Comparison of Seven Serum Thyroglobulin Assays in the Follow-Up of Papillary and Follicular Thyroid Cancer Patients

M. Schlumberger, A. Hitzel, M. E. Toubert, C. Corone, F. Troalen, M. H. Schlageter, F. Claustrat, S. Koscielny, D. Taieb, M. Toubeau, F. Bonichon, F. Borson-Chazot, L. Leenhardt, C. Schvartz, C. Dejax, I. Brenot-Rossi, M. Torlontano, F. Tenenbaum, S. Bardet, F. Bussière, J. J. Girard, O. Morel, O. Schneegans, J. L. Schlienger, A. Prost, D. So, F. Archambeaud, M. Ricard, and E. Benhamou

Institut Gustave Roussy and University Paris Sud (M.S., F.Tr., S.K., M.R., E.B.), Villejuif, France; Centre Henri Becquerel (A.H.), Rouen, France; Hôpital Saint Louis (M.E.T., M.H.S.), Hôpital de la Pitié (L.L.), and Hôpital Cochin (F.Te.), AP-HP, Paris, France; Centre René Huguenin (C.C.), Saint Cloud, France; Hospices Civils de Lyon (F.C., F.B.-C.), Lyon, France; Hopital La Timone (D.T.), Marseille, France; Centre Georges François Leclerc (M.T.), Dijon, France; Institut Bergonié (F.Bo.), Bordeaux, France; Institut Jean Godinot (C.S.), Reims, France; Centre Jean Perrin (C.D.), Clermont Ferrand, France; Institut Paoli Calmettes (I.B.-R.), Marseille, France; Hôpital de San Giovanni Rotondo (M.T.), Italy; Centre François Baclesse (S.B.), Caen, France; Centre Lacassagne (F.Bu.), Nice, France; Hôpital Bretonneau (J.J.G.), Tours, France; Centre Paul Papin (O.M.), Angers, France; Centre Paul Strauss (O.S.), Strasbourg, France; Hôpital Hautepierre (J.L.S.), Strasbourg, France; Hôpital du Mans (A.P.), Le Mans, France; Hôpital de Thionville (D.S.), Thionville, France; and Hôpital de Limoges (F.A.), Limoges, France

Background: Serum thyroglobulin (Tg) is the marker of differentiated thyroid cancer after initial treatment and TSH stimulation increases its sensitivity for the diagnosis of recurrent disease.

Aim: The goal of the study is to compare the diagnostic values of seven methods for serum Tg measurement for detecting recurrent disease both during L-T_4 treatment and after TSH stimulation.

Methods: Thyroid cancer patients who had no evidence of persistent disease after initial treatment (total thyroidectomy and radioiodine ablation) were studied at 3 months on L-T_4 treatment (Tg1) and then at 9–12 months after withdrawal or recombinant human TSH stimulation (Tg2). Sera with anti-Tg antibodies or with an abnormal recovery test result were excluded from Tg analysis with the corresponding assay. The results of serum Tg determination were compared to the clinical status of the patient at the end of follow-up.

Results: Thirty recurrences were detected among 944 patients. A control ^{131}I total body scan had a low sensitivity, a low specificity, and a low clinical impact. Assuming a common cutoff for all Tg assays at 0.9 ng/ml, sensitivity ranged from 19–40% and 68–76% and specificity ranged from 92–97% and 81–91% for Tg 1 and Tg2, respectively. Using assays with a functional sensitivity at 0.2–0.3 ng/ml, sensitivity was 54–63% and specificity was 89% for Tg1. Using the two methods with a lowest functional sensitivity at 0.02 and 0.11 ng/ml resulted in a higher sensitivity for Tg1 (81% and 78%), but at the expense of a loss of specificity (42% and 63%); finally, for these two methods, using an optimized functional sensitivity according to receiver operating characteristic curves at 0.22 and 0.27 ng/ml resulted in a sensitivity at 65% and specificity at 85–87% for Tg1.

Conclusion: Using an assay with a lower functional sensitivity may give an earlier indication of the presence of Tg in the serum on L-T_4 treatment and may be used to study the trend in serum Tg without performing any TSH stimulation. Serum Tg determination obtained after TSH stimulation still permits a more reliable assessment of cure and patient's reassurance. (*J Clin Endocrinol Metab* 92: 2487–2495, 2007)

I N THYROID CANCER patients, the extent of neck recurrence or distant metastases has a major prognostic impact (1, 2). A cure is frequently obtained without significant morbidity in patients with a low tumor burden, but rarely in those with extensive disease, despite aggressive treatment modalities. Thus, highly sensitive tools should permit the early diagnosis of persistent and recurrent disease. Fortunately, recurrences are becoming less frequent due to earlier diagnosis of the disease and to more appropriate therapeutic

modalities. Diagnostic tools should also be specific to avoid unnecessary treatment in the great majority of patients with no evidence of disease.

Since the beginning of the eighties, follow-up strategies have combined the use of total body scan with ^{131}I radioiodine (^{131}I-TBS) and serum thyroglobulin (Tg) determination. Serum Tg is an excellent tumor marker with conventional assays with a functional sensitivity at 1.0 ng/ml. Serum Tg during L-T_4 treatment is higher than this cutoff in more than 95% of patients with distant metastases; however, it is lower than this cutoff in 20% of patients with isolated lymph node metastases (3, 4), and this proportion is likely to increase with the routine use of neck ultrasonography (US). Disease detection with serum Tg measurement is improved when it is performed after either withdrawal of L-T_4 or im injections of recombinant human TSH (rhTSH) (3–5). Under such condi-

tions, false negative serum Tg determinations are observed mostly in patients with isolated small lymph node metastases that can be evidenced on neck US (6–9). In addition, ^{131}I-TBS appeared less sensitive than neck US for the detection of neck recurrences, and uptake at distant sites was demonstrated with a diagnostic or a high activity only in patients with detectable serum Tg after TSH stimulation (7–11). This was the basis for two recent consensus reports, according to which, when there is no evidence of disease at 9–12 months after initial treatment with neck US and serum Tg determined after TSH stimulation, routine diagnostic ^{131}I-TBS can be avoided at least in low-risk patients (12, 13). When this early evaluation does not show any evidence of disease, the risk of subsequent recurrence is low (10, 11).

In these studies, the functional sensitivity of serum Tg assays was around 1 ng/ml, and the availability of kits with an improved functional sensitivity raised the possibility that measuring serum Tg during L-T$_4$ treatment might permit the early detection of persistent or recurrent disease without the need for any TSH stimulation (14–17). Indeed, the functional sensitivity of each assay is an important characteristic because it gives a measure of the reproducibility of a measurement. It is the concentration derived from the 20% between-run coefficient of variation that is determined in human serum over a period of at least 6 months, using at least two different lots of reagents and two different instrument calibrations during the testing period (17). It gives the value above which any concentration can be considered detectable in clinical practice.

