor at the 1-year follow-up

Thyroid
Real-world performance of ATA risk estimates in predicting 1-year DTC outcomes: A prospective multicenter study of 2000 patients (ITCO Study #3) (DOI: 10.1089/thy.2020.0272)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

evaluation, independent of treatment centers. This is true in spite of the fact that the likelihood of a "less-than-excellent" response varies across treatment centers, probably as a result of between-center differences in surgical volumes, case mixes, the availability of diagnostic tools, and/or other factors.

It should be stressed that our findings apply exclusively to the prediction of the response to initial treatment documented at the 1-year visit. Risk stratification is in fact a dynamic, ongoing process in which the likelihood of recurrence is periodically re-assessed and the management strategy modified as needed (30). Our current findings cannot provide any indication on how the system will perform in predicting the longer-term evolution of a DTC. It is important to note, however, that most DTC recurrences are identified within the first five years of follow up after initial treatment (31). Moreover, recent evidence suggests that persistent disease observed at the 1 year follow-up visit is associated with worse outcomes than "recurrences" identified later (12). Prediction of this early outcome may thus have particular clinical relevance.

A limitation of our study is related to the inclusion of non-PTC cases in our cohort, in particular FTCs and Hürthle-cell thyroid carcinomas. There is indeed growing evidence that these tumors behave differently from each other and also from PTCs (24). These tumors represented only 5.2% of the DTCs in our cohort. Therefore, our current findings can shed no light on the specific performance of the ATA risk stratification system in patients with these less common thyroid cancer histologic subtypes.

In conclusion, the ATA risk stratification system is a reliable predictor of short-term outcomes in patients with DTC in real-world clinical settings characterized by appreciable treatment-center heterogeneity in terms of size, location, level of care, diagnostic resources, and local management strategies.

Acknowledgments. We thank all the collaborators of the ITCO Network: Rosa Falcone, Valeria Ramundo, Marco Biffoni, Laura Giacomelli (Dept. of Translational and Precision Medicine, and Dept. of Surgical Sciences, Sapienza University of Rome); Michela Massa (Dept. of Medical Sciences, Fondazione IRCCS Casa Sollievo della Sofferenza, S. Giovanni Rotondo); Efisio Puxeddu (Dept. of Medicine, University of Perugia); Paolo Limone

Thyroid

Real-world performance of ATA risk estimates in predicting 1-year DTC outcomes: A prospective multicenter study of 2000 patients (ITCO Study #3) (DOI: 10.1089/thy.2020.0272)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

(Division of Endocrinology, Diabetology, and Metabolism, Mauriziano Umberto I Hospital, Turin); Armando Patrizio (Department of Clinical and Experimental Medicine, University of Pisa); Poupak Fallahi (Department of Translational Research of New Technologies in Medicine and Surgery, University of Pisa); Michela Marina (Dept. of Medicine and Surgery, Azienda Ospedaliero-Universitaria and University of Parma); Ilaria Messuti (Dept of Oncology, Division of Endocrinology and Metabolism, Humanitas-Gradenigo Hospital, University of Turin); Giovanni Savoia, Piernicola Garofalo (Division of Endocrinology Cervello Hospital, Palermo); Maria Grazia Deiana, Federica Presciuttini (Dept. of Endocrinology, AOU Sant'Andrea, Sapienza University of Rome); Marco Centanni, Camilla Virili (Dept of Medico-surgical Sciences and Biotechnologies, Sapienza University of Rome); Valeria Calzolaro (Geriatrics Unit, Dept. of Clinical and Experimental Medicine, University of Pisa); Maria Beatrice Panarotto (ASST-Spedali Civili, Brescia); Ezio Ghigo (Division of Endocrinology, Diabetology and Metabolism, Dept. of Medical Sciences, University of Turin); Andrea Palermo (University Hospital Campus Bio-Medico, Rome); Salvatore Tumino (Dept. of Clinical and Experimental Medicine, University of Catania); Gianluca Aimaretti (Dept. of Translational Medicine, Università del Piemonte Orientale, Novara); Maria Grazia Chiofalo, Vincenzo Marotta (Istituto Nazionale dei Tumori, IRCCS Fondazione Pascale, Naples); Annamaria D'Amore (Division of Endocrine Surgery, Fondazione Policlinico Gemelli, Catholic University, Rome); Alice Nervo, Marco Gallo, Alessandro Piovesan, Alberto Ragni, Francesco Felicetti (Dept. of Medical Sciences, Molinette Hospital, University of Turin); Luca Chiovato, Martina Molteni, Giulia Bendotti (Unit of Internal Medicine and Endocrinology, Laboratory for Endocrine Disruptors, Dept. of Internal Medicine and Therapeutics, Istituti Clinici Scientifici Maugeri IRCCS, University of Pavia); Lorenzo Bresciani, Laura Locati (Istituto Nazionale dei Tumori di Milano).

