CrossMark

ORIGINAL ARTICLE

Papillary thyroid carcinomas with biochemical incomplete or indeterminate responses to initial treatment: repeat stimulated thyroglobulin assay to identify disease-free patients

Livia Lamartina[1] · Teresa Montesano[1] · Fabiana Trulli[1] · Marco Attard[2] ·
Massimo Torlontano[3] · Rocco Bruno[4] · Domenico Meringolo[5] · Fabio Monzani[6] ·
Salvatore Tumino[7] · Giuseppe Ronga[1] · Marianna Maranghi[1] · Marco Biffoni[1] ·
Sebastiano Filetti[1] · Cosimo Durante[1]

Abstract Papillary thyroid cancer (PTC) patients treated with thyroidectomy and radioiodine remnant ablation (RRA) often have detectable TSH-stimulated thyroglobulin (Tg) levels without localizable disease after primary treatment. To assess the value of repeat stimulated Tg assays in these patients' follow-up, we retrospectively analyzed 86 cases followed in 5 Italian thyroid-cancer referral centers. We enrolled 86 patients with PTCs treated with total/near-total thyroidectomy plus RRA between January 1,1990 and January 31, 2006. In all cases, the initial postoperative visit revealed stimulated serum Tg ≥ 1 ng/mL, negative Tg antibodies, and no structural evidence of disease. None received empiric radioiodine therapy. Follow-up (median: 9.6 years) included neck ultrasound and basal Tg assays (yearly) and at least 1 repeat stimulated Tg assay. Of the 86 patients analyzed (initial risk: low 63 %, intermediate 35 %, high 2 %), one (1 %) had ultrasound-detected lymph node disease and persistently elevated stimulated Tg levels at 3 years. In 17 (20 %), imaging findings were consistently negative, but the final stimulated Tg levels was still >1 ng/mL (median 2.07 ng/mL, range 1.02–4.7). The other 68 (80 %) appeared disease-free (persistently negative imaging findings with stimulated Tg levels ≤ 1 ng/mL). Mean intervals between first and final stimulated Tg assays were similar (5.2 and 4.8 years) in subgroups with versus without Tg normalization. Reclassification as disease-free was significantly more common when initial stimulated Tg levels were indeterminate (<10 ng/mL). In unselected PTC cohorts with incomplete/indeterminate biochemical responses to thyroidectomy and RRA, periodic remeasurement of stimulated Tg allows most patients to be classified as disease-free.

Introduction

The incidence of papillary thyroid cancers (PTC) has tripled since the mid-1970s [1], but outcomes are good in the vast majority of cases (approximately 98 % according to the Surveillance, Epidemiology, and End Results Program Database). Nevertheless, life-long surveillance is still the rule in many centers because disease recurrence has been observed as late as 40 years after the diagnosis [2, 3].

The cornerstones of the surveillance process are the neck ultrasound examination and assays of serum levels of

---

Livia Lamartina and Teresa Montesano have equally contributed to this paper.

✉ Sebastiano Filetti
   sebastiano.filetti@uniroma1.it

1   Dipartimento di Medicina Interna e Specialità Mediche, Dipartimento di Scienze Chirurgiche, Università di Roma Sapienza, Viale del Policlinico, 155, 00161 Rome, Italy

2   Unità Operativa di Endocrinologia, Azienda Ospedaliera Ospedali Riuniti Villa Sofia-Cervello, 90146 Palermo, Italy

3   Unità Operativa di Endocrinologia, Istituto di Ricovero e Cura a Carattere Scientifico, Casa Sollievo della Sofferenza, 71013 San Giovanni Rotondo, Italy

4   Unità di Endocrinologia, Ospedale di Tinchi-Pisticci, 75020 Matera, Italy

5   Unità Operativa Semplice Dipartimentale di Endocrinologia, Ospedale di Bentivoglio, 40010 Bologna, Italy

6   Dipartimento di Medicina Clinica e Sperimentale, Università di Pisa, 56126 Pisa, Italy

7   Dipartimento di Medicina Clinica e Sperimentale, Università di Catania, 95123 Catania, Italy

thyroglobulin (Tg), which is produced exclusively by follicular thyrocytes. The serum Tg assay is a widely available, low-cost procedure with high sensitivity and specificity for detecting persistent or recurrent thyroid cancer [4]. However, detectable Tg levels may reflect the residual presence of normal thyrocytes and/or the persistence or recurrence of neoplastic thyroid tissue. This diminishes the specificity of the marker, particularly in the growing population of PTC patients who do not undergo postoperative radioiodine remnant ablation (RRA) [5]. However, Tg assay positivity (without structural evidence of disease) is also reported in up to 36 % of the PTC patients who do undergo ablation [6–10]. As for sensitivity, serum Tg assays may also fail to identify small volumes of neoplastic thyroid tissue, especially when they are performed during thyroid hormone suppression of thyroid-stimulating hormone (TSH). Measurement of TSH-stimulated levels of Tg is thus recommended to maximize sensitivity [4]. This approach is complicated by the fact that withdrawal of thyroid hormone replacement therapy—the method traditionally used to increase TSH levels—subjects patients to the unpleasant symptoms of hypothyroidism for approximately 3–6 weeks [11]. This problem can now be avoided with the use of recombinant human TSH (rhTSH) [12], although the gain in tolerability is offset to some extent by a non-negligible cost increase related to the price of the drug.

