Clinical utility of an ultrasensitive thyroglobulin assay in the follow-up of patients with differentiated thyroid cancer: can the stimulation test be avoided in patients with an intermediate recurrence risk?

Utilità clinica del dosaggio ultrasensibile della tireoglobulina nel follow-up dei pazienti con carcinoma differenziato della tiroide: il test di stimolazione potrebbe essere evitato nei pazienti con rischio di recidiva intermedio?

A. FLORES-REBOLLAR1, I. PÉREZ-DÍAZ1, S. LAGUNAS-BÁRCENAS1, B. GARCÍA-MARTÍNEZ1, R. RIVERA-MOSCOSO2, R. FAGUNDO-SIERRA3

1 Department of Internal Medicine; 2 Planning and Quality Improvement Division; 3 Central Laboratory, Instituto Nacional de Ciencias Médicas y Nutrición “Salvador Zubirán”, México City, México

SUMMARY

Serum thyroglobulin (Tg) measurement during suppression with levothyroxine (LT4) using an ultrasensitive assay (OnT4-Tg) has been proposed as a replacement of TSH-stimulated Tg measurement (OffT4-Tg) in management of patients with differentiated thyroid cancer (DTC). The aim of this study is to evaluate the capacity of an ultrasensitive Tg assay in predicting an OffT4-Tg > 2.0 ng/mL based on the OnT4-Tg in patients with DTC and an intermediate recurrence risk. We analysed 101 patients with DTC and an intermediate (n = 92) or high risk of recurrence (n = 9) who were treated with total thyroidectomy and ablation with 131I, and followed for an average of 6 years. OnT4-Tg was undetectable in 64 of 101 patients; OffT4-Tg was < 2.0 ng/mL in 61 of these 64 patients, all with negative imaging results. Furthermore, 37 of 101 patients had detectable OnT4-Tg; 32 of these 37 patients also presented OffT4-Tg > 2.0 ng/mL, and only 3 of these 32 patients had metastases detected by neck ultrasound. Considering a cutoff point of 0.1 ng/mL for OnT4-Tg, the assay had a sensitivity of 91%, specificity of 92%, positive predictive value (PPV) of 86% and the negative predictive value (NPV) of 95% when predicting an OffT4-Tg > 2.0 ng/mL (biochemical disease). The use of an ultrasensitive Tg assay allows prediction of which patients will remain disease-free even if they are at an intermediate risk of recurrence, and to decrease the need for stimulated Tg assays in two-thirds of these patients.

KEY WORDS: Thyroid cancer • Thyroglobulin • Highly sensitive thyroglobulin assay • TSH-stimulated thyroglobulin test • Neck ultrasound

RIASSUNTO

Il dosaggio della tireoglobulina (Tg) durante soppressione con levotiroxina (LT4) utilizzando la metodica ultrasensibile (OnT4-Tg) è stato proposto per sostituire il dosaggio della Tg dopo stimolazione con TSH (OffT4-Tg) nella gestione dei pazienti con carcinoma differenziato della tiroide (DTC). L’obiettivo dello studio è stato valutare la capacità del dosaggio ultrasensibile della TG di predire un valore di OffT4-Tg > 2,0 ng/mL, sulla base del valore di OnT4-Tg nei pazienti con DTC e rischio intermedio di recidiva. Sono stati analizzati 101 pazienti con DTC, di cui 92 a rischio intermedio di recidiva e 9 ad alto rischio, tutti trattati con tiroidectomia totale e ablazione con 131I, e seguiti mediamente per 6 anni. La OnT4-Tg è risultata indosabile in 64 pazienti su 101, mentre la OffT4-Tg è risultata inferiore a 2,0 ng/mL in 61 di questi 64 pazienti, tutti con imaging negativo. La OnT4-Tg è stata rilevabile in 37 pazienti su 101; 32 pazienti di questi 37 presentavano un valore di OffT4-Tg superiore a 2,0 ng/mL, e solo 3 di questi 32 avevano metastasi linfonodali apprezzabili all’ecografia del collo. Prendendo come valore cut-off 0,1 ng/mL per la misurazione della OnT4-Tg, tale dosaggio ha mostrato una sensibilità del 91%, una specificità del 92%, un valore predittivo positivo dell’86%, un valore predittivo negativo del 95%, nel predire un valore di OffT4-Tg > 2,0 ng/mL. L’utilizzo della metodica ultrasensibile del dosaggio della Tg ci permette di predire quali pazienti rimarranno liberi da malattia, anche se hanno un rischio intermedio di recidiva, e di evitare il dosaggio della Tg dopo stimolazione con TSH in due terzi di questi pazienti.

