Journal of the Endocrine Society, 2023, **7**, 1–8
https://doi.org/10.1210/jendso/bvad102
Advance access publication 3 August 2023
Clinical Research Article

Thyroglobulin Cutoff Values for Detecting Excellent Response to Therapy in Patients With Differentiated Thyroid Cancer

Jennifer A. Sipos,[1] Joseph Aloi,[2] Andrew Gianoukakis,[3,4] Stephanie L. Lee,[5] Joshua P. Klopper,[6,*] Jacqueline T. Kung,[7] Mark A. Lupo,[8] David Morgenstern,[9,] Cristina Prat-Knoll,[10] Andre Schuetzenmeister,[11] and Whitney S. Goldner[12]**

[1]Division of Endocrinology, Diabetes and Metabolism, The Ohio State University, Columbus, OH 43210, USA
[2]Division of Endocrinology, Diabetes and Metabolism, Atrium Health Wake Forest Baptist, Winston-Salem, NC 27101, USA
[3]Division of Endocrinology, The Lundquist Research Institute at Harbor-UCLA Medical Center, Torrance, CA 90502, USA
[4]David Geffen School of Medicine, University of California—Los Angeles, Los Angeles, CA 90095, USA
[5]Department of Medicine, Boston Medical Center, Boston, MA 02118, USA
[6]Department of Endocrinology, Kaiser Permanente of Colorado, Denver, CO 80920, USA
[7]Division of Endocrinology, Diabetes and Metabolism, Tufts Medical Center, Boston, MA 02111, USA
[8]The Thyroid & Endocrine Center of Florida, Sarasota, FL 34231, USA
[9]Clinical Development and Medical Affairs, Roche Molecular Systems, Pleasanton, CA 94588, USA
[10]Clinical Development and Medical Affairs, Roche Diagnostics GmbH, 68305 Mannheim, Germany
[11]Biostatistics, Roche Diagnostics GmbH, 68305 Mannheim, Germany
[12]Division of Endocrinology, Diabetes and Metabolism, University of Nebraska Medical Center, Omaha, NE 68198, USA

Correspondence: Jennifer A. Sipos, MD, 1581 Dodd Drive, 5th Floor McCampbell Hall, Columbus, OH 43210. Email: jennifer.sipos@osumc.edu, Reprint requests: Cristina Prat-Knoll. Email: cristina.prat-knoll@roche.com
*Affiliation at the time of the study. Current affiliation is Veracyte Inc., South San Francisco, CA, USA.
**Affiliation at the time of the study. Current affiliation is DELFi Diagnostics, Baltimore, MD, USA.

Abstract

Context: Serum thyroglobulin (Tg) is a biochemical marker for detecting persistent or recurrent differentiated thyroid carcinoma (DTC) post-thyroidectomy. Tg can indicate DTC before structural disease (SD) is visible with imaging procedures.

Objective: This work aimed to evaluate the clinical performance of the Elecsys® Tg II assay at a Tg cutoff of 0.2 ng/mL for ruling out SD in adults with DTC after total/near-total thyroidectomy, with or without radioiodine ablation (RAI).

Methods: Patients were enrolled into 2 cohorts: longitudinal (Tg assessed every 6 months over 2 years under thyroid-stimulating hormone [TSH] suppression therapy following thyroidectomy with or without RAI) and cross-sectional with confirmed SD (Tg assessed once >12 weeks after thyroidectomy). Analyses were performed for both cohorts combined and in the longitudinal cohort.

Results: The study included 530 clinically evaluable samples, the majority (n = 424 samples) from patients who had not received RAI treatment. Following correction for SD prevalence (4.97% in the longitudinal cohort), an Elecsys Tg II cutoff of 0.2 ng/mL ruled out SD with a negative predictive value of 99.9% (95% CI, 99.5%-100%). The assay had excellent sensitivity (98.5%-100%) and acceptable specificity (53.4%-53.5%) for detecting SD (Tg ≥ 0.2 ng/mL) for both cohorts combined and in the longitudinal cohort, with similar findings in RAI-treated and non-RAI-treated subgroups.

Conclusion: In this cohort of DTC patients post-thyroidectomy, a Tg cutoff of 0.2 ng/mL was highly effective for ruling out the presence of SD under TSH-suppressed conditions, including in patients who had not received RAI treatment.

Key Words: thyroglobulin, thyroglobulin assay, differentiated thyroid carcinoma, structural disease recurrence, radioiodine

Abbreviations: ATA, American Thyroid Association; DTC, differentiated thyroid carcinoma; NPV, negative predictive value; PPV, positive predictive value; RAI, radioiodine ablation; SD, structural disease; Tg, thyroglobulin; TSH, thyrotropin (thyroid-stimulating hormone).

Thyroid cancer is the most common endocrine malignancy in the United States, with approximately 43 800 new cases recorded in 2022, approximately 2.3% of all new cancer diagnoses [1]. Incidence has risen sharply in recent decades, largely due to increasing rates of incidental detection of small tumors (<1 cm in diameter) through improved imaging modalities and potential environmental and lifestyle-related factors, such as increasing rates of obesity [2-4]. Differentiated thyroid cancer (DTC), including mostly papillary and follicular cancers, constitutes the vast majority (>90%) of all thyroid cancers [5, 6] and has an excellent prognosis, with an overall 5-year survival rate of ≥95% [1, 3, 7].

