Predicting the Relapse of Hyperthyroidism in Treated Graves' Disease with Orbitopathy by Serial Measurements of TSH-Receptor Autoantibodies

Authors

Mareile Stöhr[1], Michael Oeverhaus[1], Simon D. Lytton[2], Mareike Horstmann[1], Denise Zwanziger[3], Lars Möller[3], Achim Stark[4], Dagmar Führer-Sakel[3], Nikolaos Bechrakis[1], Utta Berchner-Pfannschmidt[3], J. Paul Banga[5], Svenja Philipp[3], Anja Eckstein[1]

Affiliations

1 Department of Ophthalmology, University Hospital Essen, Essen, Germany
2 SeraDiaLogistics, Munich, Germany
3 Department of Endocrinology, Diabetes and Metabolism, University of Duisburg-Essen, Essen, Germany
4 Practice for General Medicine, Essen, Germany
5 Emeritus, King's College London, London, UK

Key words

binding assays, bioassays, TSH receptor antibodies, TRAb, Graves' disease, Graves' orbitopathy

received 03.09.2020
accepted after revision 18.01.2021
published online 22.02.2021

Bibliography

Horm Metab Res 2021; 53: 235–244
DOI 10.1055/a-1373-5523
ISSN 0018-5043

Georg Thieme Verlag KG, Rüdigerstraße 14,
70469 Stuttgart, Germany

Correspondence

Dr. med. M. Stöhr
Department of Ophthalmology
University Hospital Essen
Hufelandstr. 55
45147 Essen
Germany
Tel.: +49 2017232976, Fax: +49 2017236755
Mareile.stoehr@uk-essen.de

ABSTRACT

The aim of this study was to investigate the potential of the new TSH-receptor antibody (TRAb) assays to predict remission or relapse of hyperthyroidism in patients with Graves' disease (GD) and Graves' orbitopathy (GO). TRAbs were measured retrospectively in sera from a cohort of GD patients with GO (n = 117; remission n = 38 and relapse n = 79–Essen GO biobank) with automated binding immunoassays: TRAb Elecsys (Cobas Roche) and TRAb bridge assay (IMMULITE, Siemens), and the TSAb (thyroid stimulating Ab) cell-based bioassay (Thyretain, Quidel Corp.). To identify relapse risk/remission of hyperthyroidism patients were followed up at least 10 months after the end of antithyroid drug therapy (ATD) therapy. ROC plot analysis was performed to calculate cut-off levels of TRAb and TSAb for prediction of relapse and remission of hyperthyroidism. Cut-off serum levels are provided for timepoints around 3, 6, 10, and 15 months after the beginning of ATD. Repeated measurements of TRAb increase the rate of relapses predictions to 60 % (Elecsys), 70 % (IMMULITE), and 55 % (Thyretain). Patients with remission have consistently TRAb levels below the cut off for relapse in repeated measurements. The cell-based bioassay was the most sensitive – and continued to be positive during follow up [at 15 months: 90 % vs. 70 % (IMMULITE) and 65 % (Elecsys)]. Identification of relapsing hyperthyroidism is possible with automated immunoassays and cell-based bioassay especially with serial TRAb measurements during the course of ATD therapy. Patient who need eye surgery may profit from an early decision towards definitive treatment.

Introduction

Autoimmune thyroid diseases, comprising of Graves' disease (GD) and destructive thyroiditis [Hashimito's disease (HD)] are common autoimmune disorders. GD is characterized by a diffuse lymphocytic infiltrate in the thyroid gland with the presence of autoantibodies to thyroid antigens with those directed to the TSH-receptor (TSHR) being pathognomonic. The pathogenic antibodies to TSHR in GD are termed TRAbs and comprise of different subpopulations including stimulating antibodies, which lead to development of uncontrolled hyperthyroidism. A proportion of GD patients also develop orbitopathy, which has been linked to TSHR expression in orbital fibroblasts.

Patients with newly diagnosed Graves' hyperthyroidism are usually medically treated for 12–18 months with antithyroid drugs [1]. Approximately 50 % of patients treated with antithyroid drug therapy relapse when treatment is withdrawn after achieving euthyroidism during their treatment period.