The objective of the present study was to estimate the diagnostic accuracy of seven Tg assays in a large series of patients with papillary and follicular thyroid cancer who had no evidence of disease. Patients were prospectively studied according to the recommended follow-up protocol (12, 13). Serum Tg measurement was obtained at 3 months after ablation on L-T$_4$ treatment, and at 9–12 months at the time of a control ^{131}I-TBS, after withdrawal of either L-T$_4$ or rhTSH. Either complete remission or persistent/recurrent disease was assessed at the 9–12 month study and during subsequent follow-up.

Patients and Methods

Patients

Patients with papillary and follicular thyroid carcinoma were enrolled if they were treated in the following manner: 1) they underwent a total thyroidectomy and, depending on local protocols, central neck lymph node dissection resulting in apparent complete resection of neoplastic foci. 2) Postoperatively, they received ^{131}I for thyroid remnant ablation [1.1–3.7 GBq (30–100 mCi)], and no focus of uptake was detected outside the thyroid bed on the TBS performed 3–5 d after ^{131}I therapy. 3) They then received L-T$_4$ treatment, and were seen 3–4 months later so that the adequacy of the L-T$_4$ dose could be monitored.

Study protocol

Patients were enrolled in the study at the 3- to 4-month examination, after obtaining their informed consent. The study protocol consisted first in obtaining a serum Tg1 sample at 3–4 months after the initiation of L-T$_4$ treatment when serum TSH was less than 0.5 μU/ml. Patients with serum TSH more than 0.5 μU/ml had a 25-μg increase in the daily dose of L-T$_4$, and further serum TSH and Tg1 samples were obtained 3 months later that were used in the study. Second, at 9–12 months after the initiation of L-T$_4$ therapy and after TSH stimulation, a second serum

sample (Tg2) and a ^{131}I-TBS were obtained. At that time, TSH stimulation was achieved either by withdrawing L-T$_4$ treatment for 4–5 wk until serum TSH exceeded 30 μU/ml, or through injections of rhTSH (0.9 mg im for 2 consecutive days). The method chosen for TSH stimulation was at the local doctors' discretion. The Tg2 sample was obtained on the day of ^{131}I administration in case of withdrawal (the sample was also used for TSH measurement) or 1–3 d after the second injection in case of rhTSH stimulation.

Between June 2000 and October 2003, 968 patients from 27 centers were enrolled in the study. Data obtained for each patient included: age, gender, date and extent of surgery, histological type of the primary thyroid tumor, pathological tumor node metastasis classification (18), date of ^{131}I ablation and results of postablation TBS, and the dates of Tg1 and Tg2 sampling. Serum Tg1 and Tg2 samples were aliquoted and stored frozen in a central laboratory at −80 C until use. The method used for TSH stimulation, the serum TSH levels corresponding to Tg1 and Tg2, and the date of the ^{131}I-TBS were also recorded. At the time of the protocol initiation, neck US was not a routine procedure in many centers; however, it was routinely performed during the subsequent follow-up in the majority of patients.

Then, patients were followed up according to local procedures, and outcome was recorded annually. However, because annual follow-up of all thyroid cancer patients was not the policy in several participating centers, follow-up data were not available for a number of patients. At each evaluation, the results of each test were recorded and patients were classified as without or with evidence of disease. In the latter case, the site of recurrence was recorded. When results were doubtful, for instance ^{131}I uptake in the neck that seemed to be located outside the thyroid bed, every attempt was made during the subsequent follow-up to classify patients as free of disease or exhibiting persistent or recurrent disease.

The protocol was approved by the Ethics Committee at Bicêtre Hospital, France and by the local Clinical Research committee of each participating center. All patients gave their written informed consent before participating in the study.

Control ^{131}I-TBS at 9–12 months

TBS was performed with a standard protocol, 2 d after the administration of 148–185 MBq (4–5 mCi) [131]I and consisted in a TBS (scanning with a double-head γ-camera at low speed for a minimum of 30 min or at least 140,000 counts) with spot images (scanning for a minimum of 10 min or for at least 60,000 counts) in any doubtful region. Pregnancy was excluded before each administration of radioiodine. Iodine contamination was prevented by careful instructions and was sought by interviewing the patient and, when doubt persisted, by urinary iodine measurement.

Ten sessions were organized during which all ^{131}I-TBS images were blindly reviewed by a panel of at least 10 experts, endocrinologists or nuclear medicine physicians, participating in the study. A consensus was obtained for each TBS and was recorded. Foci of uptake outside the thyroid bed in the neck and at distant sites were recorded. Foci of uptake in the neck were classified as suspicious when they were located near the thyroid bed or the salivary glands and confirmation was sought in all patients with other imaging modalities, including neck US and by fine needle aspiration biopsy when abnormalities were observed at neck US. Similarly, distant sites with uptake were submitted to other imaging procedures, including computed tomography scan of the chest, computed tomography scan or magnetic resonance imaging for bone foci, and fluorodeoxyglucose positron emission tomography scan.

Assays

Once the study was completed, aliquots of each sample of Tg1 and Tg2 were sent to one of the three participating laboratories (Lyon, Saint-Louis, and Institut Gustave Roussy). Serum Tg measurement was performed in duplicate for all samples with each of the seven kits. A given kit was used to measure all samples in a given laboratory according to the manufacturer's instructions. Results were reported as the average value of the duplicates.

The seven kits used in the study were Tg-Kryptor (B.R.A.H.M.S., Berlin, Germany), Immulite TG (Diagnostic Products Corporation, Los Angeles, CA), Thyro (Cis Bio International, Gif-sur-Yvette, France), Tg

Advantage (Nichols Institute Diagnostics, San Clemente, CA), DYNOtest Tg-Plus (B.R.A.H.M.S.), Tg Access (Beckman-Coulter, Fullerton, CA), and e-Iason TgCa (Iason, Graz-Seierberg, Austria). Functional sensitivity, described by the manufacturers is 0.9 ng/ml for three assays (Tg Kryptor, Immulite Tg, and Thyro), 0.3 ng/ml for Tg Advantage, 0.2 ng/ml for Dynotest Tg-Plus, 0.11 ng/ml for Tg Access, and 0.02 ng/ml for e-Iason TgCa.

