

Analytical Performance and Clinical Utility of a Bioassay for Thyroid-Stimulating Immunoglobulins

Johannes J. Leschik,^{1*} Tanja Diana, MS,^{1*} Paul D. Olivo, MD, PhD,² Jochem König, PhD,³ Ulrike Krahn, MS,³ Yunsheng Li, MD, PhD,² Michael Kanitz,¹ and George J. Kahaly, MD, PhD¹

Key Words: Analytical performance; Dilution analysis; Thyroid-stimulating immunoglobulins

DOI: 10.1309/AJCPZUT7CNUEU7OP

Abstract

The analytical performance and the clinical utility of a thyrotropin receptor (TSHR)–stimulating immunoglobulin (TSI) bioassay were compared with those of a TSHR-binding inhibitory immunoglobulin (TBII) assay. Limits of detection (LoD) and quantitation (LoQ), assay cutoff, and the half-maximal effective concentration (EC₅₀) were measured. Dilution analysis was performed in sera of hyperthyroid patients with Graves disease (GD) during antithyroid treatment (ATD). Titer was defined as the first dilution step at which measurement of TSI or TBII fell below the assay cutoff. The LoD, LoQ, cutoff, and EC₅₀ of the bioassay were 251-, 298-, 814-, and 827-fold lower than for the TBII assay. There were 22%, 42%, 23%, and 14% more positive samples in the TSI bioassay at dilutions of 1:3, 1:9, 1:27, and 1:81 (P < .0001), respectively. Responders to ATD demonstrated marked differences in titers compared with nonresponders. The bioassay detected lower levels of TSHR autoantibodies, and the dilution analysis provided similar predictive values of both assays in GD.

Graves disease (GD) is characterized by production of autoantibodies to thyroid-associated antigens such as the thyrotropin receptor (TSHR).¹ The pathophysiology of the hyperthyroidism of GD is related to thyroid stimulatory autoantibodies (TSAb), also termed *thyroid-stimulating immunoglobulins* (TSI), that activate the TSHR on thyroid follicular cells, leading to unregulated thyroid hormone production.^{2,3} These functional autoantibodies mimic the receptor's natural ligand by stimulating cyclic adenosine monophosphate (cAMP)–dependent signal transduction, but other anti-TSHR antibodies antagonize the TSHR by either blocking TSH binding or interacting with TSHR epitopes that inhibit cAMP production.^{4,5}

Currently, 2 different methods are used to assess the level of antibodies directed against the TSHR. TSHR-binding inhibitory immunoglobulin (TBII) assays detect immunoglobulins that inhibit the binding of thyroid-stimulating hormone (TSH), or the human monoclonal stimulating autoantibody M22, to purified or recombinant TSHR. Commercially available TBII tests display high sensitivity and specificity for TSHR autoantibodies but unfortunately do not measure the functional activity of immunoglobulins and do not differentiate antibodies with stimulatory, blocking, or neutral activity.^{6,7} TSI bioassays, in contrast, are specific for TSHR-stimulating autoantibodies since they are based on measurement of cAMP production in a cell line stably transfected with the receptor.⁶ Recently, it has been reported that TSI show more significant association with clinical features of Graves eye disease than TBII.^{8,9} Furthermore, the TSI bioassay was compared with a TSHR-binding assay in a study in which sera from patients with GD were tested following serial dilution into normal serum.¹⁰ Certain sera, which were positive by both methods,

showed extinction of their binding activity between the 1:11 and 1:22 dilution but were still positive in the bioassay at a dilution of 1:300. These data suggest that the TSI assay can measure lower amounts of anti-TSHR antibodies in the sera of patients with GD.

To further evaluate the performance of the TSI reporter bioassay, we assessed autoantibody titers in hyperthyroid patients with GD enrolled in a prospective clinical trial of antithyroid drugs (ATD). We also compared the analytical performance of 2 commercially available anti-TSHR assays, a TSI bioassay and a TBII assay, using the thyroid-stimulating monoclonal autoantibody M22.¹¹

Materials and Methods

Patients

Forty hyperthyroid patients with GD (mean \pm SD age 44.8 ± 14 years; range, 21–76 years), of whom 31 (77.5%) were female and 16 (40%) were smokers, were prospectively followed for 36 weeks (24 weeks of treatment and subsequent 12 weeks of follow-up). Serum samples were collected both prior to as well as 4, 8, 12, 24, and 36 weeks after starting ATD treatment. GD was defined as the presence of hyperthyroidism (suppressed baseline serum TSH with increased serum-free thyroid hormone levels), positive TSHR autoantibodies, and a typical ultrasound imaging of a hypoechoic goiter. Methimazole therapy starting dose was 10 mg/d and 20 mg/d when serum-free thyroid hormone values were 3- or 5-fold increased, respectively. The dose of methimazole monotherapy was titrated every 4 weeks according to the serum-free thyroid hormone levels during the first 3 months and then at week 24 and, if required, at week 36. Response to therapy was defined as chemical euthyroidism (thyroid hormone levels in the normal range) at week 24 and persistent clinical and biochemical euthyroidism at week 36, 12 weeks after discontinuing ATD. Nonresponders to ATD were defined as hyperthyroid at week 24 and/or as having a biochemical relapse within the 12-week follow-up period. The trial was approved by the Ethical Committee of the Medical Association in the State of Rhineland-Palatinate, Germany, and all extensively informed patients gave their written consent prior to their inclusion in the study.