Numerous studies have been performed to identify markers to predict relapse/remission when treatment is withdrawn. Associations to the course of hyperthyroidism were described to TRAb levels, smoking, male gender, large goitre, hypoechogenic and hypervascular gland, and severity of hyperthyroidism [2–7].

Antibody measurements using variable cut offs have been proven to be useful in comparison to use the diagnostic cut offs [8]. The significantly improved diagnostic sensitivity and specificity of second generation assays allowed the calculation of variable cut offs for different time points during the course of antithyroid drug therapy (ATD).

As early as 6 months after the beginning of ATD therapy predictions were possible. With a TRAb value of 10 IU/l or more measured 6 months after beginning of ATD therapy almost all patients (97 %) developed relapse, whereas only 3 % went into remission (positive predictive value [PPV] of 96.4 %). Below that cut-off, no outcome prediction could be made since low TRAb levels were associated with remission or relapse [3]. Other studies confirmed this observation and defined additional cut offs for other timepoints. Quadbeck et al. defined a TRAb cut-off value of 10 IU/l to distinguish relapse from remission (PPV 83 %) 12 months after initiating ATD therapy [4]. Eckstein et al. defined a slightly lower threshold of 7.5 IU/l at 12 months (PPV 97 %) [6]. Finally, Carella et al. (2006) defined the cut-off (3.85 IU/l) at 18 months after initiating ATD therapy (PPV 96.7 %) to predict upcoming relapse vs. remission [5]. In parallel to the decline of serum TRAb with increasing duration of ATD therapy the cut off values decreased as well over time [9].

Combined presence of high serum TRAb titer (> 6 IU/l) and high anti-TPO-antibody titer (> 5.000 IU/l) at about 4 months after the beginning of ATD therapy lead to a 100 % PPV for disease relapse [10].

According to the strong association of TRAb levels with the prognosis of hyperthyroidism and GO, severe GO itself is also associated with a significantly lower remission rate of hyperthyroidism (8 %) [6].

In the meantime, third generation automated binding immunoassays have recently become available which lead to further improvements in TRAb measurement [11–14]. Additionally two new assay systems are available for clinical use. One commercially available bioassay (Thyretain) detects the stimulatory immunoglobulin fraction [15]. The extended cut off levels of signal to reference ratios of Thyretain bioassay have proven suitable for the prediction of relapse. In addition, the authors show that TSAb levels decrease significantly in the GO remission group [17]. However, up until now, the bioassay applications in routine clinical diagnostic is limited by cell culture requirements and labor intensiveness compared with automated binding immunoassays.

The other new automated technology of TRAb chemiluminescent bridge immunoassay (Siemens, IMMULITE) was introduced 2015. It has been shown that the assay performs comparably to other third generation assays in prediction models which use clinical and laboratory parameters before the start of ATD [18].

The aim of this study was to investigate and compare the potential of the different assay techniques (binding, bridge and bioassay) to predict remission or relapse of hyperthyroidism during the course of ATD therapy. In particular, the objective was to also to determine whether serial TRAb measurements have additional diagnostic impact and if the bioassay technology was superior to conventional automated binding assays.

Materials and Methods

Patients

This is a retrospective study with serum samples from the University Duisburg-Essen Biobank for patients with Graves' disease (founded in 2000) and data bank from our joint thyroid eye clinic. The analysis of TRAbs using third generation assay by Roche started in 2008 and patients were consecutively included in this clinical trial since then up to 2017.

The recruitment and research protocols were reviewed and approved by the Institutional ethics commission (06–3211 and 14–5965-BO), and written informed consent was obtained from all study participants in compliance with the Declaration of Helsinki.

Classification according the course of Graves' hyperthyroidism

Patients were included in the database if they presented in the eye clinic within the first six months of the onset of eye or thyroid disease. The mean duration of the onset of eye disease was 3.83 months. Only patients who have been treated with antithyroid drugs (ATD) for at least 10 months and have been followed up for at least 10 months or longer after the end of therapy were eligible. Patients were considered to be in remission if at the last visit their FT4/TSH level were within the normal range. Patients were considered to have relapse if their FT4/TSH level run out of the normal range towards hyperthyroidism within 12 months after cessation of ATD. They were also called relapse if hyperthyroidism repeatedly reoccurred with dose reduction of antithyroid drug treatment or if ATD could not be stopped. A few of these patients subsequently underwent definitive treatment (thyroidectomy) in the observation period. From the timepoint of definitive treatment they were not included in the ROC analysis anymore.