These considerations justify attempts to rationalize the use of stimulated Tg testing during follow-up of patients with PTCs. Repetition of these assays is of little benefit once the patient meets the criteria for classification as disease-free [13–16]. For those treated with total/near-total thyroidectomy followed by RRA, this means the absence of clinical and imaging evidence of tumor [i.e., neck US and/or diagnostic ^{131}I whole-body scan (dxWBS)] and basal and stimulated Tg levels that are ≤1 ng/mL, in the absence of thyroglobulin antibodies (TgAb) [4]. The negative predictive value of all of these findings is almost 100 % [17–21]. Clinical or imaging (cross-sectional or functional) findings compatible with persistent macroscopic disease usually prompt additional treatment or at least a more active form of surveillance [4]. There is no consensus, however, regarding the management of individuals with detectable Tg levels in the absence of localizable disease, and the benefit from repeating TSH-stimulated testing in this patient population is uncertain. Reported outcomes in these cases vary widely, with imaging-documented evidence of disease emerging in 8–50 % of patients. In similar percentages (13–50 %), however, the Tg level reportedly diminishes progressively and at some point becomes undetectable, thereby allowing the patient to be classified as disease-free [6–10].

The goals of the present study were (1) to shed light on the natural course of PTC associated with detectable stimulated Tg levels in the absence of localizable disease as the best response at the initial post-treatment assessment and (2) to assess the value of repeated assay of TSH-stimulated Tg levels in the follow-up of these patients.

Patients and methods

Study design and setting

This retrospective, observational study was conducted with institutional review board approval in 8 hospital-based primary thyroid cancer referral centers in Italy (2 nuclear medicine units, 6 thyroid disease units). Since 1990, these centers have used similar protocols for the postoperative follow-up of patients with PTC. For those who undergo total or near-total thyroidectomy with RRA (the focus of the present study), the first follow-up visit is scheduled 6–18 months after the initial treatment. It includes assays of basal and TSH-stimulated serum Tg levels (with simultaneous assessment of TgAb levels) [22] and neck ultrasound studies performed with multifrequency probes (7.5–12 MHz) by an experienced staff endocrinologist. In the past, dxWBS was also performed routinely during the first visit in some centers, but since 2006 this procedure has been used only as a second-line study [23]. Subsequent visits are scheduled yearly and include basal Tg assay, TgAb levels, and neck ultrasound. If the stimulated Tg assay is positive at the initial visit, it is repeated periodically (frequency at the discretion of the physician) as long as levels remain in the detectable range. Additional functional and/or cross-sectional imaging studies are ordered as needed. If suspicious lymph nodes are found and are eligible for surgical removal, fine needle aspiration biopsy is performed for cytology, measurement of Tg levels in the needle washout fluid and, since 1997, high-sensitivity assay of washout fluid for Tg and TSH receptor mRNA [24, 25]. All Tg and TgAb assays are done in the center's own laboratory. Between 1991 and 2006, a number of Tg assays with functional sensitivities of 1 ng/mL were used in the participating centers. Between 2002 and 2005, all centers that enrolled patients began using immunometric serum Tg assays with a functional sensitivity of 0.1–0.2 ng/mL. As for TgAb, all assays used on study patients were immunometric or radioimmunometric with institutional cut-offs that varied from center to center.

Cohort selection

We reviewed the thyroid cancer databases of each participating center and identified patients consecutively diagnosed with PTC between January 1, 1990 and January 31, 2006 who met all the following eligibility criteria: (1) primary treatment consisting of total or near-total thyroidectomy followed by RRA; (2) surgical pathology report confirming diagnosis of PTC; (3) all diagnostic procedures

during follow-up carried out by the staff of the participating study center; (4) all serum Tg values measured with assays with a functional sensitivity of 1 ng/mL or less and accompanied by negative Tg antibody assay results; (5) complete findings for each year of follow-up (in accordance with the protocol described above). From this set of patients, we then selected those with (1) a stimulated serum Tg level of >1 ng/mL and negative clinical or imaging findings at the first postoperative visit and (2) at least one other stimulated Tg assay performed ≤10 years after the first. To focus on the natural history of these tumors during follow-up, we then excluded from our analysis any cohort members who had received further therapy after the primary treatment.

Data collection and end points

Data on cohort patients were collected from the medical records and stored in an electronic database. They included demographics, data on the primary treatment (surgery and RRA), tumor characteristics, documentation of structural disease during the follow-up period, and any additional treatments that were performed. The prespecified primary endpoint was the spontaneous decline of the stimulated Tg level to ≤1 ng/mL in the presence of persistently negative findings on neck ultrasound and on any other imaging studies that may have been performed (i.e., the criteria for classification as disease-free established by the American Thyroid Association) [4]. The secondary endpoints were assessed at the end of follow-up and included (1) the persistence of stimulated Tg levels >1 ng/mL and (2) the emergence of clinical/imaging evidence of disease.

Statistical analysis

Differences between categorical variables (expressed as numbers and percentages) were evaluated with the Chi-square test or the Fischer Exact test for small samples; differences between continuous variables (reported as medians and ranges) were assessed with the Mann–Whitney test. A P value <0.05 was considered statistically significant. Stat view 5.0.1 software (SAS Institute inc., Cary, NC) was used for all statistical analyses.