Writing support was provided by Marian Everett Kent, BSN, and funded by the Fondazione Umberto Di Mario. The study was supported by the Sapienza University of Rome Research Grant (RM11916B83A211FC) to C.D.

Disclosure Statement. Authors have nothing to disclose.

Corresponding author:

Sebastiano Filetti, MD

Dipartimento di Medicina Traslazionale e di Precisione, Sapienza University of Rome

Viale del Policlinico 155, I-00161, Rome, Italy

Phone: +39-06-499775133, e-mail: sebastiano.filetti@uniroma1.it

Thyroid

Real-world performance of ATA risk estimates in predicting 1-year DTC outcomes: A prospective multicenter study of 2000 patients (ITCO Study #3) (DOI: 10.1089/thy.2020.0272)

This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

References

1. Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, Pacini F, Randolph GW, Sawka AM, Schlumberger M, Schuff KG, Sherman SI, Sosa JA, Steward DL, Tuttle RM, Wartofsky L 2016 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 **26**:1-133.

2. Lamartina L, Grani G, Durante C, Filetti S 2018 Recent advances in managing differentiated thyroid cancer. F1000Res **7**:86.

3. Cooper DS, Doherty GM, Haugen BR, Hauger BR, Kloos RT, Lee SL, Mandel SJ, Mazzaferri EL, McIver B, Pacini F, Schlumberger M, Sherman SI, Steward DL, Tuttle RM 2009 Revised American Thyroid Association management guidelines for patients with thyroid nodules and differentiated thyroid cancer. Thyroid **19**:1167-1214.

4. Pitoia F, Jerkovich F 2019 Dynamic risk assessment in patients with differentiated thyroid cancer. Endocr Relat Cancer **26**:R553-R566.

5. Park S, Kim WG, Song E, Oh HS, Kim M, Kwon H, Jeon MJ, Kim TY, Shong YK, Kim WB 2017 Dynamic Risk Stratification for Predicting Recurrence in Patients with Differentiated Thyroid Cancer Treated Without Radioactive Iodine Remnant Ablation Therapy. Thyroid **27**:524-530.

6. Pitoia F, Jerkovich F, Urciuoli C, Schmidt A, Abelleira E, Bueno F, Cross G, Tuttle RM 2015 Implementing the Modified 2009 American Thyroid Association Risk Stratification System in Thyroid Cancer Patients with Low and Intermediate Risk of Recurrence. Thyroid **25**:1235-1242.

7. Momesso DP, Vaisman F, Yang SP, Bulzico DA, Corbo R, Vaisman M, Tuttle RM 2016 Dynamic Risk Stratification in Patients with Differentiated Thyroid Cancer Treated Without Radioactive Iodine. J Clin Endocrinol Metab **101**:2692-2700.

8. Lamartina L, Durante C, Lucisano G, Grani G, Bellantone R, Lombardi CP, Pontecorvi A, Arvat E, Felicetti F, Zatelli MC, Rossi R, Puxeddu E, Morelli S, Torlontano M, Crocetti U, Montesano T, Giubbini R, Orlandi F, Aimaretti G, Monzani F, Attard M, Francese C, Antonelli A, Limone P, Rossetto R, Fugazzola L, Meringolo D, Bruno R, Tumino S, Ceresini G, Centanni M, Monti S, Salvatore D, Spiazzi G, Mian C, Persani L, Barbaro D, Nicolucci A, Filetti S 2017 Are Evidence-Based Guidelines Reflected in Clinical Practice? An Analysis of Prospectively Collected Data of the Italian Thyroid Cancer Observatory. Thyroid **27**:1490-1497.