The patients were not explicitly ordered continuously. To establish a temporal reference, the examination visits were summarized to specific time periods (TP): TP 1 = 0–4 months; TP 2 = 5–8 months; TP 3 = 9–12 months; TP 4 = 13–16 after beginning of ATD therapy. Patient data and sera were not available from all patients at all timepoints.

Classification according the course of Graves' orbitopathy

Grading of Graves' orbitopathy was assessed using NOSPECS score [no signs or symptoms (N); only sign (lid retraction), no other symptoms (O); soft tissue involvement (S); proptosis (P); eye muscle involvement (E); corneal involvement (C); and sight loss due to optic nerve compression (S)]. We discriminated between mild GO (NOSPECS < 5) and severe GO (NOSPECS ≥ 5). Clinical examination and classification were carried out by one ophthalmologist with experience in examining GO patients.

Assay technologies

The Elecsys Anti-TSH Receptor (TRAb) electro-chemiluminescent immunoassay (Roche Diagnostics, Mannheim, Germany) is the assay which is used for the clinical routine at the University Duisburg-Essen. The TRAb assay by Roche is a competitive binding immunoassay, also named TSH binding inhibition immunoglobulin (TBII) assay, using the M22 monoclonal antibody which binds TSHR with high affinity. Therefore, the assay determines the competition for binding to TSHR between the TRAbs and the antigen binding fragments (Fab) of antibodies labelled with ruthenium [19]. Hence, TRAb levels are assessed indirectly via the quantification of the bound antibodies. The measured luminescence signal is inversely proportional to the TRAb-Level. The assay does not distinguish between blocking and stimulating immunoglobulins [20]. Here, the highest sensitivity (99 %) and specificity (99 %) values to diagnose GD could be calculated at a cut-off level of 1.75 IU/l [11, 21].

The IMMULITE 2000 TSI Assay ranks among the automated two-step chemiluminescent immunoassays containing a pair of recombinant human TSH receptors used in a bridging format. It consists of a bridge link between TRAbs and the two receptors (capture and signal antibodies) [22]. The capture receptor is characterized by a THSR chimera of the rat luteinizing hormone/chorionic gonadotropin receptor (LH/CG). That implies that the receptor has epitopes, which should be recognized only by TSAb [22], which is however controversial. The binding receptor is immobilized on a solid phase. TSI binds to these during a 30-min incubation period. After centrifugal purification of the supernatant, in the second step the signal receptors are added for 30 min, which bind to the second arm of the complexed TSI. These receptors are labelled with alkaline phosphatase, thereby converting an added chemiluminescent substrate in the last step. The generated light signal is detected with the result that the quantity of TSI is directly correlated to the assessed chemiluminescent signal. A new automated assay for the detection of stimulating TRAb with this bridge technology has been developed. At a cut-off level of 0.55 IU/l the highest sensitivity (100 %) and specificity (about 99 %) were seen for diagnosing GD [13, 23].

The bioactivity of TSH-receptor stimulating antibodies (TSAb) were assessed with the cell-based bridging bioassay Thyretain (Quidel Corp., San Diego, CA, USA) according the manufacturer's instructions [15]. The assay uses Chinese hamster ovary (CHO) cells expressing a chimeric TSH-R (MC4) and cAMP-dependent luciferase expression. Percentages of specimen-to-reference ratio (SRR %) values were calculated according the following formula:

$$\text{SRR\%} = \frac{\text{average TSAb specimen relative light units}}{\text{average reference standard relative light units}} \times 100$$

All three assays were performed following the manufacturer's instructions.

Statistical analysis

Data were collected by using Microsoft Office Excel (Version: 2016). For metric data, median values and range or the mean and standard deviation (SD ±) were calculated and differences were evaluated with Student's *t*-test (two-tailed) if D'Agostino–Pearson omnibus normality test showed normal distribution, if not, with Mann–Whitney-test. Fisher's exact test and the Chi^2- test were used to evaluate group distributions of binary variables (e. g., gender, smoking, inactivation and prior anti-inflammtory therapy). Additionally, analysis with ANOVA test was performed to compare the different assays. Linear regression was performed to analyze the correlation between the different assays. p-Values equal to or less than 5 % (≤ 0.05) were considered as significant. All calculations were performed with GraphPad Prism (Prism 6 for Windows, Software Inc., San Diego, CA, USA, Version 6.01). Receiver operator curve (ROC) analysis was performed and analyzed to select the best cut-off with highest clinical sensitivity and specificity. A two-tailed p-value < 0.05 was considered to be statistically significant.

