A higher cut-off for Thyroid-stimulating immunoglobulin (TSI) could better predict relapse in Graves’ disease?

**Short title:** Higher cut-off for TSI and relapse in Graves’ disease


**Citation:** ©2021 Rosita Fontes. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Abstract**

Background: TSH receptor (TSHr)-stimulating immunoglobulins (Igs) can be used as diagnostic markers of Graves’ disease (GD). Thyroid-stimulating immunoglobulin (TSI) assays exclusively detect these specific Igs. Materials and Methods: This was a prospective longitudinal study in which hyperthyroid patients with GD and toxic nodular goitres were evaluated at diagnosis. GD patients were also evaluated at antithyroid drug (ATD) withdrawal. An automated chemiluminescent assay measured TSI. According to the manufacturer TSI less than 0.55 IU/L was a non-reactive result. The authors evaluated the Sensitivity (Se) and Specificity (Sp) of the cutoff point provided by the TSI assay manufacturer, and tested other cutting points through a ROC curve, to assess relapse risk of Graves’ disease. Results: At diagnosis, were evaluated 92 (85.2%) GD patients aged 41.2 ± 2.0 years, and 16 patients (14.8%) with toxic multinodular goiter (TMNG) or toxic adenoma (TA), aged 60.8 ± 4.8 years. They were re-evaluated after 18 ± 4 months with methimazole (MMI) treatment. The follow-up after treatment suspension was of 20 ± 6 months. At diagnosis, the TSI Se and Sp were 98.9% and 100%, respectively. At ATD withdrawal, despite a high Se (95.5%), Sp was low (59.6%). By adjusting the cut-off to 1.11 (TSI < 1.11 IU/L non-reactive), TSI presented the best Sp (89.4%) with a small decrease in Se (93.3%) in predicting GD relapse. Conclusions: TSI had high Se and Sp in GD differential diagnosis with nodular goiters. In the assessment for GD relapse, by raising the cutting point to 1.11 IU/L, a better Sp was obtained at the expense of a small drop in Se. A larger sample is needed to support a higher TSI cut-off point in the clinical routine to assess GD relapse after ATD.

**Keywords:** TSI, hyperthyroidism; diagnosis; treatment

**Introduction**

Hyperthyroidism is a consequence of excessive thyroid hormone production by the thyroid gland. In iodine sufficient areas, the most common cause of hyperthyroidism is Graves’ disease (GD), responsible for about 80% of cases [1, 2]. GD is an autoimmune disease mediated by immunoglobulins (Igs) that activate the TSH receptor (TSHr). This leads to TSH-independent thyroid hyperplasia and unregulated thyroid hormone (TH) production and secretion, usually accompanied by goiter [3]. It follows in frequency as causes of hyperthyroidism, toxic multinodular goiter (TMNG) and toxic adenoma (TA), that become autonomous producing TH independent of stimulus from either TSH or TSHr antibodies [4, 5].

As GD is an autoimmune disease, patients can experience relapse after stopping anti-thyroid drugs (ATD) [6]. The relapse rate after ATD withdrawal can reach 50%, depending on several factors. Measurement of TSHr stimulatory Igs can be an indirect indicator of GD activity, and be useful when assessing the prognosis [1, 7].

The thyrotropin receptor antibody (TRAb) assays measure both thyroid-stimulating and thyroid-blocking Igs. These assays have been more used
in special situations in which there is doubt about GD diagnosis. At the end of treatment, for relapse prediction, there could be doubt about ATD withdrawal when using this test, because TRAb can also detect TSHr-blocking Igs [8, 9].

The thyroid-stimulating immunoglobulin (TSI) assays measure only stimulating Igs. As thyrotoxicosis by GD is caused by stimulus of TSHr by these Igs, theoretically TSI assays would be the best marker of the disease, and should be capable of better predicting relapse [10, 11].

Laboratory and retrospective clinical studies of TSI performance compared with TRab showed a good correlation between the two tests [6, 11, 12, 13, 14, 15]. However, there are still too few prospective studies following GD patients to judge whether the TSI performance can accurately predict hyperthyroidism relapse after ATD treatment in GD.

This study aimed to evaluate, prospectively, autoimmunity before treatment for GD TMNG and TA, and recurrence risk and at the end of treatment with ATD for GD, through TSI measurement.

**Materials and Methods**

This is a prospective longitudinal study performed in a tertiary care service specialized in treating thyroid diseases of a sufficient iodine region [16, 17]. The inclusion criteria [1,2] were: age 20 years and over, suppressed TSH, and free T4 (FT4) and total T3 (TT3) above the reference intervals (RI), with symptoms and signs of hyperthyroidism for more than 3 months, presence of goiter, with or without GD ophthalmopathy, of any gender. An ultrasound examination of the thyroid determined if goiter was diffuse or nodular. Those with nodular goiter underwent thyroid scintigraphy: if there was an increase in $^{131}\text{I}$ uptake in the corresponding region to one or more nodules, they were included as TA or TMNG [4]. Before being included in the study, hormonal measurements were confirmed in more than one blood sample.

