Analytical and diagnostic accuracy of "second generation" assays for thyrotrophin receptor antibodies with radioactive and chemiluminescent tracers

D Villalta, E Orunesu, R Tozzoli, P Montagna, G Pesce, N Bizzaro, M Bagnasco

Aims: To investigate the analytical and diagnostic accuracy of thyrotrophin (TSH) receptor antibody assays using recombinant human TSH receptors.

Methods: Sera from 68 patients with Graves' disease, 23 patients with autoimmune thyroiditis, and 119 healthy controls were evaluated in four different laboratories using both radioactive and chemiluminescent tracers. Functional sensitivity, interlaboratory precision, optimal cutoff values for Graves' disease, and the correlation between the two methods were evaluated.

Results: Functional sensitivity was 0.98 IU/litre for both assays. Interlaboratory precision, expressed as per cent coefficient of variation over a wide range of antibody concentrations, varied from 5.7% to 15.1% for the radioligand, and from 6.6% to 19.9% for the chemiluminescence assay. The two methods (radioactive and chemiluminescent) were closely correlated. All the sera from untreated or relapsing patients with Graves' disease gave TSH receptor antibody values above 2.1 IU/litre, whereas in none of the healthy controls did values exceed 2.5 IU/litre. Receiver operating curve analysis allowed an optimal cutoff point to be defined at 1.99 IU/litre, according to a sensitivity of 100% and specificity of 99.1%.

Conclusions: These data show the high analytical and diagnostic accuracy of the human TSH receptor assays, both with radioactive and chemiluminescent tracers, when both functional sensitivity and interlaboratory reproducibility are considered. These two methods could be proposed as first line diagnostic markers for Graves' disease.
METHODS

Our study included serum samples from 68 patients with GD diagnosed according to the standard clinical criteria (suppressed TSH, raised free T3 and free T4, goitre, sonography, and signs of Graves’ ophthalmopathy when present), namely: 19 untreated patients with active GD; 32 patients undergoing antithyroid drug treatment; six patients in clinical remission; and 11 patients who had relapsed after antithyroid drug treatment withdrawal. Among the 68 patients with GD evaluated in our study, 23 were classified as having ophthalmopathy of any degree and five as having “severe ophthalmopathy”, defined as the presence of one of the following: pronounced proptosis, reduction of visual acuity attributable to optic neuropathy, extraocular muscle dysfunction with resulting diplopia, or extremely severe periorbital swelling with conjunctival hyperaemia and chemosis. No patient had pretibial myxoe dema. By means of palpation and thyroid ultrasound evaluation, the thyroid volume was classified as greatly increased in seven of the 68 patients and multinodularity (defined as the presence of two or more nodules) was present in five patients.

In addition, we studied 18 patients with Hashimoto’s thyroiditis (diagnosed on the basis of clinical or subclinical hypothyroidism, enlargement and hypoecogenicity at thyroid sonography, and the presence of antithyroid peroxidase and/or antithyroglobulin antibodies), five patients with atrophic thyroiditis (diagnosed on the basis of clinical hypothyroidism, atrophy and hypoecho genicity at thyroid sonography, and the presence of antithyroid peroxidase and/or antithyroglobulin antibodies), and 119 healthy controls (without a personal or familial history of thyroid disease; median age, 41.3 years; male to female ratio, 1.83 : 1; all euthyroid, negative for antithyroid peroxidase and antithyroglobulin autoantibodies).

For the quantitative determination of IgG antibodies to human thyroid peroxidase and human thyroglobulin, we used two step enzyme linked immunosorbent assays (anti-TPO IEMA; RADIM, Rome, Italy; TgAb IEMA WELL; RADIM). Antithyroid peroxidase antibodies > 100 U/ml and antithyroglobulin antibodies > 150 U/ml were considered to be positive.

Sera were collected in the outpatient clinic of the department of internal medicine, University of Genoa, Italy, and aliquoted into small volumes for frozen shipment to the participating laboratories. Antibody measurements were performed in duplicate in four different laboratories: two of them used the DYNOn test-Trak assay and the other two used the LUMIt est-Trak assay, following the manufacturer’s instructions.