PAROLE CHIAVE: Cancro della tiroide • Tireoglobulina • Dosaggio ultrasensibile della tireoglobulina • Test della tireoglobulina dopo stimolazione con TSH • Ecografia del collo
Introduction

In recent years, the incidence of differentiated thyroid cancer (DTC) has notably increased worldwide. It is the most frequent endocrine neoplasia and accounts for 1-2% of all cancers overall; it is more common in females than in males (4:1). After total thyroidectomy (TT) and 131I therapy, most patients achieve disease-free status, but at follow-up approximately 15-20% have evidence of persistence and/or recurrence of DTC.

Thyroglobulin is a glycoprotein produced only by normal or well-differentiated malignant thyrocytes, and after TT and 131I therapy, Tg levels should be undetectable; otherwise, detectable Tg levels would suggest disease recurrence or persistence. The diagnostic precision of Tg as a tumour marker is high when the patient is treated with levothyroxine (OnT4-Tg), and it increases after stimulation with TSH either as a result of LT4 withdrawal or the administration of human recombinant TSH (rhTSH), the availability of which is limited in Mexico. Usually, Tg is considered endogenously stimulated after LT4 withdrawal (OffT4-Tg), or exogenously stimulated (rhTSH administration), if levels are above 2.0 ng/mL or 1.0 ng/mL, respectively, and these values suggest that the patient is at risk of disease recurrence or persistence.

The measurement sensitivity of Tg kits has increased in the past few years whereby the functional sensitivity (FS) of immunometric assays improved from ~1.0 ng/mL to a FS < 0.1 ng/mL; this has opened the possibility of measuring OnT4-Tg and detecting persistent or recurrent disease without the need for stimulation with TSH.

Although some studies have reviewed the functionality of these new ultrasensitive assays by measuring Tg during patient follow-up and in the prediction of TSH-stimulated Tg, most have been conducted in cases of low-risk DTC; in general, the results suggest that TSH stimulation may be unnecessary when basal Tg is undetectable, thus causing fewer unwanted effects in patients and lowering overall costs.

Serum Tg measurement and neck ultrasound (US) are standard procedures in follow-up of patients with DTC. The aim of this study is to evaluate the capacity of a serum Tg ultrasensitive assay in predicting an OffT4-Tg > 2.0 ng/mL based on OnT4-Tg in patients with DTC and an intermediate recurrence risk (ATA), using an OnT4-Tg cutoff of 0.1 ng/mL.

Materials and methods

Patient selection

One hundred and one patients with DTC were retrospectively evaluated; the group included 88 females and 13 males, with an average age at diagnosis of 39.9 years (range 10-75). They were diagnosed and treated between 1985 and 2010 at the Thyroid Clinic of the Instituto Nacional de Ciencias Médicas y Nutrición “Salvador Zubirán”; the analysed serum Tg values were obtained between January 2010 and May 2015. The study was conducted in a tertiary care hospital and approved by the Ethics Committee.