Exclusion criteria were age below 20 years, GD ophthalmopathy candidates for treatment with glucocorticoids, need for hospitalization due to the hyperthyroidism severity, pregnancy, postpartum period of up to 6 months, and use of medicines known as possible analytical or physiological interference on the measurement of TSH and/or TH in the past 3 months (medicines and contrasts containing iodine and amiodarone in the past 6 months) [6, 8].

TSH and TH were measured in patients with untreated GD, TMNG or TA before starting treatment with ATD. Once starting treatment, these hormones were measured after 1, 2, and 3 months, and every 4 months, or at shorter intervals as needed [2]. GD patients were followed until they reached criteria for discontinue ATD; BMNT and TA were followed until the definitive treatment. During treatment, the patients were seen at the thyroid specialized clinical, and everyone maintained contact with the main researcher to communicate any adverse effects or other factors that could lead to treatment interruption or non-adherence.

GD patients were evaluated again at the time of ATD withdrawal if they meet all the following criteria: 1) at least 12 to 24 months of treatment; 2) clinical euthyroidism in the last 3 months, and in use of low ATD dose during this period; 3) TSH not suppressed and within the reference interval (RI); and 4) normal FT4 and TT3 [2].

After withdrawing the medication, hormonal measurements were repeated after 1, 2, and 4, and then every 4 months until at least 24 months if patients did not relapse [2]. If during the follow-up period they presented symptoms and signs compatible with hyperthyroidism they were considered for relapse and were reassessed at any time other than planning. If they had a relapse, the TSI and hormonal measurements were repeated at the time of relapse diagnosis. GD relapse was considered if it occurred within at least 12 months after ATD discontinuation. Until February 2020, patients were followed up at the thyroid clinic to confirm clinical euthyroidism and perform the tests. Between March 2020 and March 2021 they were evaluated by teleconsultations due to the COVID-19 pandemic. From March to June 2021 they were again followed at the thyroid clinic. Hormonal and TSI measurements were interrupted between March and July 2020 when the laboratory unit responsible for blood collection was closed for public assistance. During this period, some evaluations were delayed by about 4 months.

TSI was determined by a chemiluminescent assay (Immulate 2000, Siemens Healthcare Diagnostics Inc), with a mean intra-assay percentage coefficient of variation (% CV) of 6.0%. According to the manufacturer, TSI less than 0.55 IU/L was a non-reactive result. TRab was measured by an electrochemiluminescent assay in Modular Analytics® E170, Roche Diagnostics Ltd, with amean intra-assay % CV of 7.7%. According to the manufacturer, TRab less than 1.75 IU/L was a non-reactive result.

Serum TSH (immunometric assay), FT4, and TT3 (competitive assays) were measured by electrochemiluminescence by Roche Diagnostics Ltd. The TSH interval reference (RI) was 0.4–4.3 mIU/L (20 to 59 years), 0.4–5.8 mIU/L (60 to 79 years), and 0.4–6.7 mIU/L (80 years and over). TSH was suppressed if $<$ 0.01 mIU/L. FT4 RI was 0.7–1.9 ng/dL (20 to 59 years) and 0.7–1.7 ng/dL (60 years or over) [18]. The TT3 reference interval was 70–210 ng/dL according to the manufacturer’s kit insert.

This study aimed to evaluate, prospectively, autoimmunity before treatment for GD TMNG and TA, and recurrence risk and at the end of treatment with ATD for GD, through TSI measurement.

R software (version 4.0.1) was used to perform all statistical calculations [19, 20]. We tested the sensitivity (Se) and specificity (Sp) of the TSI assay for the differential diagnosis between DG and TMNA/TA, and for prediction of relapse at the end of GD treatment using the cut-off point provided by the manufacturer. We tested other cut-off points that could provide a better Sp without Se prejudice, using a ROC curve. To assess the normal distribution of the data, Kolmogorov-Smirnov tests were performed. Data were log-transformed (base 10) before analysis for analytes with non-normal distribution. Means and standard deviations were used to demonstrate continuous variables. A two-tailed Mann-Whitney test was used to compare the nonparametric TSI and TSH distributions. The student’s t-test was used in the case of two data series with a normal distribution. The levels of analytes below the limit of quantification were considered one-hundredth below the limit of quantification for statistical analysis. The level of significance used was 0.05.