Results

As shown in Fig. 1, of the 365 patients who met the eligibility criteria, 93 (26 %) had TSH-stimulated serum Tg levels >1 ng/mL with no structural evidence of disease at
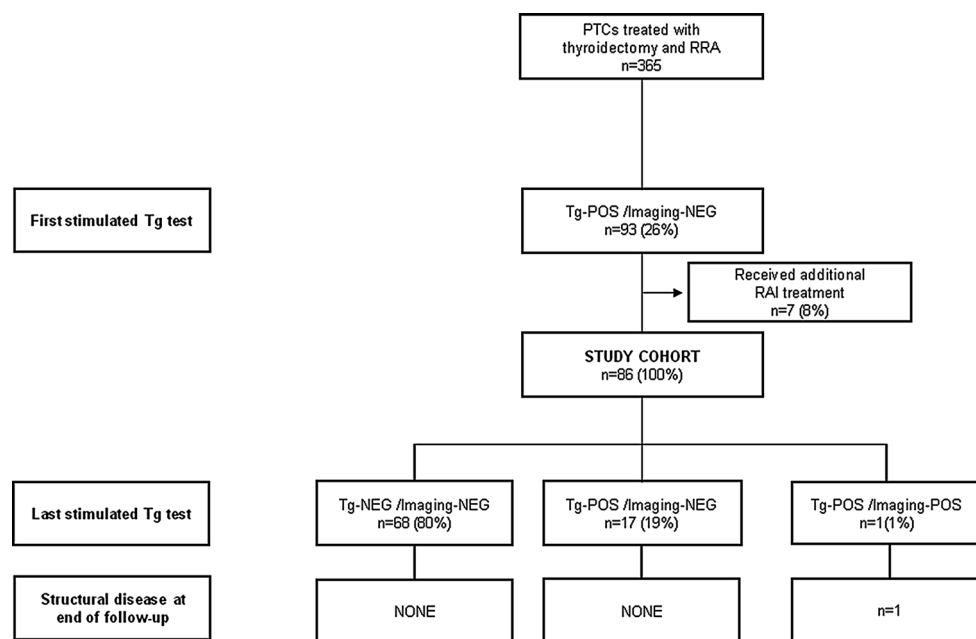

Fig. 1 Study cohort selection procedure and follow-up findings. A total of 365 cases of PTC satisfied the initial criteria for inclusion in this study. At the first postoperative visit [6–18 months after radioiodine remnant ablation (RRA)], 93 of these individuals were Tg-POS/Imaging-NEG. After exclusion of the 7 who received additional radioiodine therapy, the study cohort contained 86 patients. *Tg-NEG* stimulated serum Tg levels ≤1 ng/mL, *Tg-POS* stimulated serum Tg levels >1 ng/mL, *Imaging-NEG* no suspicious findings whatsoever on neck ultrasound or any other imaging studies that were performed, *Imaging-POS* suspicious/indeterminate findings on neck ultrasound. In all cases reported as Imaging-POS, the presence of structural disease was subsequently confirmed by biopsy and/or surgical pathology. The follow-up durations (from first to final post-treatment assessment) for this cohort ranged from 3.5 to 19.5 years (median 8.2 years)

Table 1 Baseline characteristics of the study cohort

Patient characteristics	Study cohort ($n = 86$)
Referral center—n (%)	
Center 1	42 (49)
Center 2	36 (42)
Center 3	6 (7)
Center 4	1 (1)
Center 5	1 (1)
Age (years)—median (range)	43.9 (15–87)
Females—no. (%)	66 (78)
Tumor size (mm)—median (range)	19 (2–60)
Multifocal tumors—no. (%)	35 (41)
Extrathyroidal extension—no. (%)	
None	61 (71)
Microscopic (T3)	24 (28)
Macroscopic (T4)	1 (1)
Lymph node metastases—no. (%)	19 (21)
Distant metastases—no. (%)	1 (1)
^{131}I ablation dose activity (mCi)—median (range)	66 (15–145)
ATA risk class—no. (%)	
Low	54 (63)
Intermediate	30 (35)
High	2 (2)
Initial postoperative findings (1 year)	
Serum Tg levels—median (range)	
Basal	0.71 (0.01–2.8)
Stimulated[a]	2.1 (1.05–27.9)
Post-treatment WBS[b]	
No uptake observed	56 (66)
Uptake confined to thyroid bed	29 (34)
Uptake outside the thyroid bed	–

Characteristics are those documented at the time of initial treatment (total thyroidectomy + radioiodine remnant ablation) with the exception of serum thyroglobulin levels, which are those measured at the first follow-up visit (6–18 months after initial treatment)

ATA American Thyroid Association, *mCi* millicurie, *Tg* thyroglobulin, *TSH* thyroid-stimulating hormone

[a] Assay performed according to standardized protocols (ATA 2009) after thyroid hormone withdrawal in 78 cases and after rhTSH injection in 8

[b] Data available for 85/86 patients

the first (1-year) postoperative assessment and at least one other stimulated Tg assay during follow-up. The 7 patients in this group who received additional radioiodine therapy were excluded from further analysis, leaving a study cohort of 86 patients being followed in 5 of the 8 centers. Their baseline characteristics are summarized in Table 1.

