Intermethod Variability in TSH-Receptor Antibody Measurement: Implication for the Diagnosis of Graves Disease and for the Follow-Up of Graves Ophthalmopathy

Catherine Massart,1,2* Rémy Sapin,3,4 Jacqueline Gébassier,1 Arnaud Agin,3,4 and Michèle d’Herbomez5,6

1 Unité Fonctionnelle d’Hormonologie, CHU de Rennes, France; 2 INSERM 0203 Centre d’Investigation Clinique, Université de Rennes 1, France; 3 Laboratoire d’Explorations Fonctionnelles par les Isotopes, CHRU de Strasbourg, France; 4 ULP/CNRS UMR 7191, Faculté de Médecine, Université Louis Pasteur, Strasbourg, France; 5 Laboratoire de Médecine Nucléaire, Centre de Biologie-Pathologie, CHRU de Lille, France; 6 Faculté de Médecine, Université Lille 2, France. * address correspondence to this author at: 1 Unité Fonctionnelle d’Hormonologie, CHU de Rennes, France. Fax: 02 99 28 41 45; E-mail: catherine.massart@chu-rennes.fr.

BACKGROUND: We compared the analytical and clinical performance of 3 porcine thyroid receptor antibody (TRAb) methods (1 second- and 2 new third-generation systems) with the conventional TRAb assay based on the human recombinant TSH receptor (hTRAK).

PATIENTS AND METHODS: We obtained sera from 86 patients with untreated Graves disease (GD) and 71 healthy controls. We measured TRAb concentrations by radioreceptor assay using the hTRAK (Brahms) or the porcine TSH receptor (pRRA) from Beckman-Coulter, by electrochemiluminescence immunoassay (ECLIA) with the Elecsys/Cobas (Roche), and by ELISA using the Medizym TRAb clone (Medipan).

RESULTS: Between-run assay imprecision was ≤10% and ≤7.6% for hTRAK and ECLIA, but reached 14% and 14.9% for ELISA and pRRA, respectively. Maximal specificity and sensitivity close to 100% were obtained for hTRAK, ECLIA, and ELISA. pRRA failed to detect positive TRAbs in 5 GD patients. Although calibrated against the same reference standard 90/672, the assays displayed a high intermethod variability. The results were significantly higher by ECLIA and lower by ELISA and pRRA compared with hTRAK. Patients with ophthalmopathy had higher TRAb results by ELISA and pRRA than those without eye disease.

CONCLUSIONS: Second- and third-generation TRAb assays had similar diagnostic sensitivities in the diagnostic evaluation of GD. Despite the use of the same reference standard for calibration, high intermethod variability in TRAb assay results was seen in untreated GD patients. Assay harmonization is necessary for correct interpretation in the follow-up of Graves ophthalmopathy.

In Graves disease (GD), antibodies to thyrotropin-stimulating hormone receptor (TSH-R) are detected either by their ability to stimulate adenylate cyclase [thyroid stimulating antibody (TSAb)] or by their capacity to compete for the binding sites of the TSH-R [thyroid receptor antibody (TRAb)] (1, 2). Second-generation TRAb assays using recombinant human TSH-R (hTRAK) showed high sensitivity (3–5), especially in the diagnosis of GD (6, 7). Third-generation TRAb assays, using a monoclonal thyroid stimulating antibody M22 which binds to a conformational epitope of TSH-R within the thyroid-stimulating hormone (TSH) binding pocket, have recently been described. Among them, the M22-biotin–based ELISA (8, 9) and the automated Elecsys/Cobas electrochemiluminescence immunoassay (ECLIA) (10) have been developed. However, conflicting results have been reported on the advantages of the ELISA over earlier methods (8, 9).

For a better understanding, we compared the analytical and clinical performance of different TRAb methods (2 second- and 2 third-generation assays) in a large cohort of untreated Graves patients.

Materials and Methods

The study involved 86 patients (67 women and 19 men; median age 40 years, range 18 – 66) with GD diagnosed from typical clinical signs, increased free thyroid hormone concentrations, and undetectable TSH values (<0.05 mIU/L). Seventy-one healthy sex- and age-matched blood donors, euthyroid and negative (<60 kIU/L) for antithyroperoxidase antibodies (RIA DYNOTest anti-TPOn; Brahms) served as controls. All the procedures used for study participant recruitment were in accordance with Helsinki Declaration of 1975 as revised in 1996.

Sera were stored at –20 °C or –80 °C until assay without any freeze/thaw cycle.
TSH-R ASSAYS

We performed second-generation TRAb assays using either the Dynotest TRAK human (hTRAK) from Brahms Diagnostica or the porcine RRA anti-RTSH (pRRA) from Beckman-Coulter. We performed third-generation TRAb assays using a competitive ELISA (Medizym TRAb clone; Medipan) and the automated Elecsys/Cobas ECLIA (Roche Diagnostics). All assays were calibrated according to the WHO standard 90/672 and were carried out according to manufacturer recommendations. Positive cutoff values were 1.5 IU/L, 1.75 IU/L, 0.4 IU/L, and 1.5 IU/L for hTRAK, ECLIA, ELISA, and pRRA, respectively. Concentrations between 1 and 1.5 IU/L were considered a "gray area" for both hTRAK and pRRA assays.