Calibration of each kit was controlled by using an international Tg standard (single lot number, CRM 457; BCR, Brussels, Belgium) [19] at eight known Tg values spiked into human Tg free sera obtained during L-T_4 treatment from a patient who had no evidence of disease for 10 yr and undetectable serum Tg after L-T_4 withdrawal with the Tg Access kit, and by measuring these samples with all assays [20]. Results in Table 1 clearly demonstrate the method-to-method variability. In fact, six of the seven manufacturers claimed CRM-457 standardization, but apart from Immulite TG and Tg Access methods, the other four methods (Tg-Kryptor, Dynotest Tg-Plus, Tg Advantage, and e-Iason TgCa) involve an adjustment factor of around 2 for corresponding to the expected values.

A direct search for anti-Tg antibodies (Immulite TG Ab, TgAb Nichols Advantage, TgAb Access) or the recovery test (Tg-Kryptor, Thyro, DYNOtest Tg-Plus, e-Iason TgCa) was performed according to the manufacturer's instructions for each method. Anti-Tg antibody methods are standardized against the World Health Organization 1st International Reference Preparation 65/93. Values below or equal to 40 IU/ml, 2 IU/ml, or 5 IU/ml, respectively, were regarded as negative. According to the method used, anti-Tg antibodies were found in 126–161 Tg1 samples and in 85–119 Tg2 samples. The expected values for the recovery tests were between 70–130% of recovery, and recovery was incorrect (<70% or >130%) in only 2–14 Tg1 and Tg2 samples with three kits, and in as many as 83 Tg1 samples and 69 Tg2 samples with the e-Iason TgCa assay. Sera with anti-Tg antibodies or with an abnormal recovery test result were excluded from Tg analysis with the corresponding method.

Tests and standards

For Tg1, the results were not available for 2–9% of patients, including for 14 patients for whom no Tg1 sample was available for any assay. For Tg2, the results were not available for 7–15% of patients, including for 60 (5%) patients for whom no Tg2 sample was available for any assay.

The results of serum Tg determinations were reported blindly from the clinical status of the patients. The reference standard used in the study was a combination of the ^{131}I-TBS at 9–12 months and the subsequent clinical and imaging follow-up. The results of the blinded TBS review were taken into account. Neck recurrence was confirmed by a fine needle biopsy or surgical biopsy, and ^{131}I uptake or typical imaging features confirmed distant metastases. Follow-up of patients consisted in a clinical examination with serum Tg determination under L-T_4 treatment on a yearly basis in all patients, and according to local practice, in neck US in most patients, and in a control diagnostic ^{131}I-TBS and serum Tg determination after TSH stimulation in a few patients. Serum Tg was measured in each center with the locally used method. Data on follow-up are still being collected annually. According to the protocol, patients who were not seen on a yearly basis will be evaluated again in the center at 5 yr after initial treatment.

Statistical analysis

Sample size. While designing the study, it was estimated that a total of 200 patients with a positive TBS would be required to estimate a sensitivity approximating 85% with 5% accuracy for a Tg assay. The probability of recurrence was estimated from historical series at around 20%, and 1000 patients were included in the study [3, 21].

Diagnostic accuracy. We used classical criteria to assess the clinical performance of each assay: 1) the sensitivity is the percentage of patients with recurrent disease during follow-up who had a positive Tg determination; 2) the specificity is the percentage of patients who had no evidence of disease during follow-up and who had a negative Tg determination.

Sensitivity and specificity were calculated for each assay using the aforementioned reference standard. These calculations were performed with three cutoff levels (positive or negative) for both Tg1 and Tg2: 1) a single common cutoff equal to 0.9 ng/ml for all kits; 2) for the four assays with a lower functional sensitivity, the functional sensitivity reported by the manufacturers; and 3) for the two most sensitive assays, the cutoff maximizing the sum of sensitivity plus specificity deduced from the receiver operating characteristic (ROC) (Fig. 1).

Results

Patient characteristics

A total of 968 patients from 27 centers were enrolled in the study (Table 2). Twenty-four patients did not complete the study. Among the 944 patients who completed the study and who were followed up, 504 received rhTSH injections and, in 440, L-T_4 was withdrawn for Tg2 determination and ^{131}I-TBS. Patient clinical characteristics are reported in Table 2 according to the TSH stimulation method (rhTSH or withdrawal of L-T_4). The percentage of women (79% and 76%), mean age at diagnosis (47 yr), and the pathological tumor classification and the histological type (88% and 87% of papillary carcinomas) were similar between the two groups; a slightly greater number of pN1 patients were in the rhTSH group (30%) than in the withdrawal group (24%) ($P = 0.06$). In accordance with inclusion criteria, no patient had known local and regional tumor foci or distant metastases at the time of inclusion.

Reference standard determination for each patient

As described above, the reference standard used was a combination of the control ^{131}I-TBS performed 9–12 months after ablation and clinical and imaging data obtained over a mean follow-up of 28 months (range: 8–50 months).

Recurrent disease was evidenced in 30 patients, in the thyroid bed (n = 5); in neck lymph nodes (n = 16); in lungs (n = 4); in bones (n = 1); in neck lymph nodes and lungs (n = 1); in bones and lungs (n = 1); in neck lymph nodes, lungs, and bones (n = 1); and in neck lymph nodes and skin (n = 1) (Table 3).

TABLE 1. Tg standard CRM 457 was diluted in human Tg free serum and then measured with each method