To determine the lowest concentration of serum TBII at which the assays are expected to report clinically useful results, we measured the “functional sensitivity”11. To determine the functional sensitivity of both assays, relevant interassay precision profiles for TBII were constructed with human serum pools, the lowest being less than 10% above the detection limit. The pools were measured in two of the participating laboratories by two different reagent lots of the LUMItest and six different lots of the DYNOn test in 10 different runs, following the protocol recommended by the National Academy of Clinical Biochemistry.12 To give a clinically relevant interassay precision profile for TBII, we measured the serum pools over a period of six months, because we thought that a shorter period of time would not be realistic enough to represent the clinical intervals used in the setting of thyroid outpatients.

To assess the interlaboratory precision of h-TBII assays all the sera of the above mentioned patients were assayed in the four participating laboratories. Interlaboratory coefficients of variation (CVs) were calculated from samples with different antibody concentrations covering the entire range of the standard curve. To achieve an analytical method comparison, Deming regression analysis was performed. The BlandAltman difference plot was used to visualise the comparison results.

RESULTS

To determine the inter laboratory precision for the DYNOn test, LUMIt est, and the combination of both, we used the CVs of all the sera from the patients and controls. Figure 1A shows the interlaboratory CVs, obtained from the CVs of sera with median TBII concentrations covering the entire range of the standard curve, and in particular, concentrations of 1, 2, 10, 20, and 40 IU/litre. CV values for serum TBII concentrations of 1, 2, 5, 20, and 40 IU/litre were 15.1%, 10.5%, 8.4%, 5.7%, and 9.0%, respectively, for the DYNOn test, and 19.9%, 13.0%, 9.4%, 14.3%, and 6.6%, respectively, for the LUMIt est. These results showed excellent precision for values > 2 IU/litre for the DYNOn test and good precision for the LUMIt est in the same range of TRAb concentrations, showing a CV < 15% (fig 1A).

The dilution test gave straight dilution lines for the 10 sera tested. The dilution curves paralleled the calibration curve in both assays.

Figure 1  (A) Interassay coefficients of variation (CVs) for the chemiluminescence assay, the radioreceptor assay, and the combination of both. (B) Interassay variability, evaluated in two laboratories using at least two different lots of reagents, over a period of over six months. The functional sensitivity (20% CV) corresponds to a concentration of antibodies to the thyrotrophin receptor of 0.98 IU/litre, both for the human thyrotrophin receptor chemiluminescence assay (LUMItest-Trak human) and for the radioreceptor assay (DYNOn test-Trak human).
There was a high positive correlation between the radioactive and the non-radioactive tests. The Deming regression equation was: $y = 0.957$ (95% confidence interval (CI), 0.9181 to 0.9960) $x + 0.5355$ (95% CI, 0.0474 to 1.0237) IU/litre ($R = 0.97$), where $x$ is the TBII concentration measured by the DYNOtest, and $y$ is that measured by the LUMItest (fig 2A). To visualise the comparison results, we performed the Bland-Altman difference plot (fig 2B), which shows a mean bias of 0.23 IU/litre (95% CI, −0.178 to 0.641). For high TBII values, the difference between the two methods is higher than for low values, but for clinical purposes this difference is irrelevant.

The functional sensitivity was 0.98 IU/litre for both assays, confirming the value reported by the manufacturer (fig 1B). Notably, this value is slightly lower when compared with the value reported in a previous study.10

All the sera from untreated or relapsing patients with GD gave TSH receptor antibody values above 2.1 IU/litre using both methods, whereas in none of the healthy controls did values exceed 2.5 IU/litre.