The final stimulated serum Tg assay in the study cohort was performed 1–10 years after the first (median: 5.1 years). TSH stimulation was achieved with LT4 withdrawal in 34 cases and the administration of rhTSH in the other 52 cases. In 42 (49 %) cases, the same stimulation method was used for the first and final Tg assays (LT4 withdrawal in 34 cases, rhTSH injection in 8) (see

Table 1). In the remaining 44 (51 %), the initial assay had been done after LT4 withdrawal and the final assay after rhTSH administration (see Table 1). At this point (Fig. 1), 68 (80 %) of the 86 patients presented stimulated (and basal) Tg levels that were ≤1 ng/mL together with persistently negative findings on physical examination, neck ultrasound, and any other imaging studies performed. In accordance with ATA Guidelines [4], these patients were thus reclassified as apparently disease-free, and monitoring of stimulated Tg levels was discontinued. In 17 other cases (19 %), the physical examination and neck ultrasound study were still negative, but the stimulated Tg levels remained above the 1 ng/mL cut-off (median 2.07 ng/mL,

Table 2 Baseline patient characteristics associated with negativization of TSH-stimulated Tg assay during follow-up

Patient characteristics[a]	Last TSH-Stimulated Tg assay result[b]		P value
	Tg-NEG (n = 68)	Tg-POS (n = 18)	
Initial postoperative TSH-stimulated Tg level—no. (%)			
>1 and <10 ng/mL	65 (95.6)	13 (72.2)	0.006
≥10 ng/ml	3 (4.4)	5 (27.8)	
ATA risk class—no. (%)			
Low	44 (64.7)	10 (55.6)	0.52
Intermediate	22 (32.3)	8 (44.4)	
High	2 (3)	0 (0)	
^{131}I ablation dose activity (mCi)—median (range)	67.3 (15–145)	63.3 (15–106)	0.9

Tg thyroglobulin, TSH thyroid-stimulating hormone, ATA American Thyroid Association, mCi millicurie

[a] See Table 1

[b] Tg-NEG (≤1 ng/mL); Tg-POS (>1 ng/mL)

Table 3 Sensitivity, specificity, and predictive values of stimulated Tg at the time of last stimulation test

	Sensitivity	Specificity	NPV	PPV
Stimulated Tg at last stimulation test	100 %	80 %	100 %	6 %

Tg thyroglobulin, NPV negative predictive value, PPV positive predictive value

range 1.02–4.7). (In all but 2 of the 17 (12 %), the basal Tg assay was negative (i.e., ≤1 ng/mL.) The mean interval between the first and last stimulated Tg assays in this group was not significantly different from that of the group with assay negativization (4.8 vs. 5.2 years, Mann–Whitney test, $P = 0.71$).

As for the final patient in the study cohort (Fig. 1), he had undergone total thyroidectomy, central neck dissection, and RRA (activity 106 mCi) at age 20 for an intermediate-risk (pT3, pN1a) PTC. The first postoperative assessment had revealed an rhTSH-stimulated Tg level of 1.05 ng/mL with negative findings on both neck ultrasound and dxWBS. Three years later, the rhTSH-stimulated Tg level was reassessed and found to be 1.24 ng/mL. (Basal Tg levels were undetectable at the time of both stimulation tests.) Suspicious lymph nodes were noted on neck US and cytologically confirmed to be PTC metastases. After discussion of the pros and cons of a second operation, the patient opted for active surveillance instead. In the 2.5 years that have passed since that time (5.6 years since the initial treatment), neck ultrasound studies performed at intervals of 6–12 months have revealed no change whatsoever in the size or appearance of lymph node lesions. At the final follow-up visit [median 8.2 years (range 3.5–19.5 years) after the first assessment, 9.6 years (range 4–21 years) after diagnosis], this patient was still the only one in the study cohort with documented evidence of structural recurrence.

As shown in Table 2, the likelihood of being reclassified as disease-free during follow-up in the 78 patients whose initial stimulated Tg levels were <10 ng/mL (i.e., those falling into the indeterminate range) was significantly greater than that of the 8 patients with higher Tg levels (i.e., biochemical incomplete responses) [26]. It was not significantly related to the initial American Thyroid Association (ATA) risk class or to the radioiodine dose used for remnant ablation. The sensitivity and predictive values of stimulated Tg in identifying patients destined to develop localizable disease are reported in Table 3.

Discussion

In this retrospective, multicenter study, we investigated the natural course of PTC in 86 patients whose initial assessment after primary treatment (total thyroidectomy + RRA) revealed TSH-stimulated serum Tg levels >1 ng/mL in the absence of localizable disease. This definition included patients whose biochemical response to therapy would currently be classified as *incomplete* (stimulated Tg levels ≥10 ng/mL) or *indeterminate* (stimulated levels >1 and <10 ng/mL) [26].

In 80 % of the cases, repetition of the stimulated Tg assay at some point during follow-up (median: 5.2 years after the initial assay) revealed levels below 1 ng/mL, while imaging findings remained persistently negative, thereby allowing these patients to be reclassified as disease-free, as defined by the 2009 ATA Guidelines [4]. In the remaining 20 %, the most recent stimulated serum Tg level recorded (median 4.8 years after the first assay) was still above 1 ng/mL,

although only 1 of these patients had developed structural evidence of disease. These outcomes were unchanged at the final follow-up visit (median 8.2 years after the initial assessment, range: 3.5–19.5 years): no evidence of structural disease had emerged in any of the patients reclassified as disease-free or in the 17 whose stimulated Tg levels remained >1 ng/mL in the absence of localizable disease.