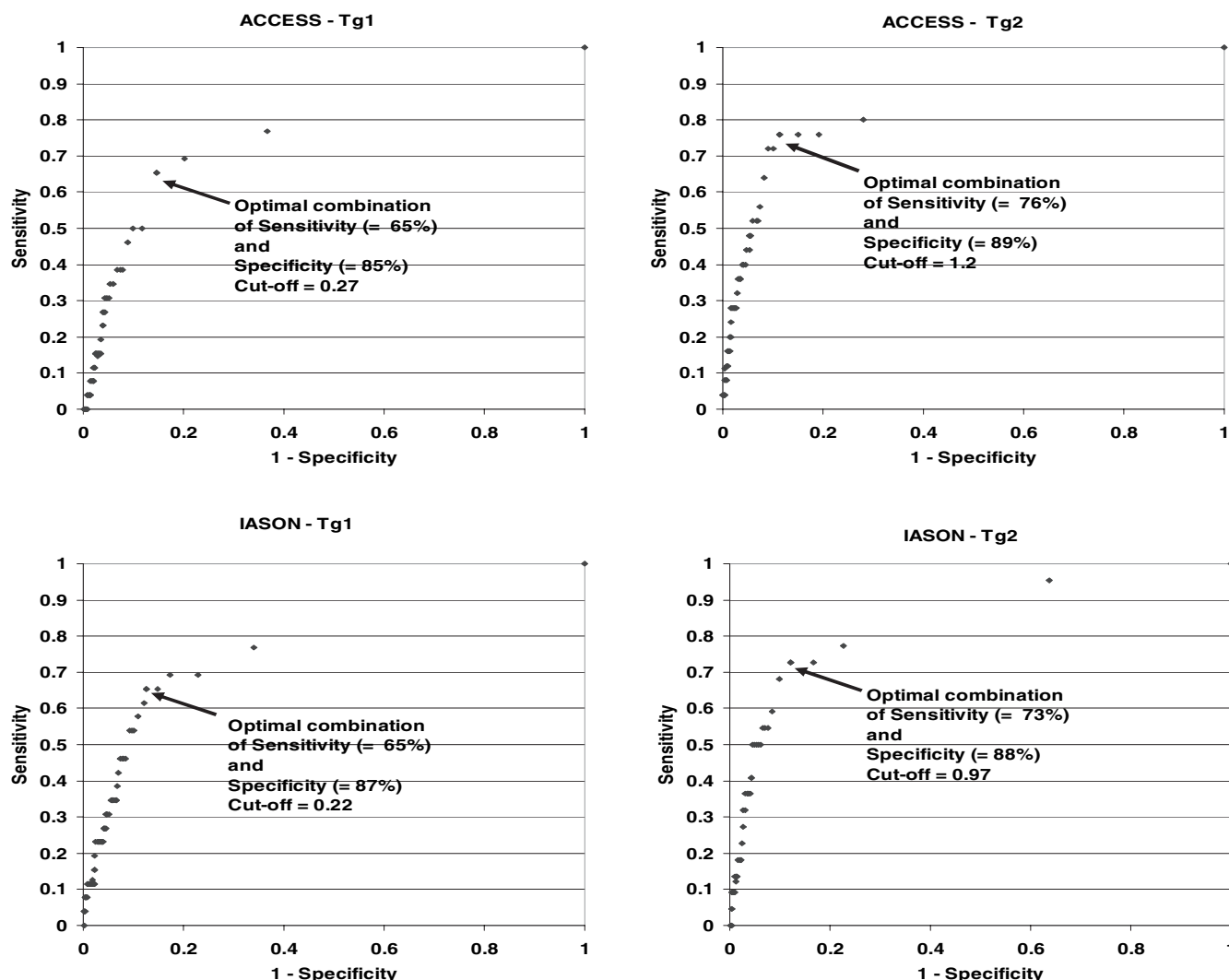| Serial dilutions of CRM 457 (ng/ml)	Calculated values (ng/ml)						
	Tg-Kryptor	Immulite TG	Thyro	Tg Access	Tg Advantage	DYNOtest	e-Iason TgCa
200	95	188	245	185	112	100	126
100	52	81	125	95	55	50	58
20	9.6	14.6	24.1	17	13.2	10	10.5
5	2.9	3.8	7.6	4.4	3.7	2.8	2
1	0.8	0.8	1.8	1	1.2	0.7	0.4
0.5	0.2	<0.2	1.4	0.5	0.5	0.2	0.2
0.25	<0.17	<0.2	0.91	0.33	<0.3	<0.1	0.13
0.10	<0.17	<0.2	<0.2	0.20	<0.3	<0.1	0.07

Fig. 1. ROC curves used to determine the Tg level that maximizes the sum sensitivity plus specificity for the Tg Access and for the e-Iason TgCa assays.

Results of the control ^{131}I-TBS

The ^{131}I-TBS performed 9 months after ablation did not show any significant uptake in the thyroid bed in 91% of the patients after rhTSH stimulation and in 81% of the patients after withdrawal (Table 4). It showed visible but low (<0.1%) uptake in the thyroid bed in 8 and 18% of the patients after rhTSH and withdrawal, respectively, and uptake in the thyroid bed greater than 0.1% of the administered activity in 1% of the patients of each group (Table 4).

The control TBS revealed suspicious foci of uptake in 23 patients. These foci of uptake corresponded to neoplastic foci in only seven patients (in two whose control ^{131}I-TBS was performed after rhTSH and in five patients whose control ^{131}I-TBS was performed after withdrawal). Uptake was visible in only three of the 16 patients with lymph node metastases, in two of the five patients with a recurrence in the thyroid bed, and in two of the nine patients with distant metastases. In the other 16 patients, suspicious foci were located in the neck (n = 15 patients) and were finally considered as thyroid remnants or salivary glands and in lungs

(n = 1) and were considered as artifacts due to a skin contamination (Table 4). These patients were then followed up for 12–50 months, and no recurrence was observed.

Based on these findings, the results of the control ^{131}I-TBS when compared with the patient's clinical status at the end of the follow-up can be classified as true positive in seven, false negative in 23, false positive in 16, and true negative in 898 patients, leading to 23% sensitivity, 98% specificity, with a clinical impact in only seven (0.7%) among 944 patients.

Serum Tg determination

Serum Tg was obtained 3 d after the second rhTSH injection in 427 patients and 1 or 2 d after the second injection in 77 patients. After withdrawal, mean serum TSH was 87 μU/ml (sd=49).

Results of serum Tg determinations (Tg1 and Tg2)

Using a common cutoff at 0.9 ng/ml, the proportion of patients with detectable Tg1 was similar with each of the

TABLE 2. Initial clinical characteristics of the patients and number of recurrences according to the pathological tumor node metastasis stage (1992 classification) and histology

	rhTSH (n = 504)	Withdrawal (n = 440)	P^a	Recurrent disease (n)^b
Women (%)	79	76	0.23	
Age mean (SD)	47 (14)	47 (13)	0.32	
TNM classification (1998)				
T<1 cm	124 (25%)	110 (25%)	0.08	6
T 1–4 cm	268 (53%)	238 (54%)		8
T>4 cm	34 (7%)	44 (10%)		3
Extrathyroid extension	78 (15%)	48 (11%)		13
N1	139 (30%)	105 (24%)	0.06	16
N0	324 (70%)	327 (76%)		14
Nx	41	8		
M0	504 (100%)	440 (100%)		
Histology				
Papillary	444 (88%)	381 (87%)	0.14	28
FMI	45 (9%)	52 (12%)		1
FWI	15 (3%)	7 (2%)		1

Pap, Papillary; FMI, follicular minimally invasive; FWI, follicular widely invasive.
^{a} *P* refers to differences between rhTSH and withdrawal groups.
^{b} Recurrent disease was more frequently observed in patients with N1 and extrathyroid extension ($P < 0.01$).

seven methods in the rhTSH group (3–14%) and in the withdrawal group (4–13%). In contrast, the proportion of detectable Tg2 was higher after withdrawal (16–33%) than after rhTSH (8–23%), with each of the seven methods ($P < 10^{-4}$). However, the number of recurrences was similar in each group (14 in the withdrawal group and 16 in the rhTSH group).