Based on data expressed in IU/litre, we performed receiver operator curve plot analysis, including healthy individuals for specificity and patients with untreated GD for sensitivity. For the DYNOtest the mean area under the curve was 0.997 (95% CI, 0.99 to 1.00) and for the LUMItest it was 0.999 (95% CI, 0.99 to 1.00). The most discriminating cutoff value was found at 1.47 IU/litre for the DYNOtest and at 1.69 IU/litre for the LUMItest (fig 3A, B). For both assays these values correspond to a sensitivity of 100% and a specificity of 99.1%. We set the decision threshold for both the LUMItest and the DYNOtest at 1.99 IU/litre, the lowest value at which the diagnostic accuracy of both assays is preserved (sensitivity, 100%; specificity, 99.1%). At this TBII concentration, the DYNOtest achieves a CV < 10% and the LUMItest a CV < 15%, both being highly satisfactory for clinical purposes.

The manufacturer suggests a cutoff value of 1.5 IU/litre; values < 1 IU/litre are considered to be negative and those between 1 and 1.5 are categorised as “grey zone”. Because of the excellent interlaboratory determined diagnostic accuracy of both assays, we were able to set a clear cutoff.

Figure 2  (A) Deming regression plot for the correlation between the LUMItest and the DYNOtest. The regression equation, represented by the solid line, is: $y = 0.957$ (95% confidence interval (CI), 0.9181 to 0.9960) $x + 0.5355$ (95% CI, 0.0471 to 1.0237) IU/litre ($R = 0.97$). (B) Bland-Altman plot displaying the per cent difference values between the two methods (y axis) and their average in IU/litre (x axis). The dashed lines represent ± 2 SD.

Figure 3  Receiver operator curve plot analysis, including the data from patients with active Graves’ disease for sensitivity and healthy controls for specificity. (A) DYNOtest-Trak; (B) LUMItest-Trak. AUC, area under the curve.
at 1.99 IU/litre, with no false negative results and only one false positive; thus, we were able to eliminate the grey zone. Therefore, values below 1.99 IU/litre, and higher than the functional sensitivity (0.98 IU/litre), should be considered as negative.

By applying this cutoff value to the patients involved in our study, the clinical accuracy proved to be excellent with both the DYNOtest and LUMItest: h-TBII were positive in 19 of 19 patients with untreated GD, one of 119 healthy controls, 21 of 32 patients with GD under treatment, two of six patients with GD in clinical remission after antithyroid drug withdrawal, 11 of 11 patients with GD who had relapsed, one of five patients with atrophic thyroiditis (probably as a result of TSH receptor blocking antibodies), and none of 18 patients with Hashimoto’s thyroiditis, with both assays showing essentially the same results (fig 4).

DISCUSSION

The classification of hyperthyroidism is the first potential clinical application of TBII measurement in the thyrotoxic patient. TBII assays based on porcine membrane derived TSH receptor preparations are still widely used but, although highly specific, they exhibit an average sensitivity of only 70%, which is far from satisfactory. The guideline of the National Academy of Clinical Biochemistry suggests that the clinical use for TBII measurement should be: “... to investigate the aetiology of hyperthyroidism when the diagnosis is not clinically obvious”; therefore, it is not considered mandatory as a first line investigation for GD. This is because about one in three patients with GD are TBII negative with the porcine derived TBII assays.

Our results show that over 99% of patients with active GD have positive serum TRAbs when tested with the new human recombinant TBII assay.

To test the diagnostic accuracy of this method further we selected an additional group of patients with GD who had a less clear clinical presentation and tested the specimens in one laboratory. The group consisted in 16 patients with GD (12 women and 4 men aged between 30 and 67 years) with no evident ophthalmopathy or goitre. The serum TBII concentrations ranged from 2.5 to 29.9 IU/litre, supporting the observations of our study and confirming the absence of a grey zone, even in patients with a less clear disease presentation.

The improvement in diagnostic accuracy and analytical quality of human recombinant receptor assays increase substantially the impact of these parameters on clinical decision making.