TSI and TBII Assays

TSI were measured with a US Food and Drug Administration (FDA)–cleared TSI reporter bioassay (Thyretain, Diagnostic Hybrids, Athens, OH), which was performed according to the manufacturer's instructions. The coefficient of variation (CV) of the triplicate measurement of a serum sample had to be below 15%. This TSI bioassay is based on Chinese hamster ovary cells (chimeric CHO-luc cells) that constitutively express a chimeric TSHR and contain a firefly

luciferase gene downstream of a promoter containing cAMP-responsive elements as described previously.¹⁰

TBII were measured using the automated electrochemiluminescence immunoassay (ECLIA) (Elecys Anti-TSHR Immunoassay, Roche Diagnostics, Mannheim, Germany) on the Cobas e411 analyzer (Roche Diagnostics). The measurements were performed according to the manufacturer's instructions. The ECLIA Elecys uses the MAb M22 labeled with a ruthenium complex. TBII are detected by inhibiting the binding of M22 to an immobilized porcine TSHR. The calibration is standardized against the National Institute for Biological Standards and Control (NIBSC) 90/672 standard.¹²

Serial Dilution Analysis

The TSI bioassay was compared with the TBII assay using 237 serum samples from 40 patients with GD. Serial 1:3 dilutions were performed on each serum sample (200 μ L) in triplicate into human TSI-negative control serum (400 μ L) up to a final dilution of 1:81. For samples still positive at 1:81, the dilution was extended to a final dilution of 1:243, 1:729, 1:2187, or 1:6561. The titer was defined as the first dilution step at which the assay results fell below the cutoff.

Analytical Performance

TSI Bioassay

Dose-response curve. M22, a stimulating monoclonal antibody (MAb) (Kronus, Star, ID; 4 μ g lyophilized), was dissolved in either 1 or 2 mL phosphate-buffered saline (PBS). Two-fold serial dilutions were prepared to achieve concentrations ranging from 0.003 to 100 ng/mL. Each M22 concentration was diluted in 1:11 control serum prior to testing in the TSI bioassay. All TSI values are expressed as percent specimen-to-reference ratio (SRR%), calculated with the following formula: $\text{SRR\%} = (\text{mean TSI}_{\text{specimen}} / \text{mean TSI}_{\text{reference}}) \times 100$. The mean $\text{TSI}_{\text{specimen}}$ is the mean of triplicate luminescence relative light unit (RLU) measurements containing M22, and the mean $\text{TSI}_{\text{reference}}$ is the mean of triplicate luminescence RLU containing 0.1 IU/mL bovine TSH. The bioassay cutoff is an SRR of 140%.

Limits of blank, detection, and quantitation. The limits of blank (LoB), detection (LoD), and quantitation (LoQ) were determined according to the Clinical and Laboratory Standards Institute (CLSI)–approved guideline for protocols determining LoD and LoQ.¹³ The LoD was measured using 8 low positive M22 concentrations ranging from 0.09 to 0.005 ng/mL. Each M22 concentration was measured 20 times in duplicate. The LoB was measured 40 times in duplicate with 1:11 diluted control serum. The LoQ was defined as the lowest M22 concentration that could be reliably repeated with an imprecision less than 20%. It could be at the same or a higher M22 concentration than the LoD in accordance with the following weighting: $\text{LoB} < \text{LoD} \leq \text{LoQ}$.

TBII Assay

Dose-response curve. M22 was diluted to prepare concentrations from 3.125 to 4000 ng/mL. All samples were measured in the Cobas e411 immunoassay analyzer (Roche Diagnostics) according to the manufacturer's instructions.

For the Elecsys Anti-TSHR Immunoassay (Roche Diagnostics; 510(k) number k080092), an LoB less than 0.5 IU/L, an LoD less than 0.8 IU/L, and an LoQ of 0.9 IU/L have been published (http://www.accessdata.fda.gov/cdrh_docs/reviews/K080092.pdf). The measurement range is 0.3 to 40 IU/L. Within-run precision testing gave CVs of less than 10%, less than 3%, and less than 4% for low (1.5-5 IU/L), moderate (5-20 IU/L), and high (20-40 IU/L) concentrations, respectively, whereas total precision testing showed CVs of less than 13%, less than 5%, and less than 6%, respectively.

Statistical Methods

With respect to the M22 dose-response curves for TSI, we report mean and 95% confidence interval (CIs) on the log 10 scale and, by raising these to the power of 10, the geometric mean with 95% CI for the metric scale in ng/mL. Each is based on 3 repeated runs, and calculations are based on means of 3 repeated measurements within each run. Calculations also are based on a Brain-Cousens¹⁴-type extension of the 5-parameter logistic model¹⁵ that allows for a slight decay of response beyond its maximum:

$$f(x, y_0, y_p, \beta, e, f, a) = y_0 + (y_1 - y_0) \frac{1 + ax}{(1 + e^{-\beta(x-e)})^f}$$

M22 dose-response curves for TBII were fit by first omitting all observations resulting in the upper measurement limit of 40 IU/L, which amounts to restriction to the dose range of 3.125 to 300 ng/mL. A 4-parameter logistic model for the log base 10 of dose (Morgan-Mercer-Flodin model) was fitted. We fitted the model to data of each of 5 runs, as well as determined the model parameters and concentrations corresponding to a response at 20 IU/L and 40 IU/L. The model-based half-maximal effective concentration (EC_{50}) parameter is not reported because the measurement range is cut off at 40 IU/L. The concentration at a response of 20 IU/L (which is midway between 0 and the maximum) is referred to as EC_{50} . For the determination of concentrations at an LoB of 0.5 IU/L, an LoD of 0.8 IU/L, an LoQ of 0.9 IU/L, and the reference limit of 1.75 IU/L, a linear fit to data for a dose ranging from 3.125 to 75 IU/L was used because it fits better over this restricted range. Positivity rates for each dilution were compared between TBII and TSI by counting the number of positive results per patient and dilution steps over all visits and comparing the resulting counts by the Student paired *t* test. Whenever they were adequate, Bonferroni-corrected test results are provided. The mean titers of TSI and TBII were compared over time by fitting a linear mixed model also including the interaction between assay method and time. We specified the

assay method as a random effect and the repeated measures for each patient and week with an unstructured covariance type, allowing the variance of TSI and TBII to be different. Sandwich estimators for the fixed-effects parameters and respective *t*-type tests and 95% confidence intervals (CI) were calculated. Least squares means of the titer within each assay method and week were computed (also separated by response) and are shown graphically with their standard errors. All *P* values are 2-sided. The analysis was performed using the statistics software programs SPSS version 19.0 (SPSS, Chicago, IL) and SAS version 9.2 (SAS Institute, Cary, NC).