9. Filetti S, Durante C, Hartl D, Leboulleux S, Locati LD, Newbold K, Papotti MG, Berruti A 2019 Thyroid cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. **30**:1856-1883

10. Leenhardt L, Erdogan MF, Hegedus L, Mandel SJ, Paschke R, Rago T, Russ G 2013 2013 European thyroid association guidelines for cervical ultrasound scan and ultrasound-guided techniques in the postoperative management of patients with thyroid cancer. Eur Thyroid J **2**:147-159.

11. Lamartina L, Grani G, Biffoni M, Giacomelli L, Costante G, Lupo S, Maranghi M, Plasmati K, Sponziello M, Trulli F, Verrienti A, Filetti S, Durante C 2016 Risk Stratification of Neck Lesions Detected Sonographically During the Follow-Up of Differentiated Thyroid Cancer. Journal of Clinical Endocrinology & Metabolism **101**:3036-3044.

12. Sapuppo G, Tavarelli M, Belfiore A, Vigneri R, Pellegriti G 2019 Time to Separate Persistent From Recurrent Differentiated Thyroid Cancer: Different Conditions With Different Outcomes. J Clin Endocrinol Metab **104**:258-265.

13. Christensen RHB 2019 ordinal—Regression Models for Ordinal Data. R package version 2019.12-10. https://CRAN.R-project.org/package=ordinal.

14. Bates D, Mächler M, Bolker B, Walker S 2015 Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, Articles **67**:1-48.

15. Team RC 2014 R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

16. Lamartina L, Grani G, Durante C, Borget I, Filetti S, Schlumberger M 2018 Follow-up of differentiated thyroid cancer - what should (and what should not) be done. Nat Rev Endocrinol **14**:538-551.

17. Castagna MG, Maino F, Cipri C, Belardini V, Theodoropoulou A, Cevenini G, Pacini F 2011 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 **165**:441-446.

18. Tuttle RM, Tala H, Shah J, Leboeuf R, Ghossein R, Gonen M, Brokhin M, Omry G, Fagin JA, Shaha A 2010 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 **20**:1341-1349.

19. Vaisman F, Momesso D, Bulzico DA, Pessoa CH, Dias F, Corbo R, Vaisman M, Tuttle RM 2012 Spontaneous remission in thyroid cancer patients after biochemical incomplete response to initial therapy. Clin Endocrinol (Oxf) **77**:132-138.

20. Pitoia F, Bueno F, Urciuoli C, Abelleira E, Cross G, Tuttle RM 2013 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 **23**:1401-1407.

21. Lee SG, Lee WK, Lee HS, Moon J, Lee CR, Kang SW, Jeong JJ, Nam KH, Chung WY, Jo YS, Lee J 2017 Practical Performance of the 2015 American Thyroid Association Guidelines for Predicting Tumor Recurrence in Patients with Papillary Thyroid Cancer in South Korea. Thyroid **27**:174-181.

22. van Velsen EFS, Stegenga MT, van Kemenade FJ, Kam BLR, van Ginhoven TM, Visser WE, Peeters RP 2019 Evaluating the 2015 American Thyroid Association Risk Stratification System in High-Risk Papillary and Follicular Thyroid Cancer Patients. Thyroid **29**:1073-1079.

23. Albano D, Bertagna F, Bonacina M, Durmo R, Cerudelli E, Gazzilli M, Panarotto MB, Formenti AM, Mazziotti G, Giustina A, Giubbini R 2018 Possible delayed diagnosis and treatment of metastatic differentiated thyroid cancer by adopting the 2015 ATA guidelines. Eur J Endocrinol **179**:143-151.

24. Grani G, Lamartina L, Durante C, Filetti S, Cooper DS 2018 Follicular thyroid cancer and Hurthle cell carcinoma: challenges in diagnosis, treatment, and clinical management. Lancet Diabetes & Endocrinology **6**:500-514.

25. Su HK, Wenig BM, Haser GC, Rowe ME, Asa SL, Baloch Z, Du E, Faquin WC, Fellegara G, Giordano T, Ghossein R, LiVolsi VA, Lloyd R, Mete O, Ozbek U, Rosai J, Suster S, Thompson LD, Turk AT, Urken ML 2016 Inter-Observer Variation in the Pathologic Identification of Minimal Extrathyroidal Extension in Papillary Thyroid Carcinoma. Thyroid **26**:512-517.

