CLINICAL STUDY

Use of the 2nd generation TRAK human assay did not improve prediction of relapse after antithyroid medical therapy of Graves’ disease

T Zimmermann-Belsing, B Nygaard1, Å K Rasmussen and U Feldt-Rasmussen

Department of Endocrinology, Rigshospitalet, and 1Department of Endocrinology, Herlev Hospital, University of Copenhagen, Denmark

(Correspondence should be addressed to Ulla Feldt-Rasmussen, Department of Endocrinology PE-2132, University Hospital Rigshospitalet, Blegdamsvej 9, DK-2100 Copenhagen, Denmark; Email: ufeldt@rh.dk)

Abstract

Objective: Antithyroid drug treatment (ATD) is used world-wide in the treatment of thyrotoxicosis in patients with Graves’ disease (GD). The main problem is a relapse rate of 30 to 50% within 2 years after the treatment has stopped. The measurement of thyrotropin receptor antibodies (TRAb) in serum has been used to confirm the diagnosis of GD in selected patients with a diagnostic specificity of 70 to 90%. However, in predicting the recurrence of thyrotoxicosis after discontinuing ATD it has been of little value. The aim of this study was to evaluate the ability of TRAb measured by the more sensitive recombinant human TSH receptor method to predict risk of recurrence of GD after discontinuing ATD.

Materials, patients and methods: One hundred and twenty nine patients with newly diagnosed GD were included. Of these, 58 had relapse of hyperthyroidism in a follow-up of at least 11 months (median 18 months, range 11 –49) after discontinuing ATD. In 122 Graves’ patients TRAb were measured at the time of diagnosis and in all patients when discontinuing ATD by a competitive radioreceptor assay using recombinant human TSH receptors (TRAK human assay).

Results: We found an increased diagnostic specificity (99%) compared with the old TRAK porcine assay. The predictive values of a positive and negative test in relation to the prediction of a relapse of GD were found to be only 55% and 62% respectively when using a cut-off level of 1.5 IU/l, and the predictive value of a positive test decreased to 49% and of a negative test to 60% at a lower cut-off limit (1 IU/l).

Conclusion: Our study confirms that the new TRAK human assay had a superior diagnostic sensitivity in comparison with the old TRAK porcine assay. Despite the higher diagnostic sensitivity of the TRAK human method, we could not find any improvement of predictive values for relapse of hyperthyroidism in the measurement of TRAb at the end of ATD.

European Journal of Endocrinology 146 173–177

Introduction

Autoantibodies directed against the thyrotopin (TSH) receptor are responsible for the hyperthyroidism of Graves’ disease (GD) (1 – 3). TSH receptor antibodies (TRAb) affect the activity of thyroid cells either by binding to the TSH receptor, activating adenylate cyclase resulting in an increase in cyclic adenosine-3′/5′ monophosphate (cAMP) (TSH receptor binding stimulating antibodies (TSAb)), or by binding to the receptor and blocking the availability to TSH (TSH receptor blocking antibodies (TBAb)) (1 – 3). TRAb are detected either by their capacity to inhibit the ability of radiolabeled TSH to bind to receptor preparations (TBIH: TSH-binding inhibitory immunoglobulins) or by their stimulatory functional effects on thyroid tissue in vitro (TSAb) (4, 5).

Until a few years ago the validated commercial routine assay for the detection of TRAb was the competitive radio receptor assay (RRA) using porcine thyroid membrane extracts (1st generation TRAK porcine assay). This assay was widely used, although only about 70 to 90% of GD patients at diagnosis had detectable levels of TRAb in the serum (3 – 5). The measurement of TRAb was determined to be of some clinical value in confirming the diagnosis of GD but was of little value in the estimation of its prognosis (6). The introduction of the recombinant human TSH receptor has allowed the development of a new RRA (2nd generation TRAK human assay), a chemiluminescence assay in coated tube format (7). The TRAK human assay has a superior diagnostic sensitivity for GD (7 – 13) but the predictive value of relapse is controversial (12, 13), as it was for the old assay (6, 14).
The aim of this study was, therefore, to evaluate the ability of TRAb, measured by the new 2nd generation TRAK human assay to predict the risk of recurrence of GD after discontinuing antithyroid drug therapy (ATD).