Received: 1 March 2023. **Editorial Decision:** 25 July 2023. **Corrected and Typeset:** 9 August 2023
© The Author(s) 2023. Published by Oxford University Press on behalf of the Endocrine Society.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs licence (https://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial reproduction and distribution of the work, in any medium, provided the original work is not altered or transformed in any way, and that the work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

The American Thyroid Association (ATA) classifies patients as having a low- (<5%), intermediate- (5%–20%), or high- (>20%) risk of recurrence [8]. Currently, >80% of patients are classified as low-risk, with most cases managed through lobectomy or total or near-total thyroidectomy alone. For intermediate- to high-risk patients, total or near-total thyroidectomy is usually followed by radioiodine ablation (RAI) plus thyroid-stimulating hormone (TSH)-suppressive therapy, based on individualized risk assessment [8]. Given that the risk for disease recurrence evolves over time, depending on clinical disease course and response to treatment, long-term monitoring is recommended for all patients [8, 9].

Serum thyroglobulin (Tg) is a critical biochemical marker for detecting persistent or recurrent DTC following total or near-total thyroidectomy [10]. As Tg can only be produced by thyroid follicular cells (benign remnant tissue or differentiated thyroid cancer), it is a good indicator of persistent or recurrent DTC even before structural disease (SD) is visible with imaging procedures such as ultrasonography or computed tomography scans [8]. Second-generation Tg immunometric assays with a limit of quantitation ≤ 0.1 ng/mL and precision at low concentrations allow for reliable Tg monitoring without the need for TSH stimulation [11-13].

Currently, standard posttreatment follow-up evaluation for DTC includes a clinical examination, periodic Tg panel (Tg and anti-Tg antibodies), and neck ultrasound, combined with other imaging methods when indicated. ATA guidelines recommend that in non-TSH-stimulated patients, without evidence of SD and Tg antibodies, Tg levels <0.2 ng/mL are consistent with an excellent response to therapy and Tg levels ≥1.0 ng/mL with a biochemical incomplete response [8]. However, for patients who have undergone less than total thyroidectomy or have had a total thyroidectomy without RAI remnant ablation, the guidelines state that while there are no known specific cutoffs to distinguish normal residual thyroid tissue from DTC, rising Tg levels over time are considered suspicious for growing thyroid tissue or cancer [8]. Notably, a recent meta-analysis of 2455 patients did not support the use of serum Tg levels in monitoring patients with low-risk DTC, post lobectomy [14].

In this study, we evaluated the clinical performance of the new Roche Elecsys® Tg II assay for ruling out SD under TSH-suppression in patients with DTC without the presence of anti-Tg antibodies who had received a total (or near-total) thyroidectomy, including a large subgroup who had not received subsequent RAI treatment. The ability of the assay to detect SD was assessed at various Tg II cutoffs, including those reflecting excellent response to therapy (Tg < 0.2 ng/mL) and biochemical incomplete response (Tg ≥ 1.0 ng/mL).

Methods

Design and Patients

This prospective, multicenter, observational study included adults in the United States with histologically confirmed DTC including papillary, follicular, and oncocytic (formerly Hürthle cell) thyroid cancer, in addition to the follicular variant of papillary thyroid cancer. Eligible patients were aged ≥22 years, had been diagnosed with DTC, and had undergone total or near-total thyroidectomy prior to enrollment; the timing of said total or near-total thyroidectomy dictated whether patients were eligible for the longitudinal (surgery within 4-12 weeks prior to enrollment) or cross-sectional (surgery >12 weeks prior to enrollment) study cohorts. Postoperative radioiodine (I-131) administration for remnant ablation or adjuvant therapy was at the treating physician's discretion. Key exclusion criteria included the presence of anti-Tg antibodies, elevated TSH levels indicating nonsuppressed/stimulated Tg, and medullary or anaplastic histology (including tumors that have any component of poorly differentiated histology if not classified as anaplastic).

Patients in the longitudinal cohort (Fig. 1) were enrolled within 4 to 12 weeks of thyroidectomy, and prior to RAI if applicable; SD status was unknown at the time of enrollment. Subsequent assessments were performed at up to 5 timepoints over 2 years following surgery: 4 to 12 weeks postsurgery (baseline) and approximately 6 months (±1 month [follow-up visit 1]) postsurgery/RAI (if applicable), and then every 6 months (±3 months) after the previous follow-up visit (follow-up visits 2-4). A neck ultrasound was performed at each visit to confirm the presence or absence of SD and any interim visits were included as clinically indicated. Additional cross-sectional or functional imaging or biopsy was at the treating physician's discretion. In another analysis from this study (to be published separately), findings from the Elecsys Tg II assay were compared with an established Tg assay (Beckman-Coulter Tg [analysis to be reported in separate publication]); therefore, evaluable samples from the longitudinal cohort for which measurements were not available on both the Elecsys and comparison assay were excluded from the statistical analysis. The cross-sectional cohort included patients with confirmed SD who were not eligible for the longitudinal cohort (ie, >12 weeks since receiving total or near-total thyroidectomy); residual samples for these patients were collected from 8 US collection study sites and the Mayo Clinic Validation Support Services (Fig. 1). SD was defined as evidence of disease on cross-sectional imaging (ultrasound, computed tomography, or magnetic resonance imaging); by functional imaging (positive RAI scan or positron-emission tomography with 2-deoxy-2-[fluorine-18] fluoro- D-glucose scan) or biopsy (cytology or histology). For the longitudinal cohort, a Board-Certified endocrinologist with extensive experience in the management of DTC (~20 years) reviewed each subject's disease status assessment (structural disease absent or present) to ensure it was adequately supported by the clinical data collected. Although the reviewer had access to data casebooks containing clinical laboratory, imaging, and relevant study data to query any discrepancies between the disease status assessment and clinical, laboratory, or imaging results, the principal investigator at each study site was responsible for the determination of the patient's disease status. Both the medical reviewer and the principal investigator were blinded at all times to the patient's Elecsys Tg II result; results of the Beckman-Coulter Tg result were available and used for patient management only (and not for definition of structural disease).

This study was conducted in accordance with all Food and Drug Administration (FDA) and local regulations, and the study protocol was approved by Institutional Review Boards from each study site. All participants provided informed written consent.