Results

Analytical Performance

TSI

M22 stimulates cAMP production in chimeric CHO-luc cells in a dose-dependent manner (Figure 1A). The concentration at which the assay cutoff level of 140% occurs is 0.017 ng/mL (95% CI, 0.004-0.068). The lower asymptotic limit was a mean \pm SE SRR% of 111% \pm 9%, and the upper limit was 593% \pm 19%. Using these parameters, the EC_{50} was calculated to be 0.20 ng/mL (95% CI, 0.100-0.395).

The results of 80 blank measurements with 1:11 diluted control serum and 40 low M22 concentration samples showed a normal distribution (Shapiro-Wilk test). Therefore, LoB and LoD were calculated according to the following formulas: $LoB = \text{mean}_{\text{blank}} + 1.645 \times SD_{\text{mean}_{\text{blank}}}$, and $LoD = LoB + 1.645 \times \text{mean } SD_{\text{low concentration sample}}$. LoB had an SRR of 76.1%, whereby the average of 80 blank measurements had an SRR of 59.9% and an SD_{blank} of 9.8%. LoD had an SRR of 90.5% and an $SD_{\text{low concentration sample}}$ of 8.7%. LoD and LoQ were both determined to be 0.02 ng/mL M22.

TBII

Dose-response curves were performed for M22 in the TBII assay (Figure 1B). Measurement error (pooled over doses 3.125-300 ng/mL over 5 repeated runs) was 0.23 ng/mL (SD). Root mean square error (RMSE) for a linear model of log response fitted on doses ranging from 3.125 to 75 ng/mL was 0.21 ng/mL, which is similar to the overall measurement error SD. Mean \pm SE LoB corresponds to 2.24 \pm 1.18 ng/mL M22; LoD, 5.02 \pm 1.17 ng/mL M22; and LoQ, 5.95 \pm 1.16 ng/mL M22.

Over the unrestricted dose range from 3.125 to 300 ng/mL, a 4-parameter logistic model for the log base 10 of dose (Morgan-Mercer-Flodin model) fitted better (RMSE 0.29 IU/L for the pooled data) than a linear model (RMSE 0.52 IU/L). We fitted the model to data of each run and determined the model parameters and concentrations corresponding to a response at 20 IU/L and 40 IU/L. Means \pm SE at 20 IU/L and

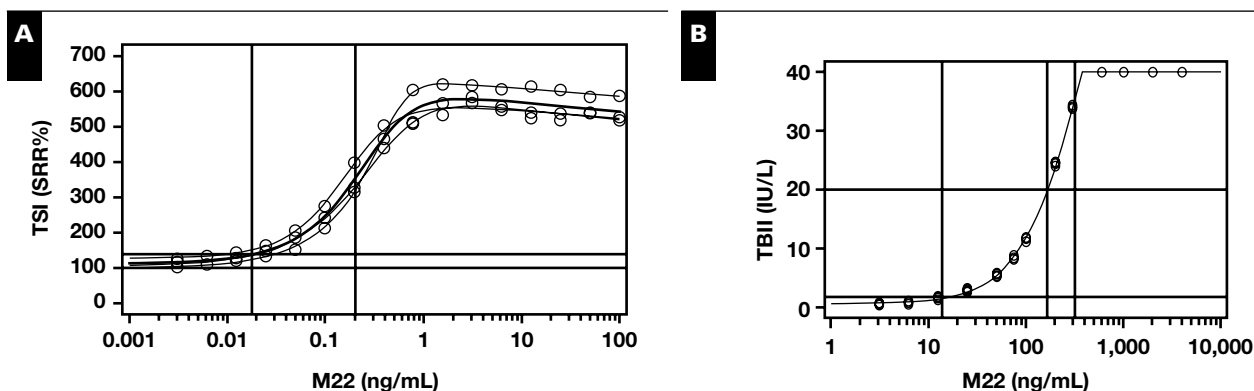


Figure 1 Dose response of a pure thyroid-stimulating human monoclonal antibody M22 in the thyroid-stimulating immunoglobulin (TSI) and thyrotropin receptor–binding inhibitory immunoglobulin (TBII) assays. **A**, Sigmoidal dose-response curve of the M22 autoantibody obtained with the chimeric TSI reporter bioassay. The M22 concentrations are plotted on the x-axis (logarithmic scale) against the corresponding percent specimen-to-reference ratio (SRR%) values on the y-axis. The dose response was measured 3 times, and the circles represent the averages of triplicate determinations of each experiment. The gray curves represent model fits for each of the 3 experiments, and the bold curve represents the model fit based on all 3 experiments. The cutoff value at an SRR of 140% and the reference level of 100% are marked by gray horizontal lines. The concentrations of the cutoff point and EC_{50} are indicated by the gray vertical lines. The model used for fitting is a 6-parameter extension of the usual 4-parameter logistic model, which allows for an asymmetric sigmoidal shape and a linear decay beyond the saturation point. **B**, Dose-response curve of the human M22 autoantibody obtained with the TBII assay. The circles present single measurements obtained in 5 runs. The M22 concentrations are plotted on the x-axis (logarithmic scale) against the corresponding IU/L values on the y-axis. A 4-parameter logistic model was fitted to the data of 5 runs where responses reaching the upper boundary value of 40 IU/L had been removed. The cutoff value at 1.75 IU/L and the response level of 20 IU/L and 40 IU/L are marked by gray horizontal lines, whereas the M22 concentrations of 13.8, 165, and 372 ng/mL, respectively, are indicated by the 3 vertical reference lines.