All patients underwent TT and dissection of the central compartment (except 4 cases); 99 papillary and 2 follicular (2.0%) carcinomas were detected and all were also treated with 131I for thyroid remnant ablation at doses between 100 and 150 mCi (3.7 to 5.5 GBq); in case of local recurrence or lymph node metastases, 200 to 250 mCi (7.4 a 9.2 GBq) were administered. The median cumulative 131I dose was 150 mCi (IQR 150 to 300 mCi). Only 29 patients (28.7%) received a second dose or more of 131I therapy, with a cumulative dosage of 250 to 750 mCi (9.2 to 27.7 GBq) due to local recurrence in cervical chain lymph nodes or distant metastases. Thirty of the 101 patients underwent cervical lymphadenectomy due to tumour recurrence. Most patients, 60 (59.4%) in clinical stage I, 7 (6.9%) were classified as stage II, 21 (20.8%) as stage III and 13 (12.9%) were stage IV (AJCC). After TT, whole body scan-post-ablation/treatment (WBS-PT) and initial neck US, patients were classified according to the modified 2009 risk stratification system of the American Thyroid Association (ATA). According to the ATA, 91.1% of participants had an intermediate risk and only nine cases (8.9%) were at high risk of recurring. The mean length of follow-up before the Tg stimulation test was 6.0 ± 4.5 years (range 3-30 years).

None of the 101 patients had structural evidence (neck US, WBS, neck and chest CT) of disease recurrence when Tg values were determined (OnT4-Tg/OffT4-Tg), and their OnT4-Tg levels were below 1.0 ng/mL and anti-Tg antibodies (TgAb) were negative. All patients underwent Tg stimulation (OffT4-Tg) by LT4 withdrawal for at least 4 weeks; before withdrawal, a basal Tg sample was obtained (OnT4-Tg) and both determinations also included TSH and TgAb values. The results of OffT4-Tg were considered accurate when combined with TSH value above 30 mIU/L.

All patients with a negative stimulated Tg, an OffT4-Tg < 2.0 ng/mL and a negative neck ultrasound, were considered disease-free and followed for 6 months with annual OnT4-Tg determinations and an annual neck US. Patients with a positive stimulated Tg, an OffT4-Tg > 2.0 ng/mL and suspicious neck images on US, underwent fine needle biopsy (FNAB) and Tg was measured in the needle washout (FNAB-Tg). In case the presence of...
metastases or local lymphatic recurrence was documented, lymphadenectomy was performed and if necessary a therapeutic $^{131}$I dose was administered. If suspicious images were absent in the neck US and the OffT4-Tg value was $> 2.0$ ng/mL, additional imaging studies such as WBS or neck and chest CT were requested. Positron emission tomography with fludeoxyglucose (18F) (18FDG-PET/CT) was only obtained in patients in whom $^{131}$I reactivity was suspected.

**Tg, TgAb and TSH determination assays**

Since 2005, our institution has used a 2nd generation immunoassay to determine chemiluminescent Tg (Access Thyroglobulin Assay; Beckman Coulter, Fullerton, CA) with a FS $< 0.1$ ng/mL; the intra-assay coefficient variation (CV) was $< 2.2\%$ and the inter-assay CV was $< 4.0\%$ with a cutoff limit of 0.1 ng/mL. TSH was measured with a 3rd generation chemiluminescent immunoassay (Access HYPERsensitive hTSH assay; Beckman Coulter, Fullerton, CA) and TgAb values were determined by competitive radioimmunoassay with an analytic sensitivity of 2.0 IU/mL; values below 30 IU/mL were considered negative (TGAB ONE STEP; Cisbio Bioassays, France).

**Imaging**

Cervical ultrasound (US) was performed with a multi frequency (7.5-10 MHz) lineal transducer, integrated to colour Doppler. USs were obtained within 1-3 months before or after the OffT4-Tg; they were repeated every 6 months throughout routine follow-up. All suspicious findings were biopsied with a fine needle (FNAB; gauge 23 to 25); tissue was aspirated with a 10 ml syringe and Tg was measured in the needle’s washout (Tg-FNAB). WBS was performed in a SPECT-CT (Symbia T2; Siemens) with a high energy collimator and a scanning speed of 10 cm/min. Diagnostic WBS were obtained 3 to 5 days after the oral administration of 3-5 mCi $^{131}$I (111 to 185 MBq) and the WBS-PT were performed on day 5 to 7 after radioiodine administration; the SPECT-CT was utilised since 2011. Chest and mediastinal computed tomography was performed in 5- to 10-mm thick sequential sections with a non-iodine contrast agent.