Sample Handling and Tg Measurement

Serum samples were collected by venous blood draw during follow-up visits and were analyzed at testing laboratories in the United States (University of Maryland, MD; Indiana

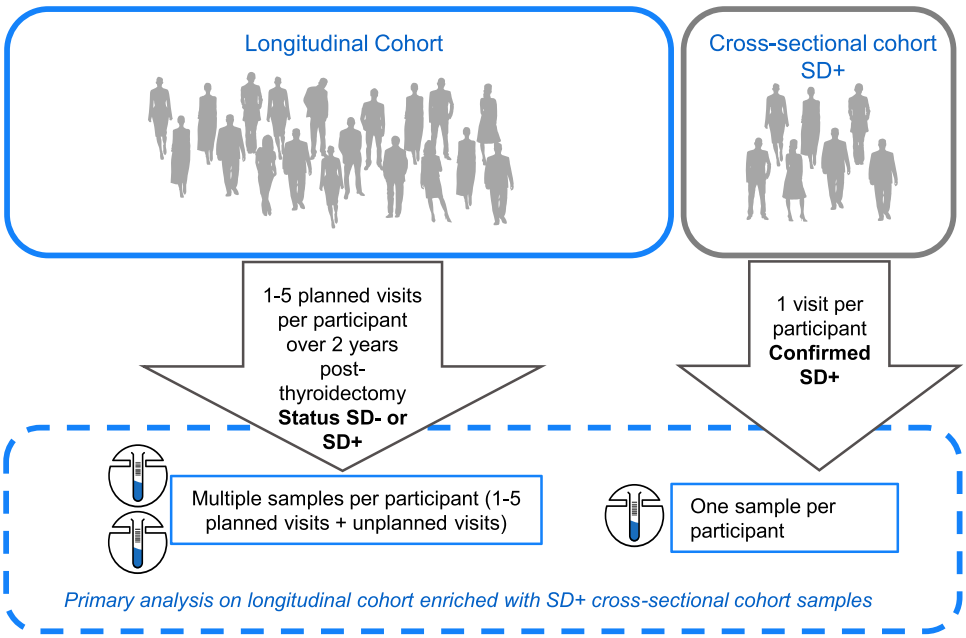

Figure 1. Study design. Abbreviations: SD, structural disease.

Health Pathology Laboratory, IN; University of California Davis Health, CA). All study samples were drawn from patients on levothyroxine therapy. Tg measurements were considered to be TSH-non-suppressed if TSH was >5.0 µIU/mL. A study visit of cross-sectional patients which showed a TSH >5.0 µIU/mL was excluded from the statistical analysis. The Elecsys Tg II assay results were measured on the cobas e 411 system and processed according to the manufacturer's instructions. This quantitative electrochemiluminescence immunoassay (ECLIA) involves a two-step, double antigen sandwich process and has a measuring range of Tg from 0.1 to 500 ng/mL, with a limit of quantitation of 0.10 ng/mL.

Objectives and Analyses

The primary objective was to show that the assay, at a Tg cutoff of 0.2 ng/mL, can be used to rule out the presence of SD over 24 months postsurgery. At a Tg cutoff value of 0.2 ng/mL, Tg measurements <0.2 ng/mL demonstrate the absence of SD. Secondary objectives were to assess clinical performance (sensitivity, specificity, negative predictive value [NPV], and positive predictive value [PPV]) at a Tg cutoff of 0.2 ng/mL.

Further outcomes included probability for being SD-positive (SD+) conditional on certain Tg measurement ranges according to ATA response criteria [8]: excellent response (Tg <0.2 ng/mL), indeterminate response (Tg ≥0.2 ng/mL and <1.0 ng/mL), and biochemical incomplete response (Tg ≥1.0 ng/mL). We also assessed probabilities for being SD+ or SD− at other Tg II cutoffs above 1 ng/mL.

Statistical Methods

Sample size was estimated using 2-by-2 contingency tables comparing disease status (presence or absence of SD) and Tg test result of <0.2 ng/mL or ≥0.2 ng/mL, for a range of sample sizes for the longitudinal cohort. It was initially calculated that 324 observations (evaluable Tg II assessments) with a minimum of 65 samples from SD+ patients would be needed to satisfy the acceptance criterion for the primary objective (NPV ≥95% with a lower bound of the two-sided 95% CI ≥90%). However, due to fewer SD+ patients than expected, more patients than initially planned were subsequently enrolled in the longitudinal cohort. To have sufficient samples from SD+ patients, samples from a cross-sectional cohort who had current evidence of SD and were not eligible for the longitudinal cohort (due to >12 weeks having elapsed from surgery) were collected (Fig. 1).

The primary objective was assessed in the full analysis set, which included eligible patients from the longitudinal and cross-sectional cohorts with clinically evaluable samples (ie, those from patients meeting all inclusion criteria, for whom an Elecsys Tg II value was available; see Supplementary Methods [15]). Other performance measurements were measured in the full analysis set or in RAI-treated and non-RAI-treated subgroups.

Unless stated otherwise, all performance measures were derived from a 2-by-2 contingency table, with clinical truth based on the presence or absence of SD and a Tg test result at the 0.2 ng/mL cutoff. Acceptance criteria for the primary objective was an NPV ≥95%, with the lower bound of the 95% bootstrap CI ≥90%. Two-sided 95% CIs of point-estimates were constructed using a bootstrap approach on the patient level (as per FDA guidance), reflecting within-subject correlation of Tg values, and based on 10 000 bootstrap iterations. Intra-patient correlation was addressed by performing CI estimation via a bootstrap approach, where bootstrapping took place at the patient level, ie, all visits/samples of a patient were included in a bootstrap iteration if that subject was randomly selected. For all bootstrap iterations, the corrected NPV was computed based on the real-world prevalence of SD (for samples from the longitudinal cohort), using Bayes Theorem; the 2.5%- and 97.5%-quantiles were derived from these values, corresponding to 95% bootstrap CI. 2-by-2 contingency tables were constructed for RAI-treated and non-RAI-treated subgroups separately. Statistical analyses were performed using R v3.4.0.