Several other retrospective studies have explored the significance of detectable Tg production in the absence of localizable disease. Reported outcomes in PTC patient populations with these characteristics vary widely [6–10] depending on the cut-off used to define detectable Tg production and the specific features of the study cohort (including tumor characteristics, initial risk of recurrence, the use of additional radioiodine treatment, duration of follow-up). Yim et al. [9], for example, retrospectively analyzed 186 cases of differentiated thyroid cancer (DTC) with stimulated Tg levels ≥2 ng/mL roughly 1 year after thyroidectomy and RRA. Repeat stimulated Tg assays during follow-up (i.e., 1–2 years after the first assay) allowed 45 (24 %) of these patients to be reclassified as disease-free. A similar picture emerged from the retrospective study by Pitoia et al. [10]. They examined a smaller cohort of 32 patients whose only anomaly at the first follow-up visit was a stimulated Tg level of >1 ng/mL. Six (19 %) were subsequently reclassified as disease-free on the basis of negative results in the repeat stimulated Tg assay. In all 6 cases, the initial stimulated Tg level had been ≤1.9 ng/mL, and assay negativization occurred within 2.5 years.

These figures are in stark contrast with the 80 % rate of patients who became disease-free during our follow-up. The results will obviously be influenced by several factors including: the type ad completeness of primary treatment, the use of second-line imaging studies, and the Tg threshold used for study enrolment. In fact, the threshold used by Yim et al. [9] was twice as high as the one we chose. Compared with the patients in our cohort, many of those included in their study probably had more markedly abnormal stimulated Tg, which is also predictive of structural recurrence [26]. However, the high disease-free rates that emerged during our follow-up are probably related mainly to the clinical profiles of our cohort. Almost two-thirds of our patients were initially classified as low risk, as opposed to less than 40 % of the cohort examined by Pitoia et al. [10]. In addition, the proportion of their cases that were considered high risk was 7 times that of our cohort (15 vs. 2 %). Similar considerations apply to the patients analyzed by Yim et al. [9], 155 (83 %) of whom had lymph node metastases at diagnosis [vs. 19 (21 %) in our study].

Although retrospective studies like these are bound to vary in terms of cohort characteristics, it is important to note that the risk category distribution of the patients in our cohort (Table 1) is much more consistent with those reported in large unselected series of DTC patients [3, 27, 28]. It is also compatible with figures from the SEER Program database, which show that almost 68 % of thyroid cancers are confined to the thyroid gland at diagnosis (and are therefore considered at low risk for recurrence according to the ATA risk stratification system), while 26 and 4 % spread to regional lymph nodes and distant sites, respectively [29].

This distinction is important because the utility of repeating stimulated serum Tg assays in patients with suspicious findings that are exclusively biochemical is related mainly to its ability to identify patients who have been cured rather than those with persistent or recurrent cancer. Indeed, while the negative predictive value of TSH-stimulated Tg assay approaches 100 %, the positive predictive value is dismally low (26 %) [30]. In our study, the negative predictive value of stimulated Tg was 100 % while the positive predictive value was only 6 %. Patients who meet the criteria for classification as disease-free/excellent response have a minimal risk of recurrence, estimated at 1–2 % [31, 32]. They can thus be followed with more relaxed protocols, which offer important potential advantages in terms of patient stress and health-care spending. While the results of the studies cited above indicate that only a minority of patients with exclusively biochemical abnormalities stand to benefit by this downgrading of risk, our experience indicates that in a more or less typical cohort of patients with PTC, in which low-risk predominates, repetition of the stimulated Tg assay can markedly reduce the number of patients who actually require close, more costly follow-ups.

In our cohort, negativization of the stimulated Tg assay during follow-up displayed no significant association with the initial ATA risk class or the radioiodine dose used for remnant ablation. However, it was significantly more common in the subgroup whose initial stimulated Tg value was >1 ng/mL but <10 ng/mL (i.e., within the so-called indeterminate range) than that in the subgroup with initial values that met the criteria for an *incomplete biochemical response*. This finding is consistent with the results reported by Pitoia et al. [10]. In their study, all cases of assay negativization occurred in patients whose initial levels were between 1.5 and 1.9 ng/mL. Yim et al. also observed that rates of negativization declined as the initial stimulated Tg level rose [9].

As in the study by Yim, to observe the natural history of patients with detectable Tg levels but no imaging evidence of disease, we chose to exclude the 7 patients that received additional courses of radioiodine treatment. As for the outcome of these 7 patients excluded from the analysis, only one developed evidence of disease on imaging during

follow-up, while 3 patients were finally classified as free from disease.

The main limitations of our study are related to its retrospective, multicenter design. Because the cohort was composed of patients diagnosed over a period of 16 years in 5 different centers, baseline stimulated Tg levels were measured with several different assays. The fact that all assays used in this cohort had functional sensitivities of 1 ng/mL or less (inclusion criterion) allowed us to *retrospectively* dichotomize the results as detectable/undetectable using the evidence-based cut-off value recommended by the ATA [4]. The use of a categorical rather than continuous variable to express Tg levels precluded quantitative analysis of the changes in stimulated Tg over time, but our findings clearly confirm the validity of this cut-off for predicting recurrence. The same cut-off was used regardless of whether Tg was assayed after LT4 withdrawal or rhTSH injection, a practice that is fully consistent with recommendations of the ATA [4] and widely used in clinical settings, where the administration of exogenous TSH is gradually replacing thyroid hormone withdrawal. Several groups, however, have shown that Tg levels measured during hormone withdrawal are significantly higher than those observed under rhTSH stimulation [33, 34]. These findings suggest that in the 44 patients in our cohort whose initial and final Tg assays were performed after LT4 withdrawal and rhTSH stimulation, respectively, some cases of assay normalization might be related to the lower intensity TSH stimulation used in the second assay. False negativity of this type is in our opinion unlikely. First of all, while Tg measured after LT4-withdrawal consistently exceeds that observed in the same patient under rhTSH stimulation, this difference rarely (7 % of cases) results in discordant classification of the two specimens as positive/negative using the 1 ng/mL cut-off [33]. Second, in our cohort, the stimulated Tg negativation rate in this subgroup whose assays were performed under different types of stimulation (36/44, 82 %) is almost identical to that observed in the cohort as a whole (68/86, 80 %). For these reasons, the high proportion of patients reclassified as disease-free in our study is probably a reliable estimate.