**Statistical analysis**

All results are expressed as median, interquartile range (IQR), mean ± standard deviation or percentages. The sensitivity, specificity, negative predictive value (NPV) and positive predictive value (PPV) were determined, as well as the accuracy of the OnT4-Tg at a cutoff point of 0.1 ng/ml. Data analysis was performed with IBM SPSS, version 21 (Armonk, NY, USA) statistical package.

**Results**

One hundred and one OffT4-Tg determinations were analysed, and the TSH values obtained during stimulation were $> 30.0$ mIU/L in all cases. The general characteristics of the study population are shown in Table I. The average TgAb concentration was $5.75 \pm 3.6$ IU/mL, the measured TSH level during LT4 withdrawal and Tg stimulation averaged $106.3 \pm 72.6$ mIU/L, with a median of $87.0$ mIU/L (IQR 58.2 to 129.7). Median OnT4-Tg and OffT4-Tg were 0.05 ng/mL (IQR 0.02 to 0.21) and 0.61 ng/mL (IQR 0.10 to 4.3), respectively. OffT4-Tg was $< 0.1$ ng/mL in 64 of the 101 patients, and in 61 of these 64 patients, OffT4-Tg was $< 2.0$ ng/mL; in all cases, the neck US was negative for nodal metastases.

**Table I. General characteristics of the 101 patients.**

<table>
<thead>
<tr>
<th>N</th>
<th>Percentages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at diagnosis (yr)</td>
<td>39.9 ± 14.9* 10-75**</td>
</tr>
<tr>
<td>Females</td>
<td>88 87.1</td>
</tr>
<tr>
<td>Males</td>
<td>13 12.9</td>
</tr>
<tr>
<td>Histological variety</td>
<td></td>
</tr>
<tr>
<td>Follicular carcinoma</td>
<td>2 2.0</td>
</tr>
<tr>
<td>Papillary carcinoma</td>
<td>99</td>
</tr>
<tr>
<td>Classic variant</td>
<td>55 54.4</td>
</tr>
<tr>
<td>Follicular variant</td>
<td>31 30.6</td>
</tr>
<tr>
<td>Tall cell variant</td>
<td>3 3.0</td>
</tr>
<tr>
<td>Diffuse sclerosing variant</td>
<td>6 6.0</td>
</tr>
<tr>
<td>Oncocytic variant</td>
<td>1 1.0</td>
</tr>
<tr>
<td>Solid/trabecular/insular growth pattern</td>
<td>3 3.0</td>
</tr>
<tr>
<td>AJCC stage</td>
<td></td>
</tr>
<tr>
<td>T Stage</td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>30 29.7</td>
</tr>
<tr>
<td>T2</td>
<td>17 16.8</td>
</tr>
<tr>
<td>T3</td>
<td>53 52.5</td>
</tr>
<tr>
<td>T4a</td>
<td>1 1.0</td>
</tr>
<tr>
<td>N Stage</td>
<td></td>
</tr>
<tr>
<td>N0</td>
<td>17 16.8</td>
</tr>
<tr>
<td>N1a</td>
<td>32 31.7</td>
</tr>
<tr>
<td>N1b</td>
<td>48 47.5</td>
</tr>
<tr>
<td>Nx</td>
<td>4 4.0</td>
</tr>
<tr>
<td>M Stage</td>
<td></td>
</tr>
<tr>
<td>M0</td>
<td>93 92.0</td>
</tr>
<tr>
<td>M1</td>
<td>8 8.0</td>
</tr>
</tbody>
</table>