Tg2 was not significantly different among patients with either no detectable uptake in the thyroid bed, detectable but low uptake (<0.1%), or with uptake greater than 0.1% on the control ^{131}I-TBS (data not shown).

Diagnostic values of the seven methods (Table 5)

Tg1 and Tg2 results were compared with the clinical status of patients at the end of the study, and the sensitivity and specificity of each assay were calculated. Tg2 obtained either after withdrawal or rhTSH were pooled for this analysis.

Assuming a common cutoff at 0.9 ng/ml, results were

similar with all kits: during L-T$_4$ treatment, sensitivity ranged from 19–40% and specificity ranged from 92–97% for Tg1. After TSH stimulation (Tg2), sensitivity attained 68–76% and specificity remained high at 81–91%.

Sensitivity and specificity were then studied according to functional sensitivity values provided by the manufacturers. With a functional sensitivity at 0.3 and 0.2 ng/ml (Tg Advantage and Dynotest Tg-Plus kits), sensitivity of Tg1 improved to 63% and 54%, respectively, and specificity was 89% for both assays. With the lowest functional sensitivity (0.02 ng/ml for e-Iason TgCa and 0.11 ng/ml for Tg Access kit), the sensitivity was even higher for both Tg1 (81% and 78%) and for Tg2 (95% and 92%), but at the expense of a considerable loss in specificity for both Tg1 (42% and 63%) and for Tg2 (23% and 44%).

Optimized thresholds were then determined for these two kits, according to ROC curves (Fig. 1): for Tg Access kit, it was equal to 0.27 and 1.2 ng/ml for Tg1 and Tg2, respectively,

TABLE 3. Characteristics of the 30 recurrences detected at control ^{131}I -TBS and during the subsequent follow-up

	Interval between surgery and detection (months)	Histology of the primary tumor	TNM (1992 classification)
Thyroid bed (n = 5)	Median: 16; [10–30]	Pap: 4 FWI: 1	T1N0:2 T3N1:1 T4N1:2
Neck lymph nodes (n = 16)	Median: 24; [13–50]	Pap: 16	T1N0:3 T1N1:1 T2N0:1 T2NI:4 T3N1:1 T4N0:2 T4N1:4
Lung (n = 4)	Median 39; [16–42]	Pap: 3 FMI: 1	T2N0:1 T3N0:1 T4N1:2
Bone (n = 1)	8	Pap: 1	T4N1
Lymph nodes + lung (n = 1)	10	Pap: 1	T2N0
Lung + bone (n = 1)	12	Pap:1	T2N0
Lung + bone + lymph nodes (n = 1)	23	Pap:1	T4N0
Lymph nodes + skin (n = 1)	14	Pap:1	T4N0

Pap, Papillary; FMI, follicular minimally invasive; FWI, follicular widely invasive; TNM, tumor node metastasis.

TABLE 4. Results of ^{131}I total body scan performed at 9 months after withdrawal or rhTSH stimulation

	rhTSH	Withdrawal
Uptake in the thyroid bed at control ^{131}I-TBS	n = 494	n = 435
No visible uptake	452 (91%)	352 (81%)
Visible uptake		
<0.1%	37 (8%)	78 (18%)
>0.1%	5 (1%)	5 (1%)
Suspicious uptake at control ^{131}I-TBS	n = 9	n = 14
Not related to neoplastic disease	7 (6 in the neck and 1 in the thorax)	9 (9 in the neck)
Related to neoplastic disease	2 (1 in bones and 1 in lungs and lymph nodes)	5 (3 in lymph nodes and 2 in the thyroid bed)
Absence of uptake in patients who subsequently developed a clinical recurrence	14	9

Sixteen recurrences were observed in the rhTSH group, and 14 were observed in the withdrawal group.

and for e-Iason TgCa kit, the corresponding thresholds were 0.22 and 0.97 ng/ml, respectively. At these optimized thresholds for Tg1, the sensitivity was 65% for both assays, and specificity was 85% and 87%.

Using a kit with a lower functional sensitivity resulted in a lower clinical impact of TSH stimulation (Table 6).

With both e-Iason and Access kits, Tg2 was strongly related to Tg1 (Table 7): only 3% of patients with undetectable Tg1 had Tg2 greater than 0.9 ng/ml, including 1% with Tg2 greater than 2 ng/ml. These percentages increased with increasing Tg1. In contrast, two thirds of patients with undetectable Tg1 also had undetectable Tg2, and most of the other patients had low Tg2.

Discussion

This prospective study reflects routine practice with thyroid cancer patients, and its results can be readily applicable.

TABLE 5. Sensitivity and specificity of the seven assays according to a common cutoff at 0.9 ng/ml

	Functional sensitivity (ng/ml)	Sensitivity (%)	Specificity (%)
Tg1			
Kryptor	0.9	21	96
Immulite	0.9	40	96
Thyro	0.9	38	92
Advantage	0.9	35	94
	0.3	63	89
DYNOtest Tg Plus	0.9	19	97
	0.2	54	89
Access	0.9	35	95
	0.11	78	63
Iason	0.9	35	94
	0.02	81	42
Tg2			
Kryptor	0.9	68	91
Immulite	0.9	71	88
Thyro	0.9	76	81
Advantage	0.9	72	83
	0.3	91	67
DYNOtest Tg Plus	0.9	72	91
	0.2	81	81
Access	0.9	76	86
	0.11	92	44
Iason	0.9	73	87
	0.02	95	23

For the four methods with an improved functional sensitivity, sensitivity and specificity are also reported according to the functional sensitivity provided by the manufacturers.

It demonstrates the low frequency of recurrences in thyroid cancer patients after complete surgical excision and when no focus of uptake outside the thyroid bed is seen on the postablation TBS. This risk of recurrence is much lower than that reported in historical series (3, 21) and explains why the potential diagnostic benefits of any tool is limited. This low number of events did not prevent us from drawing several conclusions: the low sensitivity of serum Tg determination during thyroxine treatment was improved after either TSH stimulation or by using a measurement method with a lower functional sensitivity; a large number of detectable serum Tg determinations did not correspond to any detectable disease, with an even larger number when the functional sensitivity of the serum Tg assay was lower; wide discrepancies were observed between methods for screening anti-Tg antibodies, and between these methods and the recovery tests; routine diagnostic ^{131}I-TBS during follow-up afforded information in only few patients. Each of these findings deserves extensive discussion.