40 IU/L were 165.26 ± 1.26 ng/mL M22 and 371.96 ± 1.38 ng/mL M22, respectively (Figure 1B). The 5 fitted model curves highly agreed in the observed measurement range (up to 40 IU/L) but not beyond it. Unfortunately, the model-based definition of the EC_{50} depends on the upper limit of the fitted curve, because the abscissa corresponds to the mean between the upper and lower limits. The upper limits, obtained from the fitted curves, varied considerably, and hence the model-based EC_{50} varied. For this reason, we instead report values corresponding to the responses at 20 IU/L and 40 IU/L. The cutoff value of the TBII assay is 1.75 IU/L, which corresponded to a mean \pm SE M22 concentration of 13.83 ± 1.12 ng/mL.

Serial Dilution Analysis

A total of 237 serum samples were obtained from 40 thyrotoxic patients with GD prospectively followed at baseline and at weeks 4, 8, 12, 24, and 36 after ATD. Each serum sample was diluted and tested in both the TSI and TBII assays (Figure 2A). There were 22%, 42%, 23%, and 14% more positive samples in the TSI compared with the TBII assay at dilutions of 1:3, 1:9, 1:27, and 1:81, respectively, with a mean percent difference of 25% (95% CI, 17%-32%; $P < .0001$). At a dilution of 1:243 and higher, results were positive only with

the bioassay. The ratio of the positive TSI to TBII increased as the dilution increased such that the higher the dilution, the higher the positive TSI/TBII ratio (Figure 2B).

According to the manufacturer's instructions (Diagnostic Hybrids), all serum TSI values were measured in triplicate and the CV was obtained. Mean \pm SE CV values at dilutions of 1:9 and 1:81 were $5.4\% \pm 0.2\%$ (205 serum samples) and $6.2\% \pm 0.3\%$ (132 serum samples), respectively. In accordance with the manufacturer's instructions (Roche Diagnostics), we measured the TBII values once. Mean \pm SE TBII values were 3.4 ± 0.4 IU/L (160 serum samples) and 1.7 ± 0.4 IU/L (28 serum samples) at dilutions of 1:9 and 1:81, respectively. These dilution results are within the low concentration range (1.5-5 IU/L) defined by the manufacturer (Roche Diagnostics) that previously published within-run and total precision testing values of CVs less than 10% and less than 13%, respectively.

The correlation between all undiluted TSI values and the corresponding titers is shown in Figure 3A. A steady increase was noted until a titer of 3. At a TSI titer of 3 and above, the undiluted values reached a plateau. Baseline mean \pm SE TSI values at week 0 were similar in responders and nonresponders ($373\% \pm 25\%$ vs $377\% \pm 20\%$, respectively), whereas dilution analysis differentiated between the 2 groups.

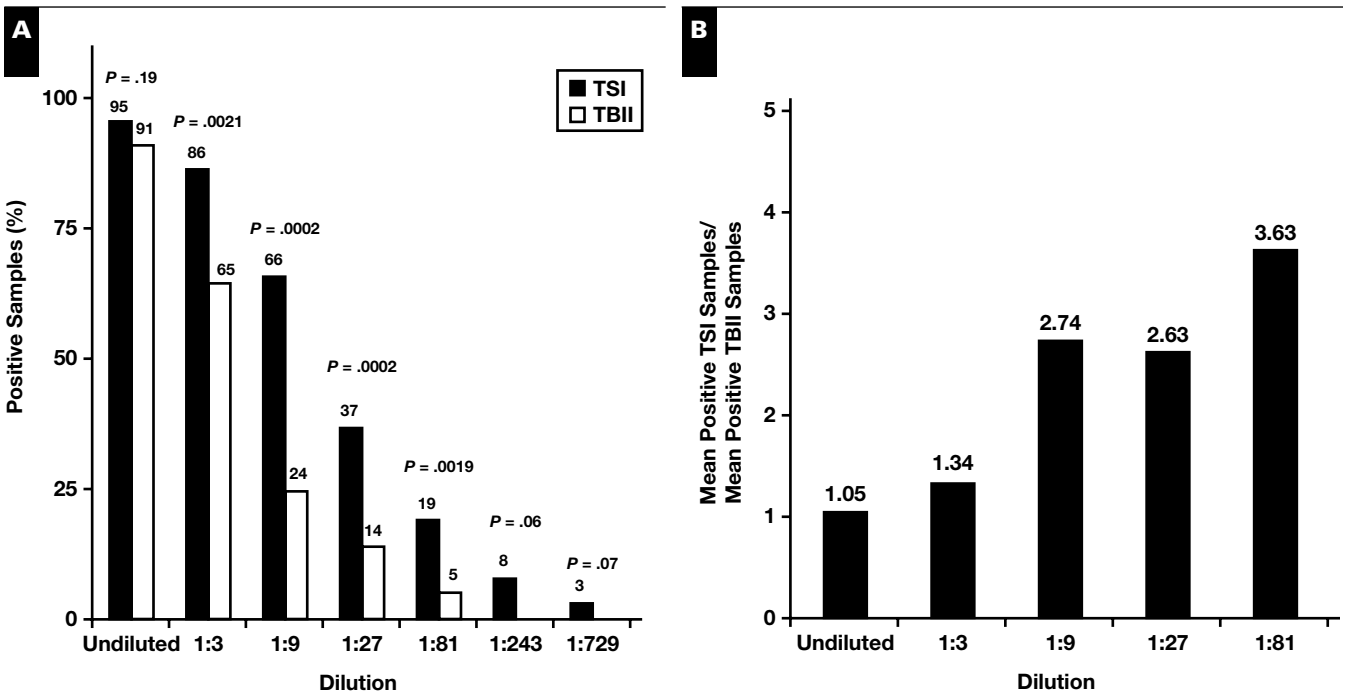


Figure 2 Comparison of thyroid-stimulating immunoglobulin (TSI) positivity vs thyrotropin receptor-binding inhibitory immunoglobulin (TBII) positivity as a function of dilution. **A**, Mean TSI (black) and TBII (white) percentage of positive samples for the indicated dilution at all time points during the 36-week observation period. Data were obtained from 228 serum samples from 38 patients with Graves disease, from whom sera were obtained at all time points during antithyroid drug treatment without exception after Bonferroni correction for multiple testing (7 tests). The difference between dilutions 1:3 through 1:81 is statistically significant at the .05 level. **B**, Ratio of the mean percentage of positive results for TSI over TBII at each dilution (see panel A). The ratios for results beyond dilution 1:81 are not shown because all TBII results were negative.