In samples with discordant results, we assayed TSAb concentrations using human thyrocyte cultures (hTSAb) and JP26/26 CHO cells transfected with the recombinant TSH-R (TSAb CHO), as described (11, 12). We also measured thyroid-stimulating blocking antibody (TSBAb) activity using JP26/26 CHO cells (5). Values >125%, >127%, and >30% were considered positive for hTSAb, TSAb CHO, and TSBAb, respectively.

STATISTICAL ANALYSIS

We analyzed data using Wilcoxon matched-paired test and Mann–Whitney test for group comparison (Statview version 4.11; Abacus Concepts). We used the method of Bland and Altman (13) for method comparison of hTRAK with other TRAb methods.

Results

Repeatability as evaluated from the within-run CV for 8 assays was ≤7.9%, ≤5.9%, ≤7.2%, and ≤11.7% for hTRAK, ECLIA, ELISA, and pRRA, respectively. Reproducibility as determined from between-run imprecision in 9 series of assays was ≤10%, ≤7.6%, ≤14%, and ≤14.9% for hTRAK, ECLIA, ELISA, and pRRA. Dilution tests performed in the zero standard or TRAb-free serum according to the manufacturers’ instructions were found to be nonlinear for the different TRAb assays. Thus, we decided to set the upper concentration at 40 IU/L or at 30 IU/L according to the highest concentration of the standard curve.

Using the cutoff values of the manufacturers, all 4 TRAb assays showed a very high specificity (95% CI) calculated at 100% (94.9%–100%), 100% (94.9%–100%), 98.6% (92.4%–100%), and 100% (94.9%–100%) for hTRAK, ECLIA, ELISA, and pRRA. The sensitivity (95% CI) was 98.8% (93.7%–100%), 100% (95.8%–100%), 98.8% (93.7%–100%), and 90.7% (82.5–95.9%). The use of the inferior limit of the gray area (1 IU/L) as cutoff for hTRAK and pRRA did not change the specificity but increased the sensitivity to 100% (95.8%–100%) and 94.2% (87.0–98.1%), respectively.

We analyzed the results of negative-TRAb GD patients (Table 1). The patient with a borderline result for hTRAK (1.2 IU/L) was found to be positive with ECLIA (2.2 IU/L) but negative with ELISA and with pRRA. TSAb and TSBAb were negative in only 1 of these 3 patients. Four patients were TRAb-negative with the pRRA but positive for the other TRAb assays and for both TSAb assays.

The TRAb concentrations obtained with the 4 immunoassays are reported in Fig. 1. The results obtained with the hTRAK were significantly different (Wilcoxon test) from those found with ECLIA (Z – 7.58; P < 0.01) and pRRA (Z – 3.60; P < 0.01).

<table>
<thead>
<tr>
<th>Patient</th>
<th>hTRAK, IU/L</th>
<th>ECLIA, IU/L</th>
<th>ELISA, IU/L</th>
<th>pRRA, IU/L</th>
<th>hTSAb, %</th>
<th>TSAb CHO, %</th>
<th>TSBAb, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.2&lt;sup&gt;a&lt;/sup&gt;</td>
<td>2.2</td>
<td>&lt;0.22&lt;sup&gt;b&lt;/sup&gt;</td>
<td>&lt;1&lt;sup&gt;b&lt;/sup&gt;</td>
<td>111&lt;sup&gt;b&lt;/sup&gt;</td>
<td>101&lt;sup&gt;b&lt;/sup&gt;</td>
<td>&lt;5%&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>2</td>
<td>2.8</td>
<td>2.8</td>
<td>0.69</td>
<td>1.3&lt;sup&gt;b&lt;/sup&gt;</td>
<td>103&lt;sup&gt;b&lt;/sup&gt;</td>
<td>117&lt;sup&gt;b&lt;/sup&gt;</td>
<td>12%&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>3</td>
<td>2.4</td>
<td>3.1</td>
<td>0.62</td>
<td>1.1&lt;sup&gt;a&lt;/sup&gt;</td>
<td>155</td>
<td>169</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>4.1</td>
<td>5</td>
<td>3.8</td>
<td>1.4&lt;sup&gt;b&lt;/sup&gt;</td>
<td>275</td>
<td>305</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>2.7</td>
<td>2.7</td>
<td>2.2</td>
<td>&lt;1&lt;sup&gt;b&lt;/sup&gt;</td>
<td>131</td>
<td>135</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>2.9</td>
<td>6.8</td>
<td>1.5</td>
<td>&lt;1&lt;sup&gt;b&lt;/sup&gt;</td>
<td>133</td>
<td>172</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>4.9</td>
<td>4.4</td>
<td>2.3</td>
<td>&lt;1&lt;sup&gt;b&lt;/sup&gt;</td>
<td>136</td>
<td>162</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>2.6</td>
<td>3.1</td>
<td>1.9</td>
<td>&lt;1&lt;sup&gt;b&lt;/sup&gt;</td>
<td>162</td>
<td>131</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> Borderline results (gray area) as reported by the manufacturers.