Indeed we have been faced with missing data, but their low numbers both for Tg1 and Tg2 are unlikely to change significantly the conclusions of this study.

Persistent or recurrent disease was demonstrated in 30 patients (<3%) with a mean follow-up of 28 months, with two thirds of recurrences occurring in the thyroid bed or in neck lymph nodes. Neck US was not performed on a routine basis at 9–12 months, because it was not routinely used at the time of the study initiation, but previous studies on similar patients have shown that the risk of subsequent recurrence is small (0.5% at 10 yr) when neck palpation and control diagnostic TBS are normal and serum Tg remains undetectable after TSH stimulation (10, 11). In the remaining nine patients, distant metastases could already be suspected in seven, with the increased Tg1 as shown with all kits.

The present study confirms that a single radioiodine ablation is successful in 99% of patients, as shown by no detectable uptake in the thyroid bed and when present, less than 0.1% of the administered activity on a control ^{131}I-TBS (7–11). When persistent uptake was present in the thyroid bed, it did not appear to be clinically relevant and was not related to the serum Tg level either during L-T$_4$ treatment or after TSH stimulation. TBS revealed suspicious foci of uptake in only seven of the 30 patients with persistent/recurrent disease, but demonstrated suspicious foci of uptake in other 16 patients, that were finally related to normal thyroid rem-

TABLE 6. Tg1 and Tg2 in the 30 patients with recurrent disease

Recurrence	KTG1	KTG2	IMTG1	IMTG2	THTG1	THTG2	ADTG1	ADTG2	DYTG1	DYTG2	ACTG1	ACTG2	IATG1	IATG2
TB	UND	UND	UND	UND	UND	UND	UND	UND	UND	UND	UND	0.15	UND	0.09
TB	UND	UND	UND	UND	UND	UND	UND	UND	UND	UND	UND	UND	UND	UND
TB	UND	1	UND	1.5	UND	2.2	0.4	1.6	0.3	1.2	0.32	1.4	0.25	1.1
TB	147	7.1	364	ND	264	ND	108.5	ND	287	8.3	268	ND	275	ND
TB	1.8	13	3.8	31	5.1	29	6	44	2.5	19	3.8	29	3.5	10
NLN	UND	1.3	UND	2.4	UND	4.7	UND	3.8	UND	2.2	UND	2.4	0.05	2.7
NLN	UND	4.9	UND	13	UND	12	1	18	UND	6.2	0.47	12	0.44	11
NLN	UND	1.4	UND	2.5	UND	2.8	UND	3.7	UND	1.5	0.28	2.2	0.15	1.5
NLN	UND	UND	UND	UND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
NLN	UND	2.1	UND	3.9	UND	4.6	UND	5.8	UND	2.3	0.13	3.4	0.09	4.3
NLN	UND	UND	UND	UND	UND	1.2	UND	0.7	UND	0.5	0.17	0.55	0.06	0.53
NLN	UND	1.1	UND	2.2	UND	3.1	0.7	3.2	UND	1.3	0.33	1.9	0.27	1.9
NLN	UND	UND	UND	UND	UND	UND	UND	UND	UND	UND	0.15	UND	UND	0.05
NLN	UND	2.7	UND	6.7	1.1	6.7	0.8	6.7	0.6	4.4	0.56	5.7	0.71	8
NLN	UND	6.6	UND	11	1.5	21	0.9	17	0.6	7.8	0.77	13	0.68	12
NLN	UND	5.5	UND	12	UND	10	0.7	14	UND	5.5	0.51	7.5	0.43	ND
NLN	UND	UND	UND	UND	UND	UND	0.6	0.4	0.4	UND	0.33	0.25	0.34	0.19
NLN	UND	UND	UND	UND	UND	UND	UND	UND	UND	0.3	2.1	2.1	2.6	4.3
NLN	8.9	14	14	21	14	23	17	28	8.1	14	12	20	17	20
NLN	UND	61	1.3	99	2.1	120	1.6	108	0.6	57	1	105	1.3	94
NLN	2.9	28	5.2	62	6.3	64	8.8	88	2.9	33	4.5	51	6.2	ND
NLN + S	UND	0.9	1.3	1.9	1.8	1.8	2.2	3.1	0.7	1	1.3	1.8	1.1	1.4
L	UND	207	1.2	424	2.3	397	2.1	632	0.8	274	2	468	1.5	ND
L	4.2	28	10	65	11	67	12	75	5.7	36	6.1	39	10	66
L + NLN	5.1	70	13	176	12	149	14	134	5.6	85	9.2	129	8.4	ND
L	6.	7.6	24	30	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
B	0.9	13	1.7	20	2.5	25	2	27	1	12	1.7	22	2.6	26
L + B	68	659	236	300	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
L + B + NLN	UND	UND	UND	UND	UND	UND	UND	UND	UND	UND	UND	0.12	0.02	0.1
L	UND	ND	UND	ND	UND	ND	UND	ND	UND	ND	UND	ND	UND	ND

For each assay, the functional sensitivity was taken into account. UND, Undetectable; ND, not done; K, Kryptor; IM, Immulite; TH, Thyro; AD, Advantage; DY, DYNOtest; AC, Access; IA, Iason; TB, thyroid bed; NLN, neck lymph nodes; S, skin; L, lungs; B, bones.

nants, salivary glands, or artifacts. The poor interest of routine diagnostic ^{131}I-TBS during follow-up is in accordance with recent consensus reports, but ^{131}I-TBS performed with a large activity may provide diagnostic information in selected patients with other imaging abnormalities or elevated serum Tg levels (12, 13).

Discrepancies between methods for direct determination of anti-Tg antibodies or recovery tests are obvious, demonstrating a method-to-method variability despite standardization against the WHO standard. Many researchers advocate direct quantitative determination of anti-Tg antibodies (20, 22, 23). Serum Tg determination may be unreliable in patients with anti-Tg antibodies, and neck US and possibly ^{131}I-TBS should be used for follow-up.

Tg assays should be standardized against the Tg international reference preparation (CRM 457), but this is not the only source of between-assay variability (19, 20). We analyzed the results without correction for two main reasons: first, our results can be readily applied in routine practice, and second, an even larger number of samples would have been below the level of functional sensitivity after correction, and this would not have improved the sensitivity of the method. For the same reason, we decided to exclude sera with detectable anti-Tg antibodies or with incorrect recovery from serum Tg determination with the corresponding assay.

Although the standard procedure for Tg measurement is 72 h after the second rhTSH injection, peak rhTSH stimulated Tg occurs earlier (24–48 h after the second injection) in 32% of patients (11). Thus, all serum Tg determinations were taken into account.