Table 1. Patient demographics

Demographics	RAI-treated (N = 103)	Non-RAI-treated (N = 117)	Overall (N = 220[a])
Mean age, years (range)	55.7 (24.0–97.0)	53.1 (23.0–85.0)	54.0 (23.0–97.0)
Sex, n (%)			
Female	52 (50.5)	84 (71.8)	136 (61.8)
Male	51 (49.5)	33 (28.2)	84 (38.2)
Race, n (%)			
African American/Black	5 (4.9)	5 (4.3)	10 (4.5)
Asian	2 (1.9)	3 (2.6)	5 (2.7)
Native Hawaiian or Other Pacific Islander	1 (1.0)	0	1 (0.5)
White	87 (84.5)	97 (82.9)	184 (83.6)
Other	2 (1.9)	7 (6.0)	9 (4.1)
Not available	6 (5.8)	5 (4.3)	11 (5.0)
ATA risk class, n (%)			
Low	19 (18.4)	66 (56.4)	85 (38.6)
Intermediate	22 (21.4)	34 (29.1)	56 (25.5)
High	62 (60.2)	17 (14.5)	79 (35.9)
Number of samples			
Median (IQR)			3.0 (2.0–4.0)
Mean			3.1

Abbreviations: ATA, American Thyroid Association; IQR, interquartile range; RAI, radioiodine ablation.
[a]One patient did not subsequently have measurements available on both Elecsys and an established comparison Tg assay device (separate analysis—see "Methods") and was therefore excluded from the full analysis set (N = 219); N is the number of patients with clinically evaluable samples.

Results

Participants and Samples

A total of 219 patients with evaluable samples were included in this study: 150 patients from the longitudinal cohort and 69 patients from the SD-enriched cross-sectional cohort. Overall, patients were predominantly female (61.8%) and White (83.6%) and had a mean age of 54.0 years (range, 23.0–97.0) (Table 1). In total, 103 patients were RAI-treated, and 117 patients were non-RAI-treated; 60.2% of RAI-treated patients were classified as high risk (ATA risk class) compared with 14.5% of non-RAI-treated.

In the longitudinal cohort, a total of 705 patient visits were completed over the 2 years, with 80 patients completing the final visit (visit 5). The number of patients who completed each visit and the number of samples provided at each visit is stated in Supplementary Table 1 [15]. In total, 463 evaluable samples were obtained, that is, those that met the inclusion criteria; data regarding TSH levels are available in Supplementary Table 2. Two samples were subsequently excluded from the full analysis set due to measurements not being available on both Elecsys and the established comparison Tg assay (separate analysis—see "Methods") (Fig. 2). In the cross-sectional cohort, a total of 76 samples were obtained and 69 were deemed evaluable. Therefore, the full analysis set comprised 530 samples from 219 patients (461 samples from the longitudinal cohort and 69 samples from the cross-sectional cohort; Fig. 2). The median number of samples per patient in the full analysis set was 3.0 (interquartile range, 2.0-4.0).

Using the Elecsys Tg II assay, 235 (44.3%) samples had Tg <0.2 ng/mL, most of which were from patients who had not received RAI treatment (Table 2). Of patients with Tg <0.2 ng/mL, only one sample—with a Tg of 0.189 ng/mL—had ultrasound findings suspicious for SD. This patient (a 32-year-old woman from the cross-sectional cohort) had a diagnosis of papillary thyroid carcinoma with extra-nodal extension and underwent initial surgery and RAI in 2013. In 2015, she had a biopsy of a left level II/III lymph node that was confirmed malignant. There was additional surgery consisting of neck dissection with complete resection of visible disease (36 nodes). During her study visit, a left level VI mass (0.9 × 0.5 cm) and a 0.6 cm level II abnormal lymph node were observed upon ultrasound; both fall below the ATA guidelines threshold for biopsy (recommended when the smallest dimension is >0.8 cm). The patient preferred surveillance over additional intervention.

In analyses of patient Tg levels over time, Tg levels were generally higher in SD+ vs SD− patients (Fig. 3). Tg levels tended to decrease over time in both SD+ and SD− patients, albeit with a proportionally greater decrease in the SD− group (Fig. 3). Further analyses of Tg levels over time are presented in the supplementary material, including by response category (excellent, indeterminate, structural incomplete, biochemical incomplete; Supplementary Fig. 1 [15]), by RAI treatment (Supplementary Fig. 2 [15]), and by ATA risk category (low, intermediate, high; Supplementary Fig. 3 [15]). Descriptive statistics of Tg levels by ATA risk category in RAI-treated and non-RAI-treated subgroups are also described (Supplementary Table 3 [15]).