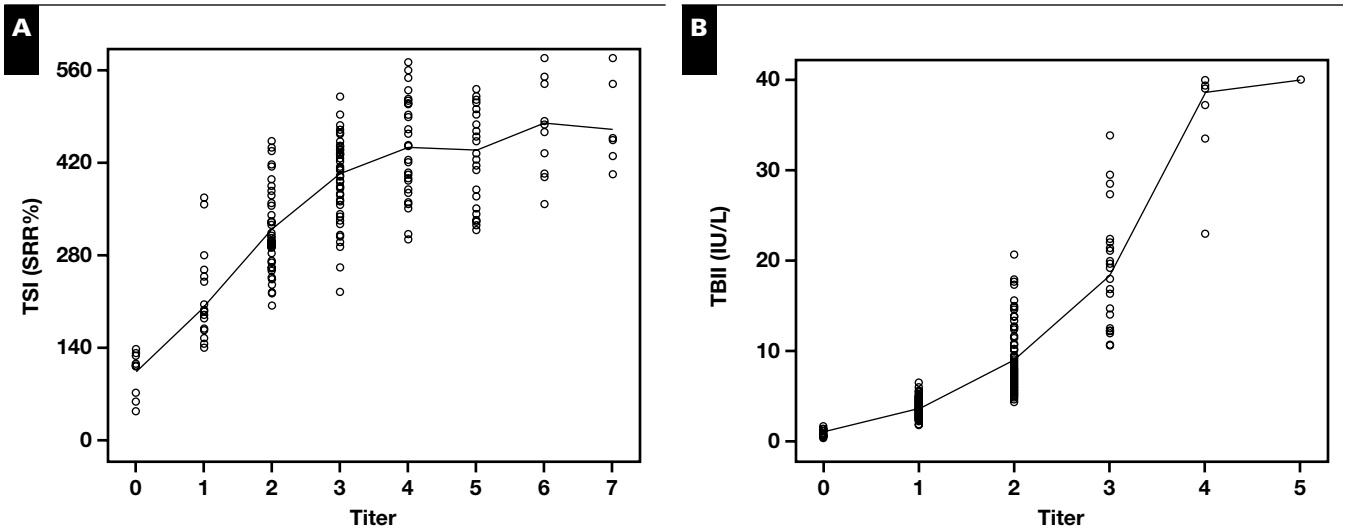


Figure 3 Undiluted baseline thyroid-stimulating immunoglobulin (TSI) and thyrotropin receptor-binding inhibitory immunoglobulin (TBII) results are plotted against the dilution titer. The circles in panels **A** and **B** represent the absolute values, and the graphs (lines) represent the mean values. The 2 figures can be regarded as clinical dose-response curves with the patient sera. **A**, Correlation of the undiluted TSI values in percent specimen-to-reference ratio (SRR%) (y-axis) and the corresponding dilution titer (x-axis). The dilution TSI titers ranged from 0 to 7. Up to a TSI titer of 3, there is a close correlation between the increase of undiluted values and the titer. **B**, Correlation of the undiluted TBII values in IU/L (y-axis) and the corresponding dilution titers (x-axis). TBII titers ranged from 0 to 5. The undiluted values increase with increasing titer up to a titer of 4 followed by a plateau at higher titers.

TSI titers of 3 or more were observed at least once in 33 of 40 (82.5%) patients and in 151 of 237 (64%) evaluated samples. In these samples, quantitative information was provided by dilution series but not by the undiluted samples (undiluted SRR% did not correlate with titer at that level). Once a “plateau” was reached above an SRR value of 400% corresponding to a TSI titer of 3, a clear differentiation of the SRR% values was possible with the help of the dilution analysis only. All nonresponders had a TSI titer of 3 and above.

In comparison, an exponential increase was noted for the TBII dilution curve reaching its plateau at a concentration of 40 IU/L [Figure 3B]. However, the plateau range of the bioassay was 3 dilution steps wider, and thus there were more values within the plateau. When analyzing titers as a function of time after ATD, both mean TSI and TBII titers demonstrated no change until week 12, followed by a linear decrease until week 36 [Figure 4]. In addition, the overall mean TSI titer was 1.13 dilution steps (95% CI, 0.84-1.42; $P < .0001$) higher than the overall mean TBII titer.

The serum thyroid-related hormone values (TSH, free T4, and free T3) of responders vs nonresponders at weeks 0 and 24 at completion of ATD treatment are shown in [Table 1]. During the 36-week observation period, a linear

Table 1
Thyroid-Related Hormones (Mean ± SE) Before and After ATD in 20 Responders vs Nonresponders at Baseline (Week 0) and After Completion of ATD Treatment (Week 24)

	Week	TSH	Free T3	Free T4
Normal range		0.4-4.9 mU/L	1.7-3.7 pg/mL	0.7-1.5 ng/dL
Responder (n = 20)	0	0.02 ± 0.01	7.33 ± 1.03	1.58 ± 0.13
	24	1.73 ± 0.36	2.82 ± 0.11	1.09 ± 0.06
Nonresponder (n = 20)	0	0.01 ± 0	10.71 ± 1.38	2.3 ± 0.17
	24	0.35 ± 0.15	5.39 ± 0.99	1.42 ± 0.13

ATD, antithyroid treatment; TSH, thyroid-stimulating hormone.

decrease of the TSI titer was noted in the ATD responder group (n = 20, 14 female, mean ± SD age 42.3 ± 14 years, 7 smokers), whereas an increase was registered at week 12 in the nonresponder group (n = 20, 17 female, mean ± SD age 47.3 ± 13 years, 9 smokers) [Figure 5]. In the nonresponders vs responders, baseline mean ± SE TSI titers were 4.0 ± 0.39 vs 2.9 ± 0.25 (difference 1.1; 95% CI, 0.19-2.01; $P = .018$), and mean ± SE TBII titers were 2.65 ± 0.29 vs 1.65 ± 0.16 (difference 1.0; 95% CI, 0.34-1.66; $P = .003$). The mean difference in TSI titers increased to 2.0 (95% CI, 1.23-2.77; $P <$