Determining a cutoff value for each assay seems to be particularly relevant for each preparation method, withdrawal or rhTSH, if we assume that there may be wide discrepancies between different methods for serum Tg determination. However, the limited number of events prohibited such determination and we focused on the diagnostic values of each assay. Using a common cutoff at 0.9 ng/ml, the percentage of patients with detectable serum Tg was 2-fold higher after withdrawal of L-T$_4$ therapy than after rhTSH, and this is consistent with the more intense stimulation induced by withdrawal as opposed to rhTSH (5). However, the percentage of recurrences was similar in the two groups of patients, and the great majority of patients with

TABLE 7. Tg2 level (percentage) as a function of Tg1 level with the two kits with the lowest functional sensitivity

Tg1 (ng/ml)	Tg2 (ng/ml) (%)				
	<0.1	0.1–0.5	0.5 to <0.9	0.9 to <2	>2
<0.1 (n = 616)[a]	69	23	4	2	1
0.1–0.5 (n = 126)[a]	8	46	17	17	13
0.5–0.9 (n = 17)[a]	0	0	24	18	59
>0.9 (n = 51)[a]	6	6	4	12	73
<0.1 (n = 521)[b]	65	29	3	2	1
0.1–0.5 (n = 231)[b]	8	57	13	12	10
0.5–0.9 (n = 29)[b]	3	10	17	28	41
>0.9 (n = 50)[b]	2	4	2	8	84

[a] Results obtained with e-Iason kit.
[b] Results obtained with Tg Access kit.

detectable serum Tg had no other evidence of disease, and this will be discussed below.

Assuming a common cutoff at 0.9 ng/ml, the sensitivity of serum Tg determination for tumor detection ranged from 19–40% during L-T$_4$ treatment and rose to 68–76% after TSH stimulation, whereas specificity was high in both conditions. This indeed is consistent with multiple previous reports, and clearly confirms the interest of TSH stimulation when using such kits and that serum Tg level is an excellent marker when obtained after TSH stimulation (4). Lowering the functional sensitivity of the method to 0.2–0.3 ng/ml improved its sensitivity for tumor detection during L-T$_4$ treatment to 54–65%, whereas specificity remained high. This was observed with two kits, using the functional sensitivity provided by the manufacturers and with the other two kits, using optimized functional sensitivity. Using lower functional sensitivity further improved sensitivity but at the expense of a decreased specificity. The clinical impact of Tg determination obtained after TSH stimulation, in terms of disease detection decreases with improved sensitivity of the assay. In that condition, the optimized thresholds were similar to those of the less-sensitive kits, around or above 0.9 ng/ml.

A large number of these low detectable Tg levels did not correspond to any demonstrable disease at the time of the study. Using a method with a low functional sensitivity, a recent report demonstrated a significant relationship between the level of Tg during L-T$_4$ treatment and the level reached after rhTSH stimulation (17), a finding that was confirmed by our study. In that study (17), much of the testing was performed several years after the initial therapy and cannot provide any real insight into the value of this Tg method at the 9–12 month evaluation. Low serum Tg levels could thus be related to normal residual thyroid tissue or to neoplastic foci that are too small to be imaged rather than to nonspecific effects in the assay, such as heterophilic antibody interference (24). Recent studies showed that two thirds of patients with detectable serum Tg after TSH stimulation at 9–12 months had lower or undetectable serum Tg after TSH stimulation a few months or years later, even without any further treatment (25–27). Indeed, these patients are being followed up and this will distinguish small tumor foci with an increasing trend in serum Tg level from irradiated cells that will disappear without any further treatment with a decreasing trend. In the present study, few patients with detectable Tg1 had lower Tg2, and this may be related to measurement or sampling errors or to a spontaneous decrease in Tg production between the time of Tg1 and Tg2 samplings.

In conclusion, the use of an assay with an improved functional sensitivity may indicate earlier the presence of Tg in the serum on L-T$_4$ treatment, and may then be used to study the trend of serum Tg by repeated serum Tg determinations during L-T$_4$ treatment without performing any TSH stimulation. This was observed with a functional sensitivity at 0.2–0.3 ng/ml without significant decrease in specificity. Using methods with a lower functional sensitivity will further improve the sensitivity but at the cost of a decreased specificity. The follow-up of these patients will permit to assess the significance of these detectable but low Tg levels. Serum Tg determination obtained after TSH stimulation at

9–12 months still permits a more reliable assessment of cure. In turn, the certainty of cure permits patient's reassurance, decreased daily dose of L-T$_4$ treatment that will avoid subclinical thyrotoxicosis, and limitation of follow-up to an annual clinical examination on L-T$_4$ treatment with serum TSH and Tg determinations. It should be kept in mind that the rate of recurrence is low in this population, and the majority of patients will have negative test with any method, and thus the relative medico-economic values for these methods should be evaluated.