Clinical Performance

Following correction for the real-world prevalence of SD (estimated from the longitudinal cohort as 4.97%), prespecified acceptance criteria were exceeded for the primary analysis

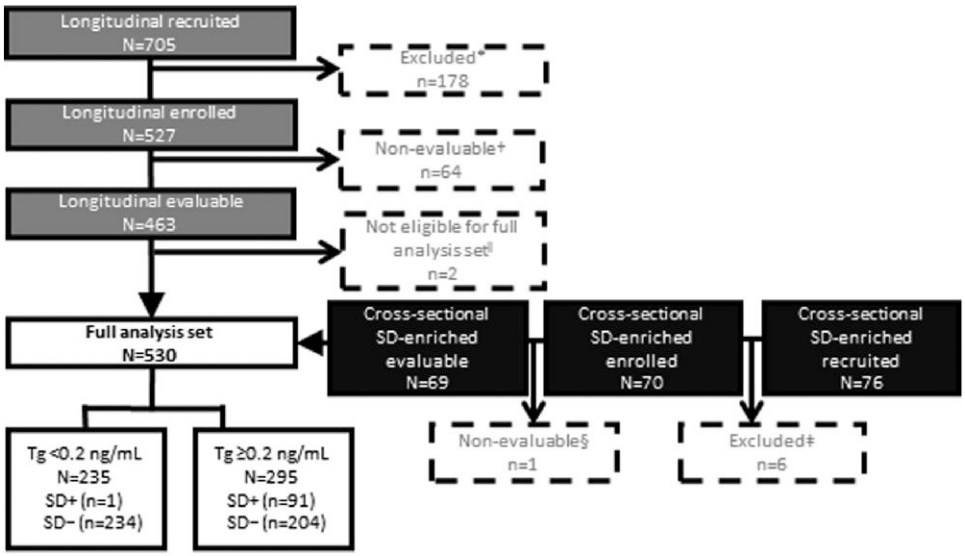

Figure 2. Sample disposition. N refers to number of samples by cohort. *Reasons for exclusion: COVID-19 (n = 57), positive for anti-Tg (n = 53), investigator, sponsor, or patient request (n = 37), other (n = 18), not eligible (n = 4), lost to follow-up (n = 4), withdrawal of consent (n = 4), and pregnancy (n = 1). †Reasons that samples were deemed not evaluable: TSH >5.0 U/mL (n = 54), neck ultrasound imaging >4 weeks from Tg sample collection (n = 4), visit did not take place (n = 3), Tg II assay value unavailable (n = 2), improper sample handling (n = 1). ‡Reasons for exclusion: not eligible (n = 4), positive for anti-Tg (n = 2). §Reason that sample was deemed not evaluable: Tg II assay value unavailable (n = 1). ‖Two samples from the longitudinal cohort did not have measurements available on both Elecsys and an established comparison Tg assay device (separate analysis—see "Methods") and were therefore excluded from the full analysis set. Abbreviations: SD, structural disease.

Table 2. Detecting the presence or absence of structural disease based on samples tested in the Elecsys Tg II assay (full analysis set)

	RAI-treated (N = 108)		Non-RAI-treated (N = 422)		Overall (N = 530)	
	Tg <0.2 ng/mL (N = 28)	Tg ≥0.2 ng/mL (N = 80)	Tg <0.2 ng/mL (N = 207)	Tg ≥0.2 ng/mL (N = 215)	Tg <0.2 ng/mL (N = 235)	Tg ≥0.2 ng/mL (N = 295)
Structural disease						
Present	1[a] (3.6)	63 (78.8)	0	28 (13.0)	1 (0.4)	91 (30.8)
Absent	27 (96.4)	17 (21.3)	207 (100)	187 (87.0)	234 (99.6)	204 (69.2)

Abbreviations: RAI, radioiodine ablation; Tg, thyroglobulin.
[a]Presumed structural disease (patient did not have biopsy confirmation); Data shown are the number (percent) of samples in each group; N is the number of clinically evaluable samples.

(full analysis set) and when analyzed in the longitudinal cohort (Table 3). The NPV was high (99.9%; 95% CI, 99.5%-100%) in the full analysis set and 100% (95% CI, 100%-100%) in the longitudinal cohort, both of which were above the 95% criterion with the lower bound of 95% CI ≥90%. The Tg II assay, at a Tg cutoff value of 0.2 ng/mL, demonstrated sensitivity of 98.5% to 100%, and specificity of 53.4% to 53.5% (Table 3). Levels of clinical performance at a Tg cutoff of 0.2 ng/mL were similar in RAI-treated and non-RAI-treated subgroups (Supplementary Table 4 [15]).

The estimated probability of being SD+ with Tg <0.2 ng/mL was 3.3% in the RAI-treated subgroup, and 0% in the non-RAI-treated subgroup (Table 4). At Tg levels representing a biochemical incomplete response (≥1.0 ng/mL), there was an 87.0% probability of SD in the RAI-treated subgroup, and 28.2% probability in the non-RAI-treated subgroup (Table 4).

We further analyzed the probability of being SD+ or SD− at several Tg cutoff values above 1 ng/mL. The probability of being SD+ increased as the cutoff value increased, for Tg values up to 5 ng/mL, in RAI-treated and non-RAI-treated patients (Supplementary Fig. 4) [14, 15]. For Tg values above 5 ng/mL, the numbers of SD+ patients were too few for meaningful interpretation.

Discussion

In this large cohort of patients with DTC following total or near-total thyroidectomy with or without subsequent RAI therapy, a Tg cutoff value of 0.2 ng/mL measured by the Elecsys Tg II assay had an NPV of 99.9% (95% CI, 99.5%-100%), with excellent sensitivity (99.9%-100%) and acceptable specificity (53.4%-53.5%) for detection of SD. The high value achieved for NPV is particularly notable given that the study population was enriched for patients with