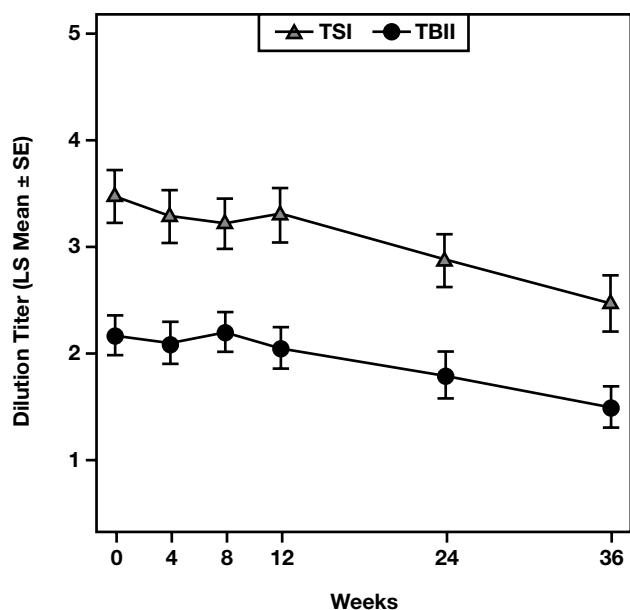


Figure 4 Mean thyroid-stimulating immunoglobulin (TSI) and thyrotropin receptor-binding inhibitory immunoglobulin (TBII) dilution titers of the 40 hyperthyroid patients with Graves disease during the 24 weeks of antithyroid drug treatment and the subsequent 12-week follow-up observation period. The mean TSI dilution titer is 1.13 (95% confidence interval, 0.84-1.42; $P < .0001$) points higher than the TBII titer at all time points. LS, least squares.

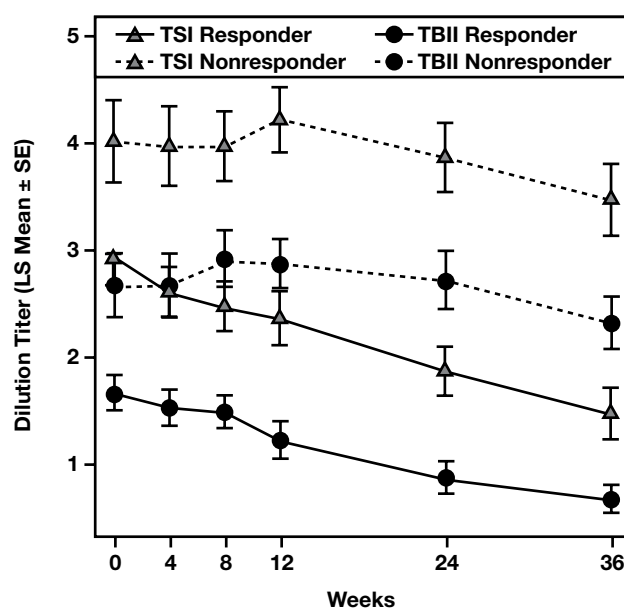


Figure 5 Mean thyroid-stimulating immunoglobulin (TSI) and thyrotropin receptor-binding inhibitory immunoglobulin (TBII) dilution titers of the 20 responders and 20 nonresponders with Graves disease during the 24 weeks of antithyroid drug treatment (ATD) and the subsequent 12-week follow-up observation period. In responders to ATD, a decrease of the TSI and TBII titers is noted already from the beginning of ATD treatment in contrast to the titer course in the nonresponders. LS, least squares.

.0001) at week 24, whereas it was 1.85 (95% CI, 1.24-2.46; $P < .0001$) for the TBII assay.

Discussion

For the first time, this prospective study evaluates the analytical performance as well as the clinical utility and the predictive value of serial dilution analysis in patients with autoimmune Graves hyperthyroidism undergoing antithyroid drug treatment. In line with our previous report¹⁶ this is a further comparative study of an FDA-cleared chimeric TSHR TSI reporter cell-based bioassay and a widely distributed automated TSHR binding (TBII) assay. Compared with the TBII assay, the bioassay is able to detect lower levels of anti-TSHR autoantibodies and has a similar predictive value in the dilution analysis with serum of patients with GD.

Analytical Performance

Compared with the TBII assay, the TSI bioassay can detect lower amounts of a stimulating monoclonal autoantibody as indicated by the lower LoD and LoQ, assay cutoff, and half-maximal effective concentration. Interestingly, the large difference between the 2 assays in LoQ and LoD as measured using M22 was not reflected in the dilution analysis with patient sera. The TSI bioassay only detects stimulating autoantibodies that activate the chimeric TSHR, resulting in rapid signal transduction and leading to induction of cAMP and luciferase expression.¹⁷ Serum from a patient with GD can encompass a mixture of various antibody types with differences in both structure as well as function.¹⁸ The third-generation TBII assay (ECLIA Elecsys, Roche Diagnostics) uses a solubilized porcine TSHR and measures competitive binding activity of patient serum in the presence of a competitive agonist, M22 labeled with a ruthenium complex.¹⁹ The readout is an electrochemiluminescent light emission that is represented in IU/L based on a standard reference material.²⁰ Stimulating, neutral, and blocking immunoglobulins bind to conformational epitopes of the TSHR within the TSH binding pocket.²¹ The assay measures the binding activity of all anti-TSHR antibodies and does not discriminate among the different functions of antibodies. The inherent differences in how each assay detects anti-TSHR antibodies might explain why the TSI bioassay exhibited greater apparent responsiveness to M22 compared with polyclonal patient sera. M22, isolated from lymphocytes of a hyperthyroid patient with GD, exhibits a powerful stimulating activity *in vivo* as well as *in vitro*.²² The antigen binding fragment (Fab) of M22 binds with high affinity to the TSH receptor^{23,24} and mimics properly the binding of thyrotropin. Furthermore, it is possible that blocking antibodies in the polyclonal sera of patients with GD bind to the TSHR and interfere with the measurement of stimulating antibodies by blocking the response of the chimeric CHO-luc cells.