The first indication for TBII measurement is the diagnosis of GD at presentation. However, the clinician can profit from the detection of TBII for other clinical indications, namely: the evaluation of GD activity at the time of antithyroid drug treatment withdrawal and during the first year after withdrawal, its concentration correlating with the chance of short term relapse of disease"; possible prediction of relapse at GD presentation—patients with high TBII may be prone to relapse after antithyroid drug withdrawal"; and the evaluation of GD during pregnancy for fetal risk as a result of transplacental passage of TBII. During pregnancy, all female patients under treatment, or with a history of GD treated either by surgery or radioiodine, should be tested for TBII in each trimester. Monitoring of TBII concentration allows risk assessment for possible consequences of transplacental passage of antibodies, such as goitre or hyperthyroidism, with growth retardation of the fetus, or emergencies at delivery, such as newborn hyperthyroidism.

Our data show the high analytical precision of the commercially available recombinant human TSH receptor radioassay with regard to interlaboratory reproducibility.

Previous studies that have compared porcine TSH receptor assays with human recombinant TSH receptor assays have shown the superior diagnostic accuracy of the latter. Our data are in agreement with these studies. Considering all these findings, in addition to the high analytical and diagnostic performance, these methods could be proposed as a first line diagnostic marker for GD. The serum concentration of TBII could allow the prompt diagnosis of GD at presentation, avoiding delay in management decisions (the start of antithyroid drug treatment or immediate radioiodine treatment), and obviating the need for other diagnostic procedures for the classification of aetiology of hyperthyroidism in thyrotoxic patients.

In view of the numerous clinical indications for TBII measurements and their potential expansion, in our opinion, the use of these second generation tests is highly recommended.

Take home messages

- The second generation of human thyrotrophin receptor assays have a high analytical and diagnostic accuracy, both with radioactive and chemiluminescent tracers, when both functional sensitivity and interlaboratory reproducibility are considered.
- At a threshold value of 1.99 IU/litre both tests had a sensitivity of 100% and a specificity of 99.1%.
- These two methods could be proposed as first line diagnostic markers for Graves’ disease and would be useful for following disease activity and treatment effects.
**Authors' affiliations**

D Villalta, Clinical Immunology and Virology Unit, City Hospital Pordenone, Via Montereale 24, 33170 Pordenone, Italy

N Bizzaro, University of Genoa, Viale Benedetto XV 6, 16132 Genoa, Italy

R Tozzoli, Laboratory, Medical and Radiometabolic Therapy Unit, DI.M.I. Autoimmunity Unit, City Hospital of Pordenone, Via Montereale 24, 33170 Pordenone, Italy

N. Sauro 25, 30027 San Donà di Piave, Italy

**REFERENCES**


**ECHO**

**Steatosis and viral genotype in chronic hepatitis C**

About 15% of people in the general population have liver steatosis. In patients with chronic hepatitis C there may be mild steatosis, which correlates with body mass index (BMI) and may be metabolic in origin, or moderate to severe steatosis that is more likely to be caused directly by the virus, especially genotype 3 virus. Steatosis may lead directly to progression of liver disease or it may add to the effect of other factors. Researchers in Switzerland and Italy have emphasised the importance of viral genotype.

The study included 755 patients with histologically and virologically proved chronic hepatitis C due to HCV genotypes 1–4 (178 with genotype 3, genotypes 5 and 6 excluded). Histology showed steatosis in 315 patients and fibrosis in 605 (187 with cirrhosis). On multivariate logistic regression analysis the factors independently associated with steatosis were fibrosis, viral genotype 3, BMI, current alcohol abuse, and age. Fibrosis was independently associated with liver disease activity score (Metavir score), age, steatosis, past alcohol abuse for more than five years, and BMI. In patients infected with virus of genotype 3 steatosis was associated with only Metavir score and fibrosis was associated with steatosis. In patients with HCV infection due to other genotypes steatosis was associated with continuing alcohol abuse and age, and fibrosis with past alcohol abuse and marginally with diabetes.

Among patients with chronic hepatitis C hepatic steatosis is associated with fibrosis only when the infection is with genotype 3 virus. Patients with steatosis and genotype 3 virus infection should be offered antiviral treatment irrespective of other clinical factors. When infection is with other genotypes of the virus abstinence from alcohol and weight control are also important.