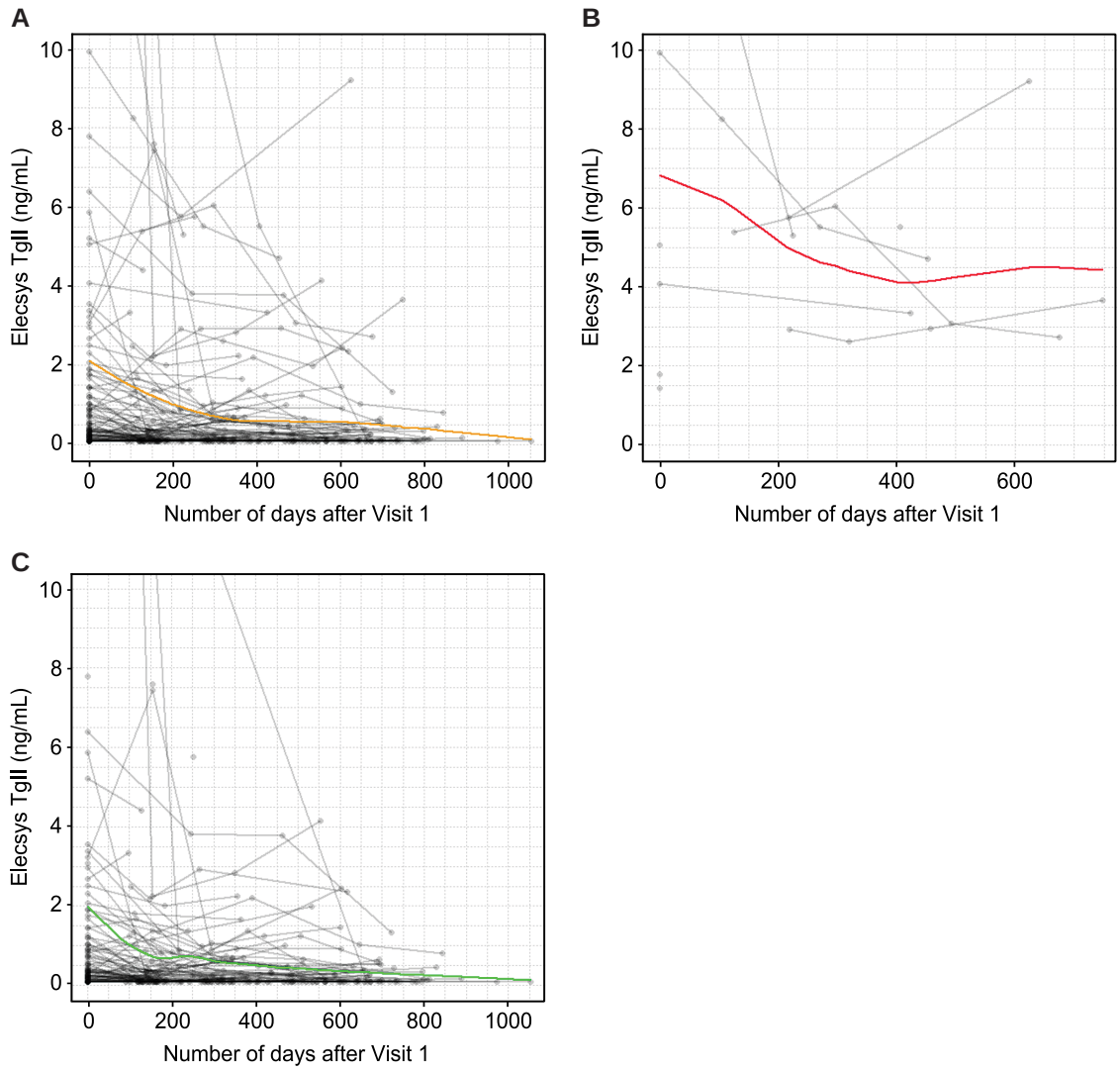

Figure 3. Analysis of Tg values (up to 10 ng/mL) over time for (A) combined, (B) SD+, and (C) SD− groups in the longitudinal cohort. Loess scatterplot smoother added to spiderplot showing a restricted range of Tg values. The Loess-smoother allows visualization of local trends by applying a locally weighted averaging strategy which can be tuned to more or less smoothing via parameter "span." Note that for time-course data, only the longitudinal cohort was used, as the cross-sectional cohort contains a single sample per subject only. Additionally, only patients with at least 2 samples were included in the total analysis. Single visits appear in the SD+ and SD− groups as these groups classify disease status at each visit, and as disease state can change over time a patient may have one visit as SD− and one as SD+. Abbreviations: SD, structural disease; Tg, thyroglobulin.

SD, which may have been expected to reduce the NPV. Similar findings were observed both for patients treated with or without RAI ablation treatment. In both RAI- and non-RAI-treated populations, only one patient with Tg <0.2 ng/mL had SD. This patient had small volume disease which was not histologically confirmed and Tg 0.189 ng/mL. As current ATA guidelines state that specific Tg cutoff levels to indicate SD are unknown in patients who have not received RAI treatment [8], these findings are highly informative, and support the use of a Tg cutoff of 0.2 ng/mL. Moreover, in patients with Tg ≥1.0 ng/mL (biochemical incomplete response) at any time during the 2-year follow-up, there was a high likelihood of SD, and the likelihood was substantially lower in non-RAI vs RAI-treated patients (28% vs 87%, respectively). This may be due to the distribution of ATA risk categories within each treatment group. Indeed, the lower frequency of SD observed in non-RAI patients was expected given the lower proportion of ATA high-risk patients in this subgroup. Also, non-RAI patients were likely to have had less extensive disease (eg, no invasion or significant lymph node metastases) leading to a lower chance of SD even if their Tg levels exceeded 1.0 ng/mL. Another possible reason for this low likelihood may be that Tg continues to be produced by the benign thyroid remnant still present in non-RAI-treated patients [8]. There have been few studies assessing Tg cutoff in non-RAI patients following total/near-total thyroidectomy, with most concerning retrospective reviews and very low rates of recurrence/persistence [16]. While our study adds a prospective study to this limited body of literature, further prospective studies are warranted to determine an optimal Tg threshold in this population.