The median interval between the first and final stimulated Tg assays in our cohort was 5.1 years, and the patients identified as disease-free at this point remained persistently negative in imaging studies during the remainder of follow-up (median 9.6 years). It is impossible to draw any definite conclusions on the optimal timing of repeat stimulated Tg assays on the basis of our findings. Nonetheless, from a cost–benefit viewpoint, the 5-year follow-up visit appears to be a reasonable choice. First, our previous experience shows that the vast majority of disease recurrences are detected within this time frame [3]. However, as noted earlier, the goal of re-testing is not to detect recurrence but to identify patients who can be considered disease-free (i.e., patients whose previously positive assay results probably reflected the presence of normal thyroid tissue remnants or even small nests of neoplastic cells that are still viable but destined to succumb with time to the effects of the initial treatment). In this context, it is interesting to note that in populations not subjected to RRA, basal Tg production declines spontaneously, and by year 5, 80 % of these patients have undetectable levels [5, 35]. In general, then, earlier re-testing is likely to be associated with lower rates of assay normalization. The intervals between the first and final stimulated Tg assays in our cohort varied widely, but it is important to note that the median length was not significantly different between patients with Tg negativization and patients with Tg >1 ng/mL at last assessment. In all probability, the increasing use of more sensitive Tg assays will eventually eliminate the need to assess Tg production under TSH stimulation, and the cost of assessing serum Tg to identify patients who are disease-free will be substantially reduced. However, a validated cut-off level that optimally identifies these patients is still not available for these new assays [30, 36–38].

In conclusion, up to one-third of PTC patients treated with thyroidectomy and RRA will have detectable stimulated Tg levels with no structural evidence of disease at 1 year. Our findings show that, in cases that fall within this "diagnostic gray zone," continued monitoring of stimulated Tg levels should be considered a key component of follow-up, *not* because of its capacity to identify patients with persistent/recurrent disease, which is in fact limited, but because it allows the *confident exclusion of this outcome* in a substantial proportion of this population.

References

1. L. Davies, H.G. Welch, Current thyroid cancer trends in the United States. JAMA Otolaryngol Head Neck Surg. **140**, 317–322 (2014)
2. E.L. Mazzaferri, S.M. Jhiang, Long-term impact of initial surgical and medical therapy on papillary and follicular thyroid cancer. Am. J. Med. **97**, 418–428 (1994)

3. C. Durante, T. Montesano, M. Torlontano, M. Attard, F. Monzani, S. Tumino, G. Costante, D. Meringolo, R. Bruno, F. Trulli, M. Massa, A. Maniglia, R. D'Apollo, L. Giacomelli, G. Ronga, S. Filetti, PTC study group: papillary thyroid cancer: time course of recurrences during postsurgery surveillance. J. Clin. Endocrinol. Metab. **98**, 636–642 (2013)

4. D.S. Cooper, G.M. Doherty, B.R. Haugen, R.T. Kloos, S.L. Lee, S.J. Mandel, E.L. Mazzaferri, B. McIver, F. Pacini, M. Schlumberger, S.I. Sherman, D.L. Steward, R.M. Tuttle, 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 **19**, 1167–1214 (2009)

5. C. Durante, T. Montesano, M. Attard, M. Torlontano, F. Monzani, G. Costante, D. Meringolo, M. Ferdeghini, S. Tumino, L. Lamartina, A. Paciaroni, M. Massa, L. Giacomelli, G. Ronga, S. Filetti, PTC study group: long-term surveillance of papillary thyroid cancer patients who do not undergo postoperative radioiodine remnant ablation: is there a role for serum thyroglobulin measurement? J. Clin. Endocrinol. Metab. **97**, 2748–2753 (2012)

6. R.T. Kloos, Thyroid cancer recurrence in patients clinically free of disease with undetectable or very low serum thyroglobulin values. J. Clin. Endocrinol. Metab. **95**, 5241–5248 (2010)

7. F. Vaisman, H. Tala, R. Grewal, R.M. Tuttle, In differentiated thyroid cancer, an incomplete structural response to therapy is associated with significantly worse clinical outcomes than only an incomplete thyroglobulin response. Thyroid **21**, 1317–1322 (2011)

8. F. Vaisman, D. Momesso, D.A. Bulzico, C.H. Pessoa, F. Dias, R. Corbo, M. Vaisman, R.M. Tuttle, Spontaneous remission in thyroid cancer patients after biochemical incomplete response to initial therapy. Clin. Endocrinol. (Oxf) **77**, 132–138 (2012)