Dilution Analysis and Clinical Relevance

In this original study, serial dilution analysis proved to be clinically useful in 64% of the prospectively evaluated serum samples, offering additional quantitative information to the baseline undiluted TSI values. In hyperthyroid patients with TSI levels above an SRR of 400%, determining the TSI titer further defines the level of TSI activity in the serum of these patients. Also, during the whole observation period, the differences of TSI titers between responders to ATD and nonresponders were slightly greater than the corresponding differences of TBII titers. Therefore, serial dilution analysis seems to be of clinical utility for treating physicians and in laboratories using the TSI bioassay during the management of thyrotoxic Graves patients. This could be acknowledged as an added value of the TSI bioassay vs its more complex methodology when proposed for its routine use as opposed to the TBII assay.

The TSI dilution titer of the nonresponder is foremost at 4 and above. This usually corresponds to undiluted values of an SRR of around 400% or higher. However, an SRR value of 400% is not a fixed line (ie, it might vary between 360% and 440%). This undiluted SRR range can also be assumed with a TSI titer of 3. Since baseline mean \pm SE TSI values at week 0 were similar in responders and nonresponders, the dilution analysis (ie, TSI titers) provided an enhanced and statistically significant quantitative discrimination between responders and nonresponders. Thus, dilution analysis markedly improved the clinical relevance and predictive role of the bioassay and may be regarded as an additional informative diagnostic tool to evaluate the serological activity of GD.

No data comparable to those from this prospective design have been reported in the literature. Previous publications investigated the value of the level of anti-TSHR antibodies for predicting the response to ATD treatment,²⁵⁻²⁸ but the present report represents the first evaluation of the usefulness of serial dilution for predicting the outcome of therapy for patients with GD. Serum dilution has been performed *in vitro* to better characterize experiment results, but none of these studies drew clinically relevant conclusions.²⁹⁻³² For example, the effects of graded dilutions of sera and IgG with known TSI activity were compared.³¹ Sera as well as IgG increased the cAMP production, but at the highest concentrations, an inhibitory effect was evident. In contrast, the serial dilutions performed in the present study were informative and clinically useful in more than 80% of the prospectively evaluated patients with GD, especially when looking at TSI values outside the linear range. Furthermore, both the baseline values as well as TSI titers during the course of therapy were clearly different in the responders to ATD medical treatment vs those who did not respond.

Considerations for Clinical Use of Both Assays in the Laboratory

Figure 3 demonstrates that the slope of the TSI curve (Figure 3A) is steeper in the range around the cutoff compared with the TBII curve (Figure 3B), which indicates that the quantitative discrimination of the bioassay is more sensitive at that level. Therefore, the bioassay might be recommended in cases where a low autoantibody level is expected. This applies to patients with mild and/or subclinical GD or with Graves hyperthyroidism of recent onset. Furthermore, and because of the risk of neonatal thyroid dysfunction in the newborn, accurate differentiation of the functional character of TSHR autoantibodies is keenly warranted during pregnancy as well as in the postpartum period. Finally, and as previously described, the close correlation of the TSI levels with the clinical activity and severity of Graves thyroid eye disease might offer additional relevant information compared with the serum TBII levels.^{8,9} On the other hand, the easier to handle binding assay measures TSHR autoantibodies at higher levels and at a broader concentration range showed by the M22 results. Thus, the TBII assay can be regarded as a useful tool for follow-up during and after specific antithyroid treatment.

From the ¹Molecular Thyroid Research Laboratory, Department of Medicine I, Johannes Gutenberg University Medical Center, Mainz, Germany; ²Diagnostic Hybrids, a Quidel Company, Athens, OH; and ³Institute of Medical Biostatistics, Epidemiology and Informatics (IMBEI), Division Medical Biometry, Johannes Gutenberg University Medical Center, Mainz, Germany.

*Mr Leschik and Ms Diana contributed equally to this article.

Address reprint requests to Dr Kahaly: Dept of Medicine I, Gutenberg University Medical Center, Mainz 55101, Germany; Kahaly@ukmainz.de.

Dr Olivo and Dr Li are Quidel employees. Dr Kahaly consults for and has received research grants from Quidel.

Acknowledgments: We are grateful to Elisa Kolbe, study nurse; Sabine Niel, laboratory scientific associate; and Nina Matheis, MS, all at the Molecular Thyroid Research Laboratory, Gutenberg University Medical Center, Mainz, Germany, for data collection and fruitful discussion.