In a previous meta-analysis of 3178 DTC patients, unstimulated Tg measurements had high NPVs (97% and 99%) at Tg cutoffs for positivity of 1 and 2 ng/mL, respectively, suggesting that TSH-stimulated measurement was not needed in the follow-up of patients with undetectable basal Tg, provided

no Tg antibodies are present [17]. However, the PPV (70% at a Tg cutoff of 1 ng/mL) was insufficient to eliminate TSH-stimulated measurement from the follow-up of patients with detectable unstimulated Tg [17]. Indeed, the use of second-generation Tg assays with functional sensitivity ≤0.1 ng/mL removes the need for TSH stimulation [11], and in low- and intermediate-risk patients, ATA guidelines recommend monitoring Tg levels without TSH stimulation. However, in high-risk patients, ATA guidelines recommend maximum TSH stimulation to assess Tg levels for 6 to 18 months after the completion of primary therapy [8].

At TSH levels ranging from undetectable to ≤5.0 μIU/mL, the Tg II assay was highly sensitive in all but one patient in our study, suggesting that a Tg cutoff value of 0.2 ng/mL may help to rule out the presence of SD in both RAI-treated and non-RAI-treated patients. This may reduce the need for TSH-stimulated Tg and whole-body scans, and reduce the frequency of ultrasound imaging visits necessary to confirm the presence of SD. As our study excluded patients with less-differentiated tumors, it is possible that tumors producing low Tg levels could be present, even when measured Tg is <0.2 ng/mL. Therefore, careful follow-up, including periodic neck ultrasound, is warranted, with consideration for additional procedures such as whole-body scan, fine-needle aspiration, or positron-emission tomography, as clinically indicated.

Table 3. Clinical performance of the Tg II assay at a Tg cutoff of 0.2 ng/mL for the detection of structural disease

	Full analysis set N = 530 Estimate % (95% CI)	Longitudinal cohort N = 461 Estimate % (95% CI)
NPV (uncorrected)	99.6 (98.6–100)	100 (100–100)
PPV (uncorrected)	23.9 (21.0–27.5)	10.0 (4.5–16.9)
NPV (corrected)[a]	99.9 (99.5–100)	100 (100–100)
PPV (corrected)[a]	10.0 (8.7–11.7)	10.1 (8.8–11.82)
Sensitivity	98.5 (95.4–100)	100 (100–100)
Specificity	53.4 (45.8–61.1)	53.5 (45.8–61.0)
Prevalence	12.9 (12.1–13.9)	4.89 (2.15–8.49)

Abbreviations: NPV, negative predictive value; PPV, positive predictive value, Tg, thyroglobulin.
[a]Corrected based on the real-world prevalence of 4.97% estimated from samples of the longitudinal cohort only; "Estimate" corresponds to the median of N = 10 000 bootstrap iterations on the patient level as well as the 95% CI, which is derived from the set of bootstrap-samples as the 2.5% and 97.5% quantiles; N is the number of clinically evaluable samples.

In a previous cost-effectiveness analysis, the cost to detect recurrence in ATA low-risk patients (taking into account post-operative blood tests, imaging scans, biopsies, and clinic visits) was more than 6- to 7-times higher per recurrence than the cost for patients with intermediate or high risk [18]. Identifying a more cost-effective follow-up strategy for this large subset of patients is thus warranted. The use of an assay that can accurately identify the absence of SD in such low-risk patients could reduce reliance on regular imaging scans and may thus represent a more cost-effective follow-up strategy for this large subset.

The strength of this study was the large, longitudinal, multi-site cohort, providing a population with diverse treatment patterns. The study also included a large proportion of patients who did not receive RAI ablation. This reflects current practice in the United States and provided an opportunity to evaluate Tg cutoffs in this particular clinical context where appropriate cutoffs have not yet been defined. A limitation of the study was the interruption caused by COVID-19 to recruitment and procedures. Other limitations include a potential introduction of bias and the enrichment of cases with SD+. As further imaging was performed at the treating physician's discretion, it could be argued that patients with an elevated Tg were more likely to undergo further evaluation to locate SD and thus become more likely to have SD discovered. However, it could be argued that the overall cohort were all investigated more rigorously than the current standard clinical practice in the United States, with ultrasound performed every 6 months regardless of ATA risk category or Tg II results. The nature of the study required the enrichment of SD+ cases to ensure that the study was powered for the primary objective. Samples were initially obtained from 19 SD+ patients at the sites enrolling into the longitudinal cohort; however, these patients were not eligible as ≥12 weeks had elapsed since their initial surgery. When this method proved insufficient to reach 65 SD+ samples, samples were obtained from an additional 50 SD+ patients at Mayo Clinic, all of whom were subject to the same inclusion/exclusion criteria as the longitudinal cohort. Further limitations were that the 2-year follow-up period may have missed patients with slowly growing recurrences.

With a Tg cutoff value of 0.2 ng/mL, the Elecsys Tg II assay showed good clinical performance for ruling out

Table 4. Probability of being positive for structural disease using the Elecsys at cutoffs of 0.2 ng/mL and 1.0 ng/mL, as per ATA response criteria (full analysis set)

	RAI-Treated (N = 108) Estimate % (95% CI)	Non-RAI-Treated (N = 422) Estimate % (95% CI)	Overall (N = 530) Estimate % (95% CI)
Probability of being SD+ and			
Tg < 0.2 ng/mL (excellent response)	3.33 (0–11.76)	0	0.41 (0–1.40)
Tg ≥ 0.2 – <1.0 ng/mL (indeterminate response)	0	0	0
Tg ≥ 1.0 ng/mL (biochemical incomplete response)	87.0 (75.6–95.7)	28.2 (18.8–39.0)	48.2 (41.0–57.1)

N is the number of clinically evaluable samples.
Abbreviations: RAI, radioiodine ablation; SD, structural disease; Tg, thyroglobulin.