26. Grani G, Lamartina L, Cantisani V, Maranghi M, Lucia P, Durante C 2018 Interobserver agreement of various thyroid imaging reporting and data systems. Endocrine Connections **7**:1-7.

27. Rachinsky I, Rajaraman M, Leslie WD, Zahedi A, Jefford C, McGibbon A, Young JE, Pathak KA, Badreddine M, De Brabandere S, Fong H, Van Uum S 2016 Regional Variation across Canadian Centers in Radioiodine Administration for Thyroid Remnant Ablation in Well-Differentiated Thyroid Cancer Diagnosed in 2000-2010. J Thyroid Res **2016**:2867916.

28. Liang TJ, Liu SI, Mok KT, Shi HY 2016 Associations of Volume and Thyroidectomy Outcomes: A Nationwide Study with Systematic Review and Meta-Analysis. Otolaryngol Head Neck Surg **155**:65-75.

29. Al-Qurayshi Z, Robins R, Hauch A, Randolph GW, Kandil E 2016 Association of Surgeon Volume With Outcomes and Cost Savings Following Thyroidectomy: A National Forecast. JAMA Otolaryngol Head Neck Surg **142**:32-39.

30. Tuttle RM, Alzahrani AS 2019 Risk Stratification in Differentiated Thyroid Cancer: From Detection to Final Follow-up. J Clin Endocrinol Metab. **104**:4087-4100.

31. Durante C, Montesano T, Torlontano M, Attard M, Monzani F, Tumino S, Costante G, Meringolo D, Bruno R, Trulli F, Massa M, Maniglia A, D'Apollo R, Giacomelli L, Ronga G, Filetti S, Group PS 2013 Papillary thyroid cancer: time course of recurrences during postsurgery surveillance. J Clin Endocrinol Metab **98**:636-642.

Thyroid

Real-world performance of ATA risk estimates in predicting 1-year DTC outcomes: A prospective multicenter study of 2000 patients (ITCO Study #3) (DOI: 10.1089/thy.2020.0272)

Table 1. Clinical and demographic features of the study cohort

No. patients	**2071**
Age in years - median (IQR)	48 (38-59)
Sex - N (%)	
Female	1546 (74.6%)
Male	525 (25.4%)
ITCO Reporting center - N (%)	40
Academic center	29 (72.5%)
Location	
Northern Italy	19 (47.5%)
Central Italy	13 (32.5%)
Southern Italy	8 (20%)
Time of thyroid cancer diagnosis - N (%)	79 (3.8%)
Not specified	813 (39.3%)
Post-surgical	1179 (56.9%)
Pre-surgical	
Initial treatment of primary tumor - N (%) [a]	
Total thyroidectomy + RRA	1191 (57.5%)
Total thyroidectomy	807 (39%)
Lobectomy	73 (3.5%)
Neck dissection - N (%)	
Not done	1283 (61.9%)
Central neck dissection	546 (26.4%)
Central and lateral neck dissection	196 (9.5%)
Lateral neck dissection	46 (2.2%)
Histology - N (%)	
Papillary thyroid cancer	1958 (94.5%)
Aggressive variants	99 (4.8%)

Classical and other variants	1862 (89.9%)
Follicular thyroid cancer	70 (3.4%)
Hürthle cell carcinoma	37 (1.8%)
Tumor size (mm) - median (IQR)	11 (6-18)
Tumor foci - N (%)	
Not specified	17 (0.8%)
Unifocal	1291 (62.3%)
Multifocal, laterality not specified	19 (0.9%)
Multifocal, unilateral	229 (11.1%)
Multifocal, bilateral	515 (24.9%)
Extrathyroidal extension - N (%)	
None	1439 (69.5%)
Microscopic	563 (27.2%)
Strap muscle invasion	
Yes	31 (5.5%)
No	367 (65.2%)
Not specified[b]	165 (29.3%)
Macroscopic (T4a)	63 (3.0%)
Macroscopic (T4b)	6 (0.3%)
Lymph node status - N (%)	
cN0/Nx	619 (29.9%)
N0	972 (46.9%)
N1a	277 (13.4%)
N1b	203 (9.8%)

[a] The categories “total thyroidectomy + RRA” and “total thyroidectomy” each include some cases of “completion thyroidectomy” performed after thyroid lobectomy.