Results

Complete patient data for the assessment of a remission or a relapse of the thyroid hyperfunction according to the criteria mentioned in the method section were available for 117 patients from the database. ▶ **Table 1** summarizes key characteristics of the patient cohort.

Patient characteristics

The mean age at the time of first manifestation of thyroid disease was 45.5 years in the remission group and 50 years in the relapse group ($p < 0.0015$). In the remission group, 7.9 % of the patients had severe GO, in the relapse group 79.7 % of the patients ($p < 0.001$). The group with a relapse of hyperthyroidism contained significantly more patients with a severe course of Graves' orbitopathy. Accordingly, patients in this group also received significantly more often anti-inflammatory treatment or surgical rehabilitative measures. A sum of 55.2 % of the patients in the remission group and 82.3 % of the patients in the relapse group were treated with steroids for anti-inflammatory use ($p < 0.001$), 23.7 % of the patients in remission and 54.4 % of the relapse patients received an orbital radiation ($p < 0.001$). Surgical intervention was required in 21.1 % remission patients and in 34.2 % relapse patients ($p < 0.001$), and decompression was required in 7.9 % patients in the remission group and in 34.2 % patients in the relapse group ($p < 0.001$).

▶ **Table 1** Patient characteristics.

	Remission	Relapse	p-Value
Number	38 (32 %)	79 (68 %)	n = 117
Age	45.5 (Min 25, Max 59)	50 (Min 20, Max 72)	p < 0.0015
Smoking	26.2 %	34.2 %	nonsmoker
Severe GO	7.9 %	79.7 %	p < 0.001
Mild GO	92.1 %	20.3 %	p < 0.001
Steroid therapy	55.2 %	82.3 %	p < 0.001
Orbital radiation	23.7	54.4 %	p < 0.001
Eye surgery	21.1 %	53.2 %	p < 0.001
Decompression	7.9 %	34.2 %	p < 0.001

Correlation of measurements of TRAb levels with the three different assays

The closest correlation was found between TRAb Elecsys binding assay and IMMULITE bridge assay ($r = 0.88$; $p < 0.0001$). Both binding assays correlated significantly to the bioassay but to a minor extent. Correlation between TSI Thyretain bioassay and IMMULITE bridge assay was $r = 0.61$ ($p < 0.0001$). The correlation between TSI Thyretain bioassay and TRAb Elecsys binding assay was $r = 0.62$ ($p < 0.0001$) (data not shown).

Levels of antibodies during the course of GD/GO

All TRAb measurements were performed at 4 timepoints (TP). TP 1 was defined 0–4 months after the beginning of ATD therapy; TP 2: 5–8 months; TP 3: 9–12 months and TP 4: 13–16 months. Using the commercially provided diagnostic cut off levels, % of positivity is displayed in ▶ **Fig. 1**. Within the first four months (TP1) positive TRAb levels can be measured in almost all patients (100 % with Thyretain Assay, 95 % IMMULITE bridge assay and 98 % TSH-R-Ab Elecsys binding assay). During the course of ATD treatment TRAb levels decrease. At TP 4, positive TRAb levels can be still measured with the Thyretain assay in 90 % of the patients while the other test systems are less sensitive after 2 years disease duration: IMMULITE bridge assay (70 %) and TSH-R-Ab Elecsys (66 %).

Since the IMMULITE assay is supposed to measure only the stimulating fraction of TRAb we compared the Immulite results directly with the Thyretain Assay in terms of positive and negative measurements. There were 24 % false positive (Thyretain negative but Immulite positive) and 9 % false negative (Thyretain positive and Immulite negative).