Materials and methods

Patients

In order to predict the risk of recurrence of GD after ATD, 129 patients with newly diagnosed GD were included and followed (Table 1). Graves’ patients from the Departments of Endocrinology at Herlev University Hospital \( (n = 97) \) and at Copenhagen University Hospital, Rigshospitalet \( (n = 32) \) were included. The criteria for GD were hyperthyroidism with or without ophthalmopathy and a diffuse uptake on a \( { }^{99m} \text{Tc} \) pertechnetate scintigraphy. The criteria establishing hyperthyroidism were clinical symptoms, increased serum concentrations of total thyroxine \( (T_4) \), increased total triiodothyronine \( (T_3) \) and decreased TSH (Table 1).

The Graves’ patients were treated with ATD (thiamazole or propylthiouracil) with no addition of L-thyroxine (ATD) for 18 months \( (6 – 30) \) (median (range)) and followed for 18 months \( (11 – 49) \) after discontinuing ATD. At the time of diagnosis, only 122 patients with GD had concentrations of TRAb measured (Table 1).

The diagnostic specificity was calculated based on 37 normal controls (23 female, median age: 38 years \( (24 – 73) \)), and a cut-off limit of 1.5 IU/l was used (8). At the time of discontinuing ATD, all 129 Graves’ patients had serum concentrations of TRAb measured, and the sensitivity and specificity of TRAb to predict relapse were calculated.

Thyroid functional tests

At Rigshospitalet, serum concentrations of TSH (reference range: 1.5 – 2.7 mU/l, lower detection limit of 0.01 mU/l) were measured by time resolved fluor-immunoassay, (hTSH Ultra, Wallacoy, Turku, Finland), \( T_4 \) by was measured by fluorescence polarisation immunoassay (IMX, Abbott) \( (1.5 – 2.7 \text{nmol/l}) \). At Herlev University Hospital, serum concentrations of \( T_4 \) (reference range: 56 – 129 nmol/l) and \( T_3 \) \( (1.0 – 2.5 \text{nmol/l}) \) were determined by Kodak Amerlite (Amersham UK, Amersham, Bucks, UK) and serum TSH was determined by Kodak Amerlite TSH-30 (Amersham UK) with a lower detection limit of 0.005 mU/l.

2nd generation TRAK

The DYNTest TRAK human (BRAHMS Diagnostica, Berlin, Germany) measures antibodies quantitatively against the human TSH receptor. Detection is based on the ability of TRAb to prevent the binding of labeled TSH to the TSH receptor. The TRAb concentration in the patient sample was calculated from a standard curve of a sample with a known TRAb concentration. The assay was calibrated according to WHO standard 90/672. In this study, and as the manufacturer recommends, TRAK human values below 1.0 IU/l were defined as negative, values above 1.5 IU/l as positive, and patients with values between 1.0 and 1.5 IU/l were monitored for the development of autoimmune thyroid disease (8). The intra-assay variation of 3 samples of 1.35, 6.2 and 29 IU/l was 15, 8 and 8% respectively (12).

Statistics

The data are presented as medians with ranges in brackets and analyzed non-parametrically using the

<table>
<thead>
<tr>
<th>Table 1 Subject characteristics and clinical parameters of 129 Graves’ patients with positive TRAb at the time of diagnosis. Median and ranges are given.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
</tr>
<tr>
<td>Male/Female</td>
</tr>
<tr>
<td>Age at time of entering the study (years)</td>
</tr>
<tr>
<td>Family history (number with positive)</td>
</tr>
<tr>
<td>Time of treatment (months)</td>
</tr>
<tr>
<td>Follow-up time after discontinuing ATD (months)</td>
</tr>
<tr>
<td>Thyrotropin at diagnosis (mU/l)</td>
</tr>
<tr>
<td>Thyroxine at diagnosis (nmol/l)</td>
</tr>
<tr>
<td>Tri-iiodothyronine at diagnosis (nmol/l)</td>
</tr>
<tr>
<td>TRAb at diagnosis (IU/l)</td>
</tr>
<tr>
<td>TRAb when discontinuing ATD (IU/l)</td>
</tr>
</tbody>
</table>

*P < 0.05; ns, not significant.
Mann-Whitney test. The predictive value (PV) of a positive (TRAb > cut-off limit) and a negative (TRAb < cut-off limit) test was estimated as follows: $PV_{\text{pos}} = \frac{\text{number of TRAb positive patients with relapse}}{\text{total number of TRAb positive patients}}$; $PV_{\text{neg}} = \frac{\text{number of TRAb negative patients without relapse}}{\text{total number of TRAb negative patients}}$.