SD (and thus defining excellent response to therapy under non-TSH-stimulated conditions) in patients with DTC following total or near-total thyroidectomy, with or without RAI treatment. The assay demonstrated excellent NPV and acceptable PPV, with values for sensitivity and specificity comparable between RAI-treated and non-RAI-treated groups. These findings support the use of the Elecsys Tg II assay and Tg cutoff of 0.2 ng/mL in clinical decision-making following initial surgical management of DTC patients, including those who have not received RAI ablation.

Acknowledgments

The authors would like to thank the patients and their families for their contribution to the study. The authors would like to thank Mike Tuttle of Memorial Sloan Kettering Cancer Center, New York, NY for his contribution to the development of the study design, Alicia Algeciras-Schimnich of Mayo Clinic, Rochester, MN, and Jennifer Cannon of Wake Forest School of Medicine, Medical Center Boulevard, Winston-Salem, NC for their help with data collection, and Poorani Goundan of Boston Medical Center, Boston, MA, USA for her assistance with data and sample collection and medical review. The authors also thank Erik Alexander for his contribution to the study and his expert medical review of the data collected.

Editorial support was provided by Róisín O'Connor, Juliet Gray, and Estelle Challinor of inScience Communications, Springer Healthcare Ltd, UK, and was funded by Roche Diagnostics International Ltd (Rotkreuz, Switzerland). ELECSYS is a trademark of Roche. All other product names and trademarks are the property of their respective owners. The Elecsys Tg II assay is cleared for clinical use in the CE-mark accepting countries. The Elecsys Tg II assay is not approved for clinical use in the United States.

Funding

This study was funded by Roche Diagnostics GmbH (Penzberg, Germany).

Author Contributions

A.S., D.M., and J.A.S. conceptualized the study. All authors were responsible for data acquisition, analysis, and interpretation. All authors were responsible for drafting and the final approval of the manuscript.

Disclosures

A.G., J.A., J.T.K., and S.L.L. have nothing to declare. A.S. and C.P.K. are employed by Roche Diagnostics. D.M. was employed by Roche Diagnostics at the time of the study and has stock options in F Hoffman-La Roche. J.A.S. and M.A.L. have received research funding from Roche. J.P.K. is employed by and has equity interests in Veracyte, Inc. W.S.G. was the site investigator for Siemens (Tg Study).

Data Availability

Some or all datasets generated during and/or analyzed during the current study are not publicly available but are available from the corresponding author on reasonable request.

References

1. Cancer Stat Facts: Thyroid Cancer. National Cancer Institute. Published 2022. Accessed January 25, 2023. https://seer.cancer.gov/statfacts/html/thyro.html
2. Hughes DT, Haymart MR, Miller BS, Gauger PG, Doherty GM. The most commonly occurring papillary thyroid cancer in the United States is now a microcarcinoma in a patient older than 45 years. *Thyroid*. 2011;21(3):231-236.
3. Morris LG, Sikora AG, Tosteson TD, Davies L. The increasing incidence of thyroid cancer: the influence of access to care. *Thyroid*. 2013;23(7):885-891.
4. Kitahara CM, Sosa JA. The changing incidence of thyroid cancer. *Nat Rev Endocrinol*. 2016;12(11):646-653.
5. Sherman SI. Thyroid carcinoma. *Lancet*. 2003;361(9356):501-511.
6. Cabanillas ME, McFadden DG, Durante C. Thyroid cancer. *Lancet*. 2016;388(10061):2783-2795.
7. Brito JP, Morris JC, Montori VM. Thyroid cancer: zealous imaging has increased detection and treatment of low-risk tumours. *BMJ*. 2013;347:f4706.
8. Haugen BR, Alexander EK, Bible KC, *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;26(1):1-133.
9. Tuttle RM, Tala H, Shah J, *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;20(12):1341-1349.
10. Indrasena BS. Use of thyroglobulin as a tumour marker. *World J Biol Chem*. 2017;8(1):81-85.
11. Giovanella L, Clark PM, Chiovato L, *et al*. Thyroglobulin measurement using highly sensitive assays in patients with differentiated thyroid cancer: a clinical position paper. *Eur J Endocrinol*. 2014;171(2):R33-R46.
12. Giovanella L, Ceriani L, Maffioli M. Postsurgery serum thyroglobulin disappearance kinetic in patients with differentiated thyroid carcinoma. *Head Neck*. 2010;32(5):568-571.
13. Gurleyik E, Dogan S. Accuracy of unstimulated basal serum thyroglobulin levels in assessing the completeness of thyroidectomy. *J Clin Med Res*. 2014;6(5):369-373.
14. Giovanella L, Ceriani L, Garo ML. Is thyroglobulin a reliable biomarker of differentiated thyroid cancer in patients treated by lobectomy? A systematic review and meta-analysis. *Clin Chem Lab Med*. 2022;60(7):1091-1100.
15. Sipos JA, Aloi J, Gianoukakis A, *et al*. Data from: Thyroglobulin Cut-off Values for Detecting Excellent Response to Therapy in Patients with Differentiated Thyroid Cancer - Supplementary Material. Figshare Digital Repository 2023. Date of deposit 3 August 2023. http://doi.org/10.6084/m9.figshare.23821455
16. Chou R, Dana T, Brent GA, *et al*. Serum thyroglobulin measurement following surgery without radioactive iodine for differentiated thyroid cancer: a systematic review. *Thyroid*. 2022;32(6):613-639.
17. Giovanella L, Treglia G, Sadeghi R, Trimboli P, Ceriani L, Verburg FA. Unstimulated highly sensitive thyroglobulin in follow-up of differentiated thyroid cancer patients: a meta-analysis. *J Clin Endocrinol Metab*. 2014;99(2):440-447.
18. Wang LY, Roman BR, Migliacci JC, *et al* Cost-effectiveness analysis of papillary thyroid cancer surveillance. *Cancer*. 2015;121(23):4132-4140.