[b] Many pathology reports did not include details of muscle invasion before the publication of 8th edition of AJCC TNM staging system.

RRA: radioiodine remnant ablation.

Table 2. Responses to treatment at 1-year evaluation

	All	TT+RRA	TT	TL
No. patients	**2071**	**1191**	**807**	**73**
Excellent response	1576 (76.1%)	921 (77.3%)	655 (81.1%)	-
Indeterminate response	376 (18.2%)	168 (14.1%)	137 (17%)	71 (97.3%)*
Biochemical incomplete response	33 (1.6%)	33 (2.8%)	-	-
Structural incomplete response	86 (4.2%)	69 (5.8%)	15 (1.9%)	2 (2.7%)

* not possible to document stability of thyroglobulin values at first evaluation.

Abbreviations: RRA, radioactive iodine remnant ablation; TT, total thyroidectomy; TL, thyroid lobectomy.

Thyroid
Real-world performance of ATA risk estimates in predicting 1-year DTC outcomes: A prospective multicenter study of 2000 patients (ITCO Study #3) (DOI: 10.1089/thy.2020.0272)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

Table 3. Subgroup analysis of 1662 patients with excellent or structural complete responses at 1 year. Likelihood of structural disease according to the estimated risk of persistent disease calculated at baseline

	Likelihood of structural incomplete response at 1 year [a]		
Estimated risk of persistent disease	**Coefficient ± SE**	**OR (95% CI)**	**p value**
Low (reference)	-	-	
Intermediate	1.54±0.3	4.67 (2.59-8.43)	<0.0001
High	2.8±0.38	16.48 (7.87-34.5)	<0.0001

[a] Versus excellent response.

Thyroid
Real-world performance of ATA risk estimates in predicting 1-year DTC outcomes: A prospective multicenter study of 2000 patients (ITCO Study #3) (DOI: 10.1089/thy.2020.0272)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

Table 4. Likelihood of less-than-excellent response (all 4 classes; ordinal analysis) according to the estimated risk of persistent disease calculated at baseline

ATA risk	Coefficient ± SE	OR (95% CI)	P
Low (reference)	-		
Intermediate	0.52±0.12	1.68 (1.34-2.10)	<0.0001
High	1.17±0.19	3.23 (2.23-4.67)	<0.0001

Thyroid
Real-world performance of ATA risk estimates in predicting 1-year DTC outcomes: A prospective multicenter study of 2000 patients (ITCO Study #3) (DOI: 10.1089/thy.2020.0272)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

Figure legends

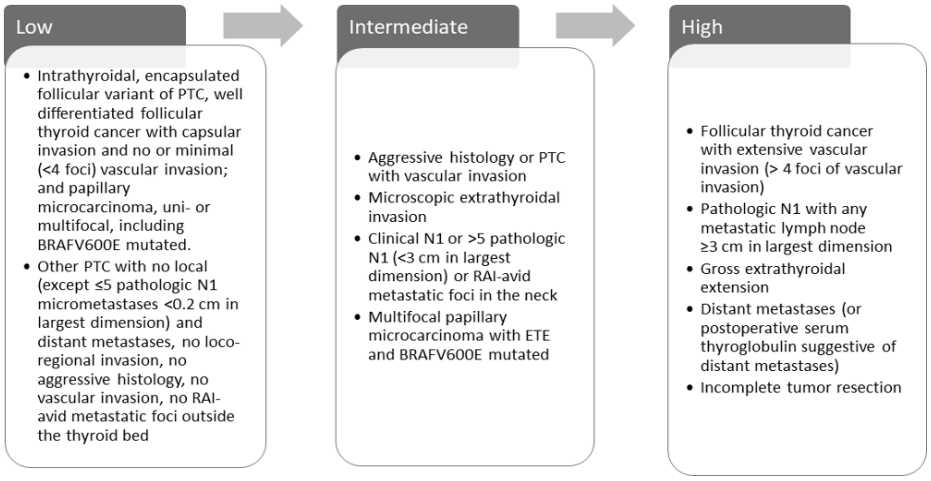

Figure 1. Risk of persistent or recurrent disease according to the ATA risk stratification system. PTC: papillary thyroid cancer, ETE: extrathyroidal extension, RAI: radioactive iodine.