**Results**

At diagnosis, 116 patients were TRAb positive (>1.5 IU/l), and 6 patients were TRAb negative (<1.5 IU/l); with a cut-off limit of 1.0 IU/l, 118 patients were TRAb positive and 4 were TRAb negative (Table 2). For the normal controls, 36 were TRAb negative and 1 TRAb positive (TRAb = 1.6 IU/l) (Table 2). The diagnostic specificity of TRAb (>1.5 IU/l) in serum ($n = 122$) was 99% and the sensitivity was 86%, while at a cut-off limit of 1.0 IU/l, the specificity was similar at 99% and the sensitivity was 90% (Table 2).

At the time of discontinuing ATD, 51 patients were TRAb positive and 78 were TRAb negative (<1.5 IU/l); using a cut-off limit of 1.0 IU/l, 74 were TRAb positive and 55 were TRAb negative (Table 3). Fifty-eight of the patients had recurrence of GD and 71 remained euthyroid. Recurrence developed after 10 months (1–42) (median (range)) after discontinuing ATD. The patients with relapse were younger compared with those without relapse. The mean $T_4$ level at diagnosis was 268 nmol/l (97–608) (median (range)). mean $T_3$ was 7.1 nmol/l (2.0–19.0) and TSH was suppressed (0.01 mU/l (<0.01–0.05)) in non-relapsing patients and were not different from relapsing thyrotoxic Graves’ patients (Table 1). Both $T_4$ and $T_3$ were normalized after 3 months of therapy. No significant changes were found in TRAb at time of diagnosis or when discontinuing ATD between non-relapsing and relapsing Graves’ patients (Table 1).

The serum levels of TRAb measured at the time of diagnosis and at the time of discontinuing ATD were independent of whether the GD patients had recurrence or not. The predictive value of the presence of TRAb > 1.5 IU/l ($PV_{\text{pos}}$) at the time of discontinuing ATD ($n = 129$) in relation to relapse was 55% and $PV_{\text{neg}}$ was 62% (Table 3). When using a cut-off limit of 1.0 IU/l, $PV_{\text{pos}}$ decreased to 49% and $PV_{\text{neg}}$ to 60%.

**Discussion**

Until recently, measurement of TRAb had to rely on assays using porcine antigen (TRAK porcine assay). Costagliola and co-workers (7) in a trial of 86 patients with untreated GD and 282 healthy individuals showed a diagnostic sensitivity of 98.8% and a specificity of 99.6% using TRAK human, with a cut-off limit of 1 IU/l (7). In the present study, we have confirmed that the TRAK human assay increased the specificity for diagnosis of GD from 80% (TRAK porcine assay) to nearly 100%, which is clearly superior to the TRAK porcine assay (5, 7–11). The increase in diagnostic specificity, however, was not accompanied by an improvement in the prediction of relapse/remission at the end of ATD in GD, and the calculated predictive values were very similar to those reported in an earlier meta-analysis (6) based on reports from the TRAK porcine assay. Whether the conflicting results are due to differences in the methodology of the TRAb methods, TRAb heterogeneity, patient selection or other factors is at present not clarified, since some populations seem to demonstrate higher predictive values of serum TRAb concentrations after ATD (6, 13–18). Thyroids of some of the TRAb positive patients with GD might not have the ability to respond with hyperthyroidism because of iodine deficiency of the thyroid due to previous ATD or because of autoimmune damage of the thyroid generated during the treatment period. These phenomena might be one of the reasons why TRAb might not be the ideal predictor.

**Table 2** Diagnostic specificity and sensitivity of the 2nd generation TRAK human assay using a cut-off limit of 1.5 IU/l or 1 IU/l.