The study was approved by the Research Ethics Committee of the Instituto Estadual de Diabetes e Endocrinologia Luiz Capriglione (IEDE), CAAE 90325418.5.0000.5266. The invited subjects agreed to participate by signing the informed consent form.

**Results**

Of 124 patients, five were excluded before onset due to severe GD ophthalmopathy who needed glucocorticoid treatment. Of the remaining 119 who started the study, 11 did not complete: four because of non-adherence to treatment; two because they moved to another city; one became pregnant; and four due to the difficulty in controlling hyperthyroidism with ATD (of these, three were submitted to $^{131}\text{I}$ treatment, and one has undergone surgery. One patient was a male transgender and was not excluded from the study.

Of the 108 patients selected, 92 (85.2%) had GD, and 16 had TMNG or TA (14.8%). Age and gender by groups are in table 1. All patients were treated with methimazole (MMI). The initial MMI dose for GD ranged from 20 to 40 mg/day, and they had a titration according to clinical and laboratory evaluation during periodic visits. They were treated for 18 ± 4 months (14-26) until clinical and laboratory parameters were reached. The MMI dose used at the time of treatment discontinuation was 2.5 – 5.0 mg/day. Those with TMNG or TA were treated with 10 -20 mg/day until the definitive treatment was done.
After ATD withdrawal GD patients had a follow-up 20 ± 6 months (12–30). Laboratory data before treatment are in Table 1. Data of the TSI performance for predicting relapse after treatment with ATD are in Table 2. In those who presented relapse, it occurred in 4 ± 3 months (2-11) after ATD withdrawal. At that time there was no statistically significant difference in age, TSH, FT4, or TT3 levels between groups that relapsed or not (p = 0. 469581, p = 0.432298, p = 0.467808, p = 0.769889, respectively).

### Table 1: Epidemiological and laboratory data of hyperthyroid patients submitted to TSI measurement before treatment.

<table>
<thead>
<tr>
<th>Disease</th>
<th>Age (years)</th>
<th>Gender</th>
<th>TSI (IU/L)</th>
<th>TSH (mIU/L)</th>
<th>FT4 (ng/dL)</th>
<th>TT3 (ng/dL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GD</td>
<td>41.2 ± 2.0</td>
<td>F = 83</td>
<td>11.0 ± 9.2</td>
<td>All patients had TSH &lt; 0.011</td>
<td>3.61 ± 1.9</td>
<td>332.16 ± 122.59</td>
</tr>
<tr>
<td>TMNG/T A</td>
<td>60.8 ± 4.8</td>
<td>F = 16</td>
<td>0.11 ± 0.05</td>
<td>0.02 ± 0.01</td>
<td>2.16 ± 0.42</td>
<td>239 ± 29.5</td>
</tr>
</tbody>
</table>

The results are presented as means and standard deviations.

### Table 2: Epidemiological and laboratory data after treating GD with ATD.

<table>
<thead>
<tr>
<th></th>
<th>N (%)</th>
<th>Age (years)</th>
<th>TSI (IU/L)</th>
<th>TSH (mIU/L)</th>
<th>FT4 (ng/dL)</th>
<th>TT3 (ng/dL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relapsed</td>
<td>45 (48.9)</td>
<td>51 ± 15.7</td>
<td>2.33 ± 2.29</td>
<td>2.42 ± 1.3</td>
<td>1.12 ± 0.22</td>
<td>120.53 ± 24.17</td>
</tr>
<tr>
<td>Did not relapse</td>
<td>47 (51.1)</td>
<td>47 ± 16.9</td>
<td>0.99 ± 0.79</td>
<td>2.36 ± 1.51</td>
<td>1.09 ± 0.21</td>
<td>119.42 ± 22.26</td>
</tr>
</tbody>
</table>

The results are presented as means and standard deviations.

At the time of ATD withdrawal of GD patients, TSI levels were significantly higher in patients who had no relapse (p = 0.0135). The TSI Se and Sp to predict relapse are in Table 3. Since TRAb was the test routinely used at that time to assess autoimmunity, the results are also shown for comparison to TSI. Assessing possible other cut-off points for the TSI that could better predict the GD recurrence, a ROC curve (image not included) demonstrated that an adjustment of the cutoff to 1.1 IU/L (TSI < 1.11 IU/L being a non-reagent result) a higher Sp was achieved at the expense of a small decrease in Se.

### Table 3: TSI assay performance in diagnosis and predicting relapse at the end of treatment of patients with hyperthyroidism. Comparison with TRAb results.