**ATA risk**

| Intermediate | 92 91.1 |
| High | 9 8.9 |

*: average and SD; **: range; AJCC: American Joint Committee on Cancer, 7 ed.; ATA: American Thyroid Association.
Thirty-seven patients had an OnT4-Tg > 0.1 ng/mL, and in 32 of these cases the OffT4-Tg was > 2.0 ng/mL. Among the 35 patients with evidence of biochemical disease (OffT4-Tg > 2.0 ng/mL), 13 had a Tg value > 10.0 ng/mL during stimulation and only 3 had suspicious lymph nodes on US, subsequently confirmed by FNAB and Tg-FNAB; 2 patients underwent therapeutic lymphadenectomy and radioactive iodine was administered as single therapy in only one case. In the remaining patients in this group, all imaging studies (neck US, neck and chest CT, WBS or 18FDG-PET/CT) were negative; among the remaining 22 patients with a positive but < 10.0 ng/mL OffT4-Tg, further imaging studies were obtained aside from the neck US, but no structural or functional disease could be identified. No therapeutic procedures were performed and patients are being followed clinically with a “wait and see” attitude.

Using these cutoff limits, the test’s sensitivity was 91%, the positive predictive value (PPV) was 86%, the specificity was 92%, the negative predictive value (NPV) was 95% and diagnostic accuracy was 92% (see Table II).

**Discussion**

In the last few decades, follow-up strategies of patients with DTC have changed. They are currently based on measurement of serum Tg and a neck US that together, yielded a NPV of 99 to 100% [11-15]. The high sensitivity of the second generation assays (functional sensitivity < 0.1 ng/mL) used to measure Tg appears to confidently predict increases in basal Tg and results of the TSH stimulated Tg either with rhTSH or LT4 withdrawal (OffT4-Tg), according to the two most often used cutoff limits for stimulation, 1.0 and 2.0 ng/mL [5].

Unlike previous studies, ours only included patients at intermediate and high risk of recurrence. We corroborated the fact that in this patient group, a negative stimulation test (OffT4-Tg) can be predicted if their basal Tg (OnT4-Tg) is < 0.1 ng/mL, if based on endogenous TSH stimulation and an OffT4-Tg cutoff limit of 2.0 ng/mL [5]. We found 64 patients with an OnT4-Tg < 0.1 ng/mL, and 61 had an OffT4-Tg < 2.0 ng/mL; the remaining 3 cases had an OffT4-Tg > 2.0 ng/mL, slightly above the cutoff limit but all below 3.1 ng/mL and with no evidence of structural disease. The NPV was very high (95%), similar to that previously reported [8-11, 16-17].

Biochemical disease was detected in 32 of 37 patients with an OnT4-Tg > 0.1 ng/mL, but structural disease requiring treatment was identified in only 3 cases. The remaining patients in this group were reclassified as an incomplete biochemical response. The PPV was higher (86%) than that reported by other groups [9, 15, 17, 18].

Controversies have arisen in terms of the use of Tg ultrasensitive assays: one in particular is the cutoff limit of unstimulated or basal Tg and although no consensus has been reached, most authors use a value of ≤ 0.1 ng/mL, as in our study. Most publications agree that ultrasensitive Tg is highly sensitive (88 to 97%) and has a NPV of 97 to 99%, suggesting that undetectable Tg in the absence of TgAb would preclude the need for a stimulated Tg test [5]. However, a recurring problem with which all publications coincide is its low specificity and PPV, with a resulting high percentage of false positive results, ranging between 15 and 20%, depending on the study [5, 19]. Thus, if a patient has a Tg value ≥ 0.1 ng/mL, the possibility of a concomitant positive stimulated Tg test is, on average, slightly below 50% (PPV between 32 and 72%) [5]; this could expose a great percentage of patients to unnecessary testing and treatments [17]. However, perhaps the number of false positive results reported in other studies is directly related to the origin of the TSH used for Tg stimulation. Most studies have used rhTSH which neither reaches the levels nor the stimulus intensity provided by endogenous TSH. Spencer et al. measured Tg after stimulation with rhTSH in 1,029 samples after 72 hours. The median TSH was no greater than 23 mIU/L TSH [6]; although there are several studies in which rhTSH was used and the serum TSH peaks were above 30 mIU/L, it appears that the rhTSH-mediated peak is significantly affected by the patient’s age, sex, height, weight, body surface area, body mass index, blood urea nitrogen, creatinine and glomerular filtration rate [20, 21]. The fact that stimulation with rhTSH is inferior to that with endogenous TSH is a well-known fact and the Tg values obtained with rhTSH are also lower [22-24]; this has prompted some authors to reconsider the traditional cutoff limits originally used in OffT4-Tg and that were directly transferred to results using rhTSH. Recently, Kowalska et al. conducted a study in 63 DTC patients, comparing in the same subjects Tg values obtained after rhTSH stimulation.
and after OffT4-Tg; they also analysed the cutoff values for rhTSH-stimulated Tg (rhTSH/Tg). They determined that the rhTSH/Tg value with the greatest sensitivity and specificity equivalent to 2 ng/mL in the OffT4-Tg was 0.6 ng/mL, while a value of 10.0 ng/mL was 2.3 ng/mL. One of the characteristics of our study is the predominance of individuals with an intermediate recurrence risk (91.1%), an unusual variable in this type of study since most include patients with a lower risk of recurrence. As in previous studies that included patients with a high risk of recurrence, we have confirmed that in most cases it is possible to predict the OffT4-Tg response based on the OnT4-Tg using an ultrasensitive immunoassay to determine Tg, in spite of an undetectable serum Tg < 0.1 ng/mL. For this reason, several authors recommend replacing the stimulated Tg test by ultrasensitive basal Tg measurement, but only in patients with DTC and a low recurrence risk, since they mainly recruited DTC patients with a low risk.