9. J.H. Yim, E.Y. Kim, W. Bae Kim, W.G. Kim, T.Y. Kim, J.S. Ryu, G. Gong, S.J. Hong, J.H. Yoon, Y.K. Shong, Long-term consequence of elevated thyroglobulin in differentiated thyroid cancer. Thyroid **23**, 58–63 (2013)

10. F. Pitoia, E. Abelleira, H. Tala, F. Bueno, C. Urciuoli, G. Cross, Biochemical persistence in thyroid cancer: is there anything to worry about? Endocrine **46**, 532–537 (2014)

11. G. Brent, R. Larsen, Normalization of TSH may require up to 8 weeks, but complete recovery usually requires several months, in *Werner & Ingbar the Thyroid: A Fundamental and Clinical Text*, eds. by L. Braverman, R. Utiger R, 8th ed. (Lippincott Williams and Wilkins Publishers, Philadelphia), p. 853–858

12. F. Pacini, P.W. Ladenson, M. Schlumberger, A. Driedger, M. Luster, R.T. Kloos, S. Sherman, B. Haugen, C. Corone, E. Molinaro, R. Elisei, C. Ceccarelli, A. Pinchera, R.L. Wahl, S. Leboulleux, M. Ricard, J. Yoo, N.L. Busaidy, E. Delpassand, H. Hanscheid, R. Felbinger, M. Lassmann, C. Reiners, Radioiodine ablation of thyroid remnants after preparation with recombinant human thyrotropin in differentiated thyroid carcinoma: results of an international, randomized, controlled study. J. Clin. Endocrinol. Metab. **91**, 926–932 (2006)

13. U. Crocetti, C. Durante, M. Attard, A. Maniglia, S. Tumino, R. Bruno, N. Bonfitto, F. Dicembrino, A. Varraso, D. Meringolo, S. Filetti, V. Trischitta, M. Torlontano, Predictive value of recombinant human TSH stimulation and neck ultrasonography in differentiated thyroid cancer patients. Thyroid **18**, 1049–1053 (2008)

14. M.G. Castagna, L. Brilli, T. Pilli, A. Montanaro, C. Cipri, C. Fioravanti, F. Sestini, M. Capezzone, F. Pacini, Limited value of repeat recombinant human thyrotropin (rhTSH)-stimulated thyroglobulin testing in differentiated thyroid carcinoma patients with previous negative rhTSH-stimulated thyroglobulin and undetectable basal serum thyroglobulin levels. J. Clin. Endocrinol. Metab. **93**, 76–81 (2008)

15. J. Klubo-Gwiezdzinska, K.D. Burman, D. Van Nostrand, L. Wartofsky, Does an undetectable rhTSH-stimulated Tg level 12 months after initial treatment of thyroid cancer indicate remission? Clin. Endocrinol. (Oxf) **74**, 111–117 (2011)

16. P.W. Rosario, M.S. Furtado, A.F. Mineiro Filho, R.X. Lacerda, M.R. Calsolari, Value of repeat stimulated thyroglobulin testing in patients with differentiated thyroid carcinoma considered to be free of disease in the first year after ablation. Thyroid **22**, 482–486 (2012)

17. M. Torlontano, U. Crocetti, L. D'Aloiso, N. Bonfitto, A. Di Giorgio, S. Modoni, G. Valle, V. Frusciante, M. Bisceglia, S. Filetti, M. Schlumberger, V. Trischitta, Serum thyroglobulin and 131I whole body scan after recombinant human TSH stimulation in the follow-up of low-risk patients with differentiated thyroid cancer. Eur. J. Endocrinol. **148**, 19–24 (2003)

18. F. Pacini, E. Molinaro, M.G. Castagna, L. Agate, R. Elisei, C. Ceccarelli, F. Lippi, D. Taddei, L. Grasso, A. Pinchera, Recombinant human thyrotropin-stimulated serum thyroglobulin combined with neck ultrasonography has the highest sensitivity in monitoring differentiated thyroid carcinoma. J. Clin. Endocrinol. Metab. **88**, 3668–3673 (2003)

19. M. Torlontano, M. Attard, U. Crocetti, S. Tumino, R. Bruno, G. Costante, G. D'Azzò, D. Meringolo, E. Ferretti, R. Sacco, F. Arturi, S. Filetti, Follow-up of low risk patients with papillary thyroid cancer: role of neck ultrasonography in detecting lymph node metastases. J. Clin. Endocrinol. Metab. **89**, 3402–3407 (2004)

20. S. Orlov, F. Salari, L. Kashat, J.L. Freeman, A. Vescan, I.J. Witterick, P.G. Walfish, Post-operative stimulated thyroglobulin and neck ultrasound as personalized criteria for risk stratification and radioactive iodine selection in low- and intermediate-risk papillary thyroid cancer. Endocrine **50**, 130–137 (2015)

21. P.W. Rosario, T.L. Siman, M.R. Calsolari, Day 3 thyroglobulin ≤1 ng/ml after recombinant human TSH just prior to radioactive iodine is predictive of low risk for persistent/recurrent disease in patients with papillary thyroid carcinoma. Endocrine **49**, 170–174 (2015)

22. T.P. Loh, H.W. Chong, S.L. Kao, Thyroglobulin and thyroglobulin autoantibodies: interpret with care. Endocrine **46**, 360–361 (2014)

23. D.S. Cooper, G.M. Doherty, B.R. Haugen, R.T. Kloos, S.L. Lee, S.J. Mandel, E.L. Mazzaferri, B. McIver, S.I. Sherman, R.M. Tuttle, American thyroid association guidelines taskforce: management guidelines for patients with thyroid nodules and differentiated thyroid cancer. Thyroid **16**, 109–142 (2006)