<table>
<thead>
<tr>
<th>Cut-off points</th>
<th>At the diagnosis</th>
<th>For relapse prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Se</td>
<td>Sp</td>
</tr>
<tr>
<td>TSI &lt; 0.55 IU/L</td>
<td>98.9%</td>
<td>100%</td>
</tr>
<tr>
<td>TSI &lt; 1.11 IU/L</td>
<td>98.9%</td>
<td>100%</td>
</tr>
<tr>
<td>TRAb &lt; 1.75 IU/L</td>
<td>93.7%</td>
<td>100%</td>
</tr>
</tbody>
</table>

The average remission rate after a course of ATD for GD is about 50% [21]. Some factors may be associated with a higher relapse rate. One factor is the high concentration of Igs capable of stimulating the TSHr. The remission rate is lower if TSHr is elevated before the ATD suspension. [22]. As there are still few studies evaluating GD recurrence with TSI assay [2, 6], we aimed to evaluate the performance of this test in GD patients treated with ATD.

TSHr-stimulating Igs are used as diagnostic markers of GD. The TSI assay only detect Igs capable of stimulating TSHr [2]. Until some years ago, TSI was assessed by bioassays, making it difficult to incorporate the complex and expensive technique into the laboratory routine [23]. In recent years, an automated assay that directly measures TSI has been used, showing high clinical value in preliminary studies [10, 11].

### Discussion

The average remission rate after a course of ATD for GD is about 50% [21]. Some factors may be associated with a higher relapse rate. One factor is the high concentration of Igs capable of stimulating the TSHr. The remission rate is lower if TSHr is elevated before the ATD suspension. [22]. As there are still few studies evaluating GD recurrence with TSI assay [2, 6], we aimed to evaluate the performance of this test in GD patients treated with ATD.

TSHr-stimulating Igs are used as diagnostic markers of GD. The TSI assay only detect Igs capable of stimulating TSHr [2]. Until some years ago, TSI was assessed by bioassays, making it difficult to incorporate the complex and expensive technique into the laboratory routine [23]. In recent years, an automated assay that directly measures TSI has been used, showing high clinical value in preliminary studies [10, 11]. In the present study, we compared results of TSI in patients with GD with those with TMNG or TA at diagnosis, since the latter are not autoimmune diseases. In almost all GD patients TSI was reactive using the cut-off provided by the manufacturer. As expected, all patients with TMNG and TA had TSI non-reactive. In this way, TSI showed high diagnostic Se and Sp for GD diagnosis, of 98.9% and 100%, respectively. These data agree with previous data that show high TSI Se and Sp for GD diagnosis [1, 6, 11, 24].

At the end of treatment, TSI Se and Sp were reassessed to predict GD relapse. At this stage, a positive result could indicate that autoimmune activity was still present and, therefore, the high probability of disease recurrence due to the presence of Igs capable of stimulating TSHr [25]. The TSI level was higher in patients who relapsed in the following months. Other
studies have found similar results, but they were retrospective [6, 26], or performed with the TRAb measurement, an assay which is not specific for the evaluation of TSTR-stimulating IgGs [25, 27]. It is also known that patients with higher TRAB-reagent titers are likely to return to hyperthyroidism and should be given the choice of continuation of MMI treatment [8]. It can be inferred that the same happens with the TSI, and that is why it is important to define exactly what level of TSI would indicate that the treatment should be maintained longer.

We observed that some patients who did not relapse had reactive TSI, but near the 0.55 cutoff. This was responsible for the low Sp when assessing ATD withdrawal, despite the high Se. This encouraged the authors to analyze if another cut-off point could lead to a higher Sp without prejudice to the Se. Among the various points tested, the one that showed the best Sp was 1.11 IU/L, at the expense of a very small loss of Se. With this cut-off point, TSI Sp was higher to that of TRAb, still with a similar Se.

In recent years, laboratories are replacing assays that measure TRAb by those that measure TSI in commercial platforms. This is likely because the latter, in addition to being automated, is faster, promoting better cost-benefit to the laboratory routine. At least one study showed that the inclusion of TSI measurement in the current diagnostic algorithms confers cost savings and shortens time to diagnosis [10].

According to the present study, TSI proved to be clinically relevant in the assessment of diagnosis, and in prediction of GD relapse. The data obtained suggest that a higher cut-off point would better contribute to the evaluation for ATD discontinuation. As the still limited number of patients studied can lead to bias, this data now requires validation through other prospective studies.

Conclusions

In this group, TSI had high Se and Sp in the diagnosis of GD and for differential diagnosis with nodular goiters. In the assessment to GD relapse upon suspension of ATD, despite the good Se, a low Sp was obtained with the cut-off provided by the manufacturer. By raising the cutting point to 1.11 IU/L, a higher Sp was obtained at the expense of a small drop in Se, which still remained high. A larger sample is needed to support a higher TSI cut-off point in the clinical routine for the assessment of GD relapse after ATD.

References