References

- Brent GA. Graves' disease. *N Engl J Med*. 2008;358:2594-2605.
- Rose NR. Predictors of autoimmune disease: autoantibodies and beyond. *Autoimmunity*. 2008;41:419-428.
- Mizutori Y, Chen CR, Latrofa F, et al. Evidence that shed thyrotropin receptor A subunits drive affinity maturation of autoantibodies causing Graves' disease. *J Clin Endocrinol Metab*. 2009;94:927-935.
- Davies TF, Yin X, Latif R. The genetics of the thyroid stimulating hormone receptor: history and relevance. *Thyroid*. 2010;20:727-736.
- Zöphel K, Roggenbuck D, Schott M. Clinical review about TRAb assay's history. *Autoimmun Rev*. 2010;9:695-700.
- Lytton SD, Kahaly GJ. Bioassays for TSH-receptor autoantibodies: an update. *Autoimmun Rev*. 2010;10:116-122.
- Morris JC III, Hay ID, Nelson RE, et al. Clinical utility of thyrotropin-receptor antibody assays: comparison of radioreceptor and bioassay methods. *Mayo Clin Proc*. 1988;63:707-717.
- Ponto KA, Kanitz M, Olivo PD, et al. Clinical relevance of thyroid-stimulating immunoglobulins in Graves' ophthalmopathy. *Ophthalmology*. 2011;118:2279-2285.
- Lytton SD, Ponto KA, Kanitz M, et al. A novel thyroid stimulating immunoglobulin bioassay is a functional indicator of activity and severity of Graves' orbitopathy. *J Clin Endocrinol Metab*. 2010;95:2123-2131.
- Lytton SD, Li Y, Olivo PD, et al. Novel chimeric thyroid-stimulating hormone-receptor bioassay for thyroid-stimulating immunoglobulins. *Clin Exp Immunol*. 2010;162:438-446.
- Sanders J, Evans M, Premawardhana LD, et al. Human monoclonal thyroid stimulating autoantibody. *Lancet*. 2003;362:126-128.
- Schott M, Hermsen D, Broecker-Preuss M, et al. Clinical value of the first automated TSH receptor auto-antibody assay for the diagnosis of Graves' disease (GD): an international multicentre trial. *Clin Endocrinol (Oxf)*. 2009;71:566-573.
- Tholen DW, Linnet K, Kondratovich M, et al. *Protocols for Determination of Limits of Detection and Limits of Quantitation*. Approved Guideline. Wayne, PA: National Committee for Clinical Laboratory Standards; 2004. NCCLS EP17-A.
- Brain P, Cousens R. An equation to describe dose responses where there is stimulation of growth at low dose. *Weed Res*. 1989;29:93-96.
- Finney DJ. Bioassay and the practice of statistical inference. *Int Stat Rev*. 1979;47:1-12.
- Kamijo K, Murayama H, Uzu T, et al. Similar clinical performance of a novel chimeric thyroid-stimulating hormone receptor bioassay and an automated thyroid-stimulating hormone receptor binding assay in Graves' disease. *Thyroid*. 2011;21:1295-1299.
- Kamijo K, Murayama H, Uzu T, et al. A novel bioreporter assay for thyrotropin receptor antibodies using a chimeric thyrotropin receptor (Mc4) is more useful in differentiation of Graves' disease from painless thyroiditis than conventional thyrotropin-stimulating antibody assay using porcine thyroid cells. *Thyroid*. 2010;20:851-856.
- Evans M, Sanders J, Tagami T, et al. Monoclonal autoantibodies to the TSH receptor, one with stimulating activity and one with blocking activity, obtained from the same blood sample. *Clin Endocrinol (Oxf)*. 2010;73:404-412.
- Yoshimura Noh J, Miyazaki N, Ito K, et al. Evaluation of a new rapid and fully automated electrochemiluminescence immunoassay for thyrotropin receptor autoantibodies. *Thyroid*. 2008;18:1157-1164.
- Hermsen D, Eckstein A, Schinner S, et al. Reproducibility of Elecsys anti-TSHR test results in a lot-to-lot comparison. *Horm Metab Res*. 2010;42:295-297.
- Latif R, Morshed SA, Zaidi M, et al. The thyroid-stimulating hormone receptor: impact of thyroid-stimulating hormone and thyroid-stimulating hormone receptor antibodies on multimerization, cleavage, and signaling. *Endocrinol Metab Clin North Am*. 2009;38:319-341.

22. Núñez Miguel R, Sanders J, Chirgadze DY, et al. Thyroid stimulating autoantibody M22 mimics TSH binding to the TSH receptor leucine rich domain: a comparative structural study of protein-protein interactions. *J Mol Endocrinol*. 2009;42:381-395.
23. Sanders J, Jeffreys J, Depraetere H, et al. Characteristics of a human monoclonal autoantibody to the thyrotropin receptor: sequence structure and function. *Thyroid*. 2004;14:560-570.
24. Rees Smith B, Sanders J, Evans M, et al. TSH receptor-autoantibody interactions. *Horm Metab Res*. 2009;41:448-455.
25. Schott M, Minich WB, Willenberg HS, et al. Relevance of TSH receptor stimulating and blocking autoantibody measurement for the prediction of relapse in Graves' disease. *Horm Metab Res*. 2005;37:741-744.
26. Schott M, Minich WB, Willenberg HS, et al. Levels of autoantibodies against human TSH receptor predict relapse of hyperthyroidism in Graves' disease. *Horm Metab Res*. 2004;36:92-96.
27. Wallaschofski H, Miehle K, Mayer A, et al. Prediction of remission or relapse for Graves' hyperthyroidism by the combined determination of stimulating, blocking and binding TSH-receptor antibodies after the withdrawal of antithyroid drug treatment. *Horm Metab Res*. 2002;34:383-388.
28. Michelangeli V, Poon C, Taft J, et al. The prognostic value of thyrotropin receptor antibody measurement in the early stages of treatment of Graves' disease with antithyroid drugs. *Thyroid*. 1998;8:119-124.
29. Michelangeli VP, Munro DS, Poon CW, et al. Measurement of thyroid stimulating immunoglobulins in a new cell line transfected with a functional human TSH receptor (JPO9 cells), compared with an assay using FRTL-5 cells. *Clin Endocrinol (Oxf)*. 1994;40:645-652.
30. Michelangeli VP, Poon CW, Arnus EE, et al. Measurement of TSH receptor blocking immunoglobulins using ³H-adenine incorporation into FRTL-5 and JPO9 cells: use in a child with neonatal hypothyroidism. *Clin Endocrinol (Oxf)*. 1995;42:39-44.
31. Vitti P, Chiovato L, Lopez G, et al. Measurement of TSAbs directly in serum using FRTL-5 cells. *J Endocrinol Invest*. 1988;11:313-317.
32. Wallaschofski H, Paschke R. Detection of thyroid stimulating (TSAB)- and thyrotropin stimulation blocking (TSBAB) antibodies with CHO cell lines expressing different TSH-receptor numbers. *Clin Endocrinol (Oxf)*. 1999;50:365-372.