<sup>b</sup> Negative result.
10^{-4}), ELISA (Z \sim 7.25; P < 10^{-4}), and pRRA (Z \sim 6.93; P < 10^{-4}). By contrast, no significant difference was found between the ELISA and the pRRA titers.

As shown in Supplemental Fig. 1A, available with the online version of this paper at www.clinchem.org/content/vol55/issue1, the method of Bland and Altman showed a positive difference between the results obtained with ECLIA and hTRAK (mean difference value 3.7 IU/L, 2 SDs of the differences 8.2 IU/L). Negative differences (online Supplemental Fig. 1B and 1C) were observed between the data with ELISA or pRRA and those with hTRAK (mean difference 4.4 IU/L, 2 SDs 10.8 IU/L for ELISA and 4.1 IU/L, 2 SDs 12.2 IU/L for pRRA). All these findings demonstrated the higher results obtained with ECLIA (median 12.2 IU/L, range 2.9–40 IU/L) and the lower values found with ELISA (median 5.2 IU/L, range 1–30 IU/L) or pRRA (median 5.2 IU/L, range 1–31 IU/L) than those observed with hTRAK (median 9.2 IU/L, range 2.5–40 IU/L).

The patients were divided into 2 groups according to the absence (group 1; n = 67) or the presence (group 2; n = 19) of ophthalmopathy before treatment (online Supplemental Table 1). TRAbs were similar in the 2 groups for hTRAK and ECLIA. In contrast, the results obtained with the ELISA were significantly higher (Mann–Whitney test) in the patients with ophthalmopathy (median 9.5 IU/L, range 2.5–29 IU/L) than in those without eye disease (median 4.3 IU/L; range 0.7–30 IU/L; Z \sim 2.11; P = 0.035). Significantly increased (Z \sim 1.98; P = 0.048) TRAbs with the pRRA were also observed in group 2 (median 8.2 IU/L, range 2.2–27 IU/L) compared with group 1 (median 4.5 IU/L; range 1–30 IU/L).

Discussion

In this study, we confirmed poor precision reported by the manufacturers for ELISA and pRRA, with CVs >10% for both methods. Our results also showed that the ECLIA displayed lower imprecision than the other assays. However, between-run variability of this assay should be confirmed with at least 3 batches of reagents, as recommended (14).

We observed high intermethod variability in TRAb results despite the use of the same reference standard in all methods for calibration. The type of TRAb—stimulating vs blocking and directed against different parts of the ectodomain of the TSH-R (15)—although a subject of debate (16), is not a likely explanation for these differences, since no TSABab was detected in any of our TSAb-negative patients.

The lower concentrations detected with the ELISA are in agreement with findings of Rees Smith et al. (8). However, the detection of lower TRAb concentrations did not increase the clinical sensitivity, since the patient with the hTRAK borderline result was found to be negative with ELISA. In contrast, a positive result was obtained for this patient with ECLIA. Thus, the use of the M22 antibody in the 2 third-generation assays did not give equal clinical sensitivity in the diagnostic evaluation of GD. These intermethod discrepancies may be due to differences in accessibility of the different epitopes of the TSH-R when coated on plastic surface for 3 assays or solubilized for ECLIA. Use of 125I-TSH vs labeled-M22 for competitive binding also may explain these conflicting data.

We found that the pRRA was less sensitive than the other methods. Thus, caution should be taken for interpretation of the results by this method in the diagnosis of GD.

TRAbs have been associated with the clinical features of ophthalmopathy in untreated GD patients (17–19). We demonstrated variable TRAb results according to the assay used. Higher TRAb concentrations were found in the patients with ophthalmopathy when measured with ELISA and pRRA, as reported for the pRRA (19).
In conclusion, the second- and third-generation TRAb assays exhibit very similar diagnostic sensitivities in the evaluation of GD. Despite use of the same reference standard for calibration, the different TRAb assays display high intermethod variability in untreated Graves patients. The lack of standardization among TRAb assays remains an important problem in the diagnosis of GD and in the follow-up of Graves ophthalmopathy. Because TRAb measurement by the hTRAK at the end of treatment appears to have questionable value (6, 20), the clinical relevance of third-generation assays needs further investigation.

Author Contributions: All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.

References

15. Morgenthaler NG, Ho SC, Minich WB. Stimulating and blocking thyroid-stimulating hormone (TSH) receptor autoantibodies from patients with Graves’ disease and autoimmune hypothyroidism have very similar concentration, TSH receptor affinity, and binding sites. J Clin Endocrinol Metab 2007;92:1038–45.

DOI: 10.1373/clinchem.2008.115162