<table>
<thead>
<tr>
<th>Gravess’ patients</th>
<th>Normal controls</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRAK &gt; 1.5 IU/l</td>
<td>116</td>
<td>1</td>
</tr>
<tr>
<td>TRAK &lt; 1.5 IU/l</td>
<td>6</td>
<td>36</td>
</tr>
<tr>
<td>Number</td>
<td>122</td>
<td>37</td>
</tr>
<tr>
<td>Diagnostic specificity</td>
<td>116/117 = 99%</td>
<td></td>
</tr>
<tr>
<td>Diagnostic sensitivity</td>
<td>36/42 = 86%</td>
<td></td>
</tr>
<tr>
<td>TRAK &gt; 1 IU/l</td>
<td>116</td>
<td>1</td>
</tr>
<tr>
<td>TRAK &lt; 1 IU/l</td>
<td>4</td>
<td>36</td>
</tr>
<tr>
<td>Number</td>
<td>122</td>
<td>37</td>
</tr>
<tr>
<td>Diagnostic specificity</td>
<td>118/119 = 99%</td>
<td></td>
</tr>
<tr>
<td>Diagnostic sensitivity</td>
<td>36/40 = 90%</td>
<td></td>
</tr>
</tbody>
</table>
in some GD populations. In our study, the follow-up period was at least 11 months, which should be sufficient to exclude the possibility of overlooking hyperthyroidism in TRAb positive patients. It has been suggested that the rate of fall of TRAb is faster in the second 6 months of ATD in those patients who will eventually achieve long-term remission (19); this was, however, in contrast to another study (20), but different assay systems were used in these studies. However, Maugendre and Massart (11) suggested that the measurement of TRAB values between 6 and 12 months after the start of ATD might be too early to predict relapses in GD. It has also been suggested that relapsing Graves’ patients at the time of diagnosis and/or at time of discontinuing ATD had higher TRAb (TRAK porcine) compared with GD patients without relapse (19); we could not confirm this in our study using the TRAK human assay (Table 1).

Autoantibodies directed to the TSH receptor might either mimic the stimulating action of TSH or inhibit the action of the hormone, and the heterogeneous nature of TRAbs is well recognized (21). TSAb is pathogenic of GD, but serum from most Graves’ patients probably contains both TSAb and TBAb activities, and the clinical effect may depend on the relative concentration and affinity of the predominating antibody (1–5). However, TBAb seem to be rare in Caucasians and more common in Asians (22).

The choice of treatment in patients with GD seems to be, at least partly, dependent on the country in which the patient has been treated (15). ATD is mainly used in Europe and Japan while radioiodine is often preferred in the USA (15). However, some endocrinologists in the USA suggest ATD (12). In the majority of European countries the patients are treated for 1–2 years and thereafter they are followed-up by measurement of thyroid hormones (23). However, some endocrinologists monitor the treatment by measurement of the TRAb level (12, 24). The main problem is a relapse rate of between 30 and 50% within 2 years after the treatment has stopped and a variety of different predictors of relapse have been looked for such as age, goiter size, radioactive iodine uptake, degree of ophthalmopathy, serum concentrations of thyroglobulin (Tg), anti-thyroid peroxidase (anti-TPO), anti-Tg or TRAb, HLA DR typing, T3 suppression test or ATD (25). Even in various combinations these indices have proven unsatisfactory as reliable predictors of remission (13, 14, 16–18, 21, 23, 26–33). This was confirmed in the meta-analysis of data from 10 prospective studies in the literature, which showed a statistically significant difference between the number of relapses in patients with or without TRAb at the end of ATD treatment. The predictive value of both a positive and a negative test was low and 25% of patients were misclassified (6) but others have disagreed with this conclusion (12).

In forecasting the transient neonatal syndromes in the offspring of women with autoimmune thyroid disease, the TRAK porcine assay was found to be reliable (23, 34).

In principle, the bioassay might be preferable in the prediction of relapse of GD after ATD because it is possible to measure both stimulating and blocking antibodies separately; however, the assays are cumbersome and expensive for routine purposes and in practice are not superior to other methods (3–5, 12, 23).

In conclusion, our study confirms that the new TRAK human assay had a superior diagnostic sensitivity in comparison with the old TRAK porcine assay. Despite the demonstrated higher sensitivity of the TRAK human method, we could not find any improvement of predictive values for relapse or remission by measuring TRAb at the end of ATD in GD.

Acknowledgements

We thank BRAHMS Diagnostic for providing the TRAK kits as well as the laboratory technicians Lisbeth Kirkegaard, Betty Fischer, Ulla K Hansen and colleagues for handling the patient samples and TRAb measurements. This work received financial support from Agnes and Knut Morks Foundation.

References