One of our study’s limitations is the relatively low number of patients, but this was the result of the exclusion of low risk patients and the participation of only one medical centre; however, it is compensated by the inclusion of patients with an intermediate risk and the use of endogenous TSH to stimulate Tg, a clearly unusual condition in published studies. The obtained stimulation was significant, with a median TSH of 87.0 mIU/L (IQR 58.0-130.0), and reaching the recently recommended cutoff limit > 80 mIU/L or > 100 of TSH mIU/L, which leads to intense Tg stimulation in the OffT4-Tg test > 2.0 ng/mL. Although low-risk patients were excluded, we must emphasise that 63% of patients had an OnT4-Tg < 0.1 ng/mL, with no evidence of structural involvement at the time of the study; therefore, in the reclassification system dependent on response to therapy (“dynamic risk”), they would be classified as excellent responders with a predictive risk of persistent/recurrent disease that is decreased in comparison with their initial estimated risk. Although this argument may affect one of this study’s strengths, we believe that this will become a daily clinical scenario when following patients with DCT and an initial high-risk for recurrence. This situation can be expected whenever only the initial or static risk is determined without considering the dynamic or delayed risk stratification; few authors have made this point clear in recently published studies on the subject. Nevertheless, 35% of our patients had biochemical evidence of disease, which is greater than the 22% reported in the intermediate risk group (this high number could be explained by the exclusion of patients with structural disease), and much greater than the 11% recorded in the low-risk group, evaluated 24 months after initiating treatment and with a dynamic risk stratification. This should be analysed and taken into account when analysing our data.

Conclusions

The use of ultrasensitive Tg immunoassays allows prediction of which patients will be disease-free (NPV 95% and PPV 86%), even if they have an intermediate recurrence risk. Therefore, if individuals are TgAb (-), clinical follow-up may be simplified by measuring Tg during suppressive treatment with LT4 and a cervical US; these measures would optimise the endogenous TSH-stimulated Tg test, a costly test that is also uncomfortable for the patient; its use could be decreased in two-thirds of patients.

References


Rosario PW, Purisch S. Does a highly sensitive thyroglobulin (Tg) assay change the clinical management of low-risk patients with thyroid cancer with Tg on T4 < 1 ng/ml determined by traditional assays? Clin Endocrinol (Oxf) 2008;68:338-42.


Castagna MG, Maino F, Cipri C, et al. Delayed risk stratification, to include the response to initial treatment (surgery and radioiodine ablation), has better outcome predictivity in differentiated thyroid cancer patients. Eur J Endocrinol 2011;165:441-6.