24. F. Arturi, D. Russo, D. Giuffrida, A. Ippolito, N. Perrotti, R. Vigneri, S. Filetti, Early diagnosis by genetic analysis of differentiated thyroid cancer metastases in small lymph nodes. J. Clin. Endocrinol. Metab. **82**, 1638–1641 (1997)

25. D. Russo, F. Arturi, A. Pontecorvi, S. Filetti, Genetic analysis in fine-needle aspiration of the thyroid: a new tool for the clinic. Trends Endocrinol. Metab. **10**, 280–285 (1999)

26. B.R. Haugen, E.K. Alexander, K.C. Bible, G.M. Doherty, S.J. Mandel, Y.E. Nikiforov, F. Pacini, G.W. Randolph, A.M. Sawka, M. Schlumberger, K. Schuff, S.I. Sherman, J.A. Sosa, D.L. Steward, R.M. Tuttle, L. Wartofsky, American thyroid association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer. Thyroid (2015). doi:10.1089/thy.2015.0020

27. R. Elisei, E. Molinaro, L. Agate, V. Bottici, L. Masserini, C. Ceccarelli, F. Lippi, L. Grasso, F. Basolo, G. Bevilacqua, P. Miccoli, G. Di Coscio, P. Vitti, F. Pacini, A. Pinchera, Are the clinical and pathological features of differentiated thyroid carcinoma really changed over the last 35 years? Study on 4187

patients from a single Italian institution to answer this question. J. Clin. Endocrinol. Metab. **95**, 1516–1527 (2010)

28. F.A. Verburg, U. Mäder, K. Tanase, E.D. Thies, S. Diessl, A.K. Buck, M. Luster, C. Reiners, Life expectancy is reduced in differentiated thyroid cancer patients ≥45 years old with extensive local tumor invasion, lateral lymph node, or distant metastases at diagnosis and normal in all other DTC patients. J. Clin. Endocrinol. Metab. **98**, 172–180 (2013)

29. Surveillance Epiemiology and End Results Program. SEER 18 2005–2011, All races, Both Sexes by SEER Summary Stage 2000. www.seer.cancer.gov. Accessed 07 June 2015

30. M. Brassard, I. Borget, A. Edet-Sanson, A.L. Giraudet, O. Mundler, M. Toubeau, F. Bonichon, F. Borson-Chazot, L. Leenhardt, C. Schvartz, C. Dejax, I. Brenot-Rossi, M.E. Toubert, M. Torlontano, E. Benhamou, M. Schlumberger, THYRDIAG Working Group: long-term follow-up of patients with papillary and follicular thyroid cancer: a prospective study on 715 patients. J. Clin. Endocrinol. Metab. **96**, 1352–1359 (2011)

31. R.M. Tuttle, H. Tala, J. Shah, R. Leboeuf, R. Ghossein, M. Gonen, M. Brokhin, G. Omry, J.A. Fagin, A. Shaha, 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 (2010)

32. M.G. Castagna, F. Maino, C. Cipri, V. Belardini, A. Theodoropoulou, G. Cevenini, F. Pacini, 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 (2011)

33. F. Pacini, E. Molinaro, F. Lippi, M.G. Castagna, L. Agate, C. Ceccarelli, D. Taddei, R. Elisei, M. Capezzone, A. Pinchera, Prediction of disease status by recombinant human TSH-stimulated serum Tg in the postsurgical follow-up of differentiated thyroid carcinoma. J. Clin. Endocrinol. Metab. **86**, 5686–5690 (2001)

34. A. Kowalska, I. Pałyga, D. Gąsior-Perczak, A. Walczyk, T. Trybek, A. Słuszniak, R. Mężyk, S. Góźdź, The cut-off level of recombinant human TSH-stimulated thyroglobulin in the follow-up of patients with differentiated thyroid cancer. PLoS One **10**, e0133852 (2015)

35. T.E. Angell, C.A. Spencer, B.D. Rubino, J.T. Nicoloff, J.S. LoPresti, In search of an unstimulated thyroglobulin baseline value in low-risk papillary thyroid carcinoma patients not receiving radioactive iodine ablation. Thyroid **24**, 1127–1133 (2014)

36. L. Giovanella, M. Maffioli, L. Ceriani, D. De Palma, G. Spriano, Unstimulated high sensitive thyroglobulin measurement predicts outcome of differentiated thyroid carcinoma. Clin. Chem. Lab. Med. **47**, 1001–1004 (2009)

37. P. Malandrino, A. Latina, S. Marescalco, A. Spadaro, C. Regalbuto, R.A. Fulco, C. Scollo, R. Vigneri, G. Pellegriti, Risk-adapted management of differentiated thyroid cancer assessed by a sensitive measurement of basal serum thyroglobulin. J. Clin. Endocrinol. Metab. **96**, 1703–1709 (2011)

38. A.M. Chindris, N.N. Diehl, J.E. Crook, V. Fatourechi, R.C. Smallridge, Undetectable sensitive serum thyroglobulin (<0.1 ng/ml) in 163 patients with follicular cell-derived thyroid cancer: results of rhTSH stimulation and neck ultrasonography and long-term biochemical and clinical follow-up. J. Clin. Endocrinol. Metab. **97**, 2714–2723 (2012)