References

1. **Travagli JP, Cailleux AF, Ricard M, Baudin E, Caillou B, Parmentier C, Schlumberger M** 1998 Combination of radioiodine (^{131}I) and probe-guided surgery for persistent or recurrent thyroid carcinoma. J Clin Endocrinol Metab 83:2675–2680
2. **Durante C, Haddy N, Baudin E, Leboulleux S, Hartl D, Travagli JP, Caillou B, Ricard M, Lumbroso JD, De Vathaire F, Schlumberger M** 2006 Long term outcome of 444 patients with distant metastases from papillary and follicular thyroid carcinoma: benefits and limits of radioiodine therapy. J Clin Endocrinol Metab 91:2892–2899
3. **Schlumberger MJ** 1998 Papillary and follicular thyroid carcinoma. N Engl J Med 338:297–306
4. **Eustatia-Rutten CFA, Smit JWA, Romijn JA, van der Kleij-Corssmit EPM, Pereira AM, Stokkel MP, Keivit J** 2004 Diagnostic value of serum thyroglobulin measurements in the follow-up of differentiated thyroid carcinoma, a structured meta-analysis. Clin Endocrinol 61:61–74
5. **Haugen BR, Pacini F, Reiners C, Schlumberger M, Ladenson PW, Sherman SI, Cooper DS, Graham KE, Braverman LE, Skarulis MC, Davies TF, De-Groot LJ, Mazzaferri EL, Daniels GH, Ross DS, Luster M, Samuels MH, Becker DV, Maxon 3rd HR, Cavalieri RR, Spencer CA, McEllin K, Weintraub BD, Ridway EC** 1999 A comparison of recombinant human thyrotropin and thyroid hormone withdrawal for the detection of thyroid remnant or cancer. J Clin Endocrinol Metab 84:3877–3885
6. **Bachelot A, Cailleux AF, Klain M, Baudin E, Ricard M, Bellon N, Caillou B, Travagli JP, Schlumberger M** 2002 Relationship between tumor burden and serum thyroglobulin level in patients with papillary and follicular thyroid carcinoma. Thyroid 12:707–711
7. **Frasoldati A, Pesenti M, Gallo M, Caroggio A, Salvo D, Valcavi R** 2003 Diagnosis of neck recurrences in patients with differentiated thyroid carcinoma. Cancer 97:90–96
8. **Pacini F, Molinaro E, Castagna MG, Agate L, Elisei R, Ceccarelli C, Lippi F, Taddei D, Grasso L, Pinchera A** 2003 Recombinant human TSH-stimulated serum thyroglobulin combined with neck ultrasonography has the highest sensitivity in monitoring differentiated thyroid carcinoma. J Clin Endocrinol Metab 88:3668–3673
9. **Torlontano M, Attard M, Crocetti U, Tumino S, Bruno R, Costante G, D'Azzo G, Meringolo D, Ferretti E, Sacco R, Arturi F, Filetti S** 2004 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
10. **Cailleux AF, Baudin E, Travagli JP, Ricard M, Schlumberger M** 2000 Is diagnostic iodine-131 scanning useful after total thyroid ablation for differentiated thyroid cancer? J Clin Endocrinol Metab 85:175–178
11. **Pacini F, Capezzone M, Elisei R, Ceccarelli C, Taddei D, Pinchera A** 2002 Diagnostic 131-iodine whole-body scan may be avoided in thyroid cancer patients who have undetectable stimulated serum Tg levels after initial treatment. J Clin Endocrinol Metab 87:1499–1501
12. **Cooper DS, Doherty GM, Haugen BR, Kloos RT, Lee SL, Mandel SJ, Mazzaferri EL, McIver B, Sherman SI, Tuttle RM, The American Thyroid Association Guidelines Taskforce** 2006 Management guidelines for patients with thyroid nodules and differentiated thyroid cancer. Thyroid 16:1–33
13. **Pacini F, Schlumberger M, Dralle H, Elisei R, Smit JWA, Wiersinga W, The European Thyroid Cancer Taskforce** 2006 European consensus for the man-

agement of patients with differentiated thyroid cancer of the follicular epithelium. Eur J Endocrinol 154:787–803

14. **Wunderlich G, Zophel K, Crook L, Smith S, Smith BR, Franke WG** 2001 A high-sensitivity enzyme-linked immunosorbent assay for serum thyroglobulin. Thyroid 11:819–824

15. **Morgenthaler NG, Froehlich J, Rendl J, Willnich M, Alonso C, Bergmann A, Reiners C** 2002 Technical evaluation of a new immunoradiometric and a immunoluminometric assay for thyroglobulin. Clin Chem 48:1077–1083

16. **Morris LF, Waxman AD, Braunstein GD** 2002 Interlaboratory comparison of thyroglobulin measurements for patients with recurrent or metastatic differentiated thyroid cancer. Clin Chem 48:1371–1372

17. **Smallridge RC, Meek SE, Morgan MA, Gates GS, Fox TP, Grebe S, Fatourechi V** 2007 Monitoring thyroglobulin in a sensitive immunoassay has comparable sensitivity to recombinant human TSH stimulated thyroglobulin in follow-up of thyroid cancer patients. J Clin Endocrinol Metab 92:82–87

18. **International Union against Cancer** 1992 Thyroid gland (ICD-OC73). In: Hermanek P, Sobin LH, eds. TNM classification of malignant tumors. 4th ed. 2nd rev. Berlin: Springer-Verlag; 35–37

19. **Feldt-Rasmussen U, Profilis C, Colinet E, Black E, Bornet H, Bourdoux P, Carayon P, Ericsson UB, Koutras DA, Lamas de Leon L, De Nayer P, Pacini F, Palumbo G, Santos A, Schlumberger M, Seidel C, Van Herle AJ, DeVijlder JJ** 1996 Human thyroglobulin reference material (CRM 457). 2nd Part: Physicochemical characterization and certification. Ann Biol Clin (Paris) 54:343–348

20. **Baloch Z, Carayon P, Conte-Devolx B Demers LM, Feldt-Rasmussen U, Henry JF, LiVosli VA, Niccoli-Sire P, John R, Ruf J, Smyth PP, Spencer CA, Stockigt JR, Guidelines Committee, National Academy of Clinical Biochemistry** 2003 Laboratory medicine practice guidelines: Laboratory support for the diagnosis and monitoring of thyroid disease. Thyroid 13:34–60

21. **Mazzaferri EL, Jhiang SM** 1994 Long-term impact of initial surgical and medical therapy on papillary and follicular thyroid cancer. Am J Med 97:418–428

22. **Mariotti S, Barbesino G, Caturegli P, Marino M, Manetti L, Pacini F, Centoni R, Pinchera A** 1995 Assay of thyroglobulin in serum with thyroglobulin autoantibodies: an unobtainable goal? J Clin Endocrinol Metab 80:468–472

23. **Spencer CA, Bergoglio LM, Kazarosyan M, Fatemi S, LoPresti JS** 2005 Clinical impact of thyroglobulin (Tg) and Tg autoantibody method differences on the management of patients with differentiated thyroid carcinomas. J Clin Endocrinol Metab 90:5566–5575

24. **Preissner CM, O'Kane DJ, Singh RJ, Morris JC, Grebe SKG** 2003 Phantoms in the assay tube: heterophile antibody interferences in serum thyroglobulin assays. J Clin Endocrinol Metab 88:3069–3074

25. **Baudin E, Do Cao C, Cailleux AF, Leboulleux S, Travagli JP, Schlumberger M** 2003 Positive predictive value of serum thyroglobulin levels, measured during the first year of follow-up after thyroid hormone withdrawal, in thyroid cancer patients. J Clin Endocrinol Metab 88:1107–1111

26. **Pacini F, Agate L, Elisei R, Capezzone M, Ceccarelli C, Lippi F, Molinaro E, Pinchera A** 2001 Outcome of differentiated thyroid cancer with detectable serum Tg and negative diagnostic ^{131}I whole-body scan: comparison of patients treated with high ^{131}I activities versus untreated patients. J Clin Endocrinol Metab 86:4092–4097

27. **Toubeau M, Touzery C, Arveux P, Chaplain G, Vaillant G, Berriolo A, Riedinger JM, Boichot C, Cochet A, Brunotte F** 2004 Predictive value for disease progression of serum thyroglobulin levels measured in the postoperative period and after ^{131}I ablation therapy in patients with differentiated thyroid cancer. J Nucl Med 45:988–994