Thyroid
Real-world performance of ATA risk estimates in predicting 1-year DTC outcomes: A prospective multicenter study of 2000 patients (ITCO Study #3) (DOI: 10.1089/thy.2020.0272)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

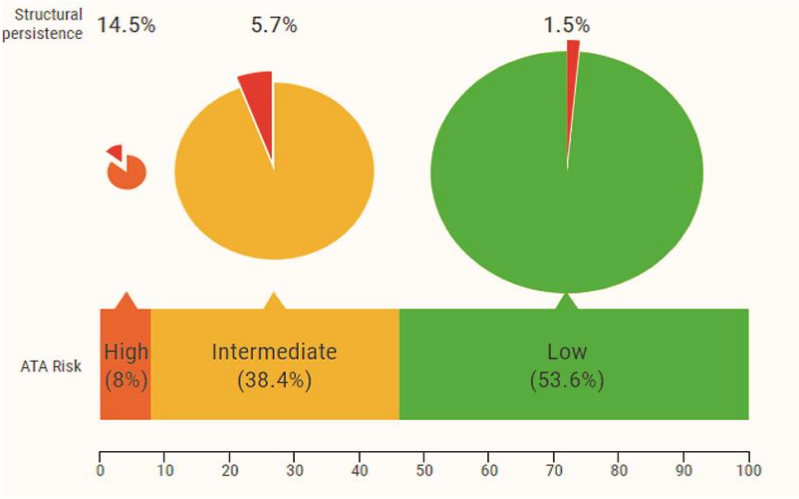

Figure 2. Prevalence of patients classified as low, intermediate, and high risk, and their rates of structural persistent disease.

Thyroid
Real-world performance of ATA risk estimates in predicting 1-year DTC outcomes: A prospective multicenter study of 2000 patients (ITCO Study #3) (DOI: 10.1089/thy.2020.0272)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

Supplemental Table 1: Criteria for classifying responses to treatment

	Initial treatment		
Responses to treatment	**TT+RRA**	**TT alone**	**Lobectomy**
Excellent	Negative imaging AND	Negative imaging AND	Negative imaging AND
	Undetectable TgAb AND	Undetectable TgAb AND	Undetectable TgAb AND
	Tg <0.2 ng/mL OR S-Tg <1 ng/mL	Tg <0.2 ng/mL	Stable Tg levels
Biochemical incomplete	Negative imaging AND	Negative imaging AND	Negative imaging AND
	Tg ≥1 ng/mL OR S-Tg ≥10 ng/mL OR Rising TgAb levels	Tg >5 ng/mL OR Rising Tg values with similar TSH levels OR Rising TgAb levels	Rising Tg values with similar TSH levels OR Rising TgAb levels
Structural incomplete	Imaging evidence of disease (regardless of serum Tg or TgAb levels)		
Indeterminate	Nonspecific imaging findings OR	Nonspecific imaging findings OR	Nonspecific imaging findings
	Faint uptake in thyroid bed on RAI scanning OR		

Thyroid

Real-world performance of ATA risk estimates in predicting 1-year DTC outcomes: A prospective multicenter study of 2000 patients (ITCO Study #3) (DOI: 10.1089/thy.2020.0272)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

	Tg 0.2-1 ng/mL OR S-Tg 1-10 ng/mL OR TgAb stable OR declining in patient with no imaging evidence of disease	Tg 0.2–5 ng/mL OR TgAb levels stable or declining in the absence of structural or functional disease	

As modified from the 2015 ATA ongoing risk stratification (response to therapy) system by the European Society of Medical Oncology Guidelines[9]

Abbreviations: ATA, American Thyroid Association; RAI, radioactive iodine; RRA, radioactive iodine remnant ablation; S-Tg, TSH-stimulated serum thyroglobulin; Tg, serum thyroglobulin; TgAb, anti-Tg antibodies; TT, total thyroidectomy.

Thyroid
Real-world performance of ATA risk estimates in predicting 1-year DTC outcomes: A prospective multicenter study of 2000 patients (ITCO Study #3) (DOI: 10.1089/thy.2020.0272)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.