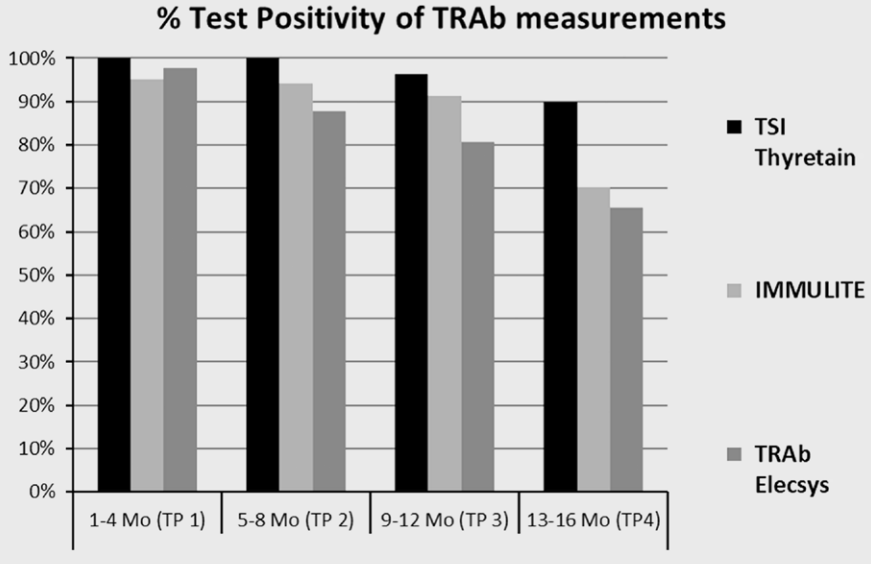
The course of the median antibody levels measured with all three assays are displayed in ▶ **Fig. 2**. TRAb levels are significantly higher in the relapse group in comparison to the remission group. A continuous decrease of TRAb towards normal median (at TP 4 when measured with Elecsys and IMMULITE assays) occurs in the remission groups, while TRAb stay wide in the positive range for the relapse group.

Cut off levels for predictions of relapse during the course of ATD therapy

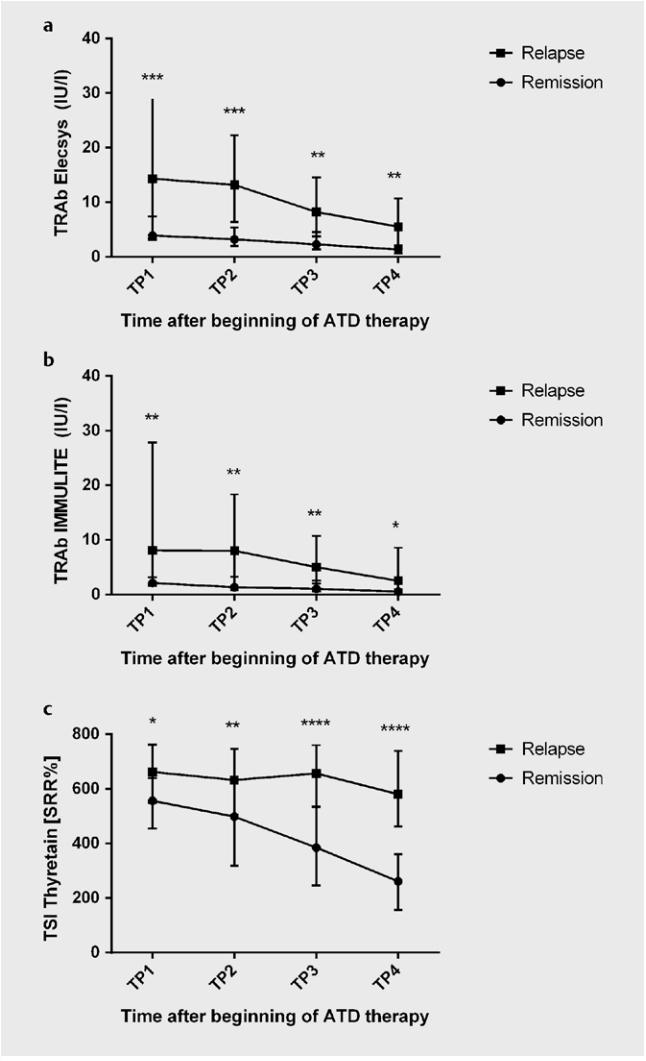
Receiver operator curve (ROC) analysis was performed and analyzed to select the best cut-off value with the highest specificity (≥ 90 %) to identify the patients who will relapse. At 90 % specificity TRAb thresholds, sensitivity levels and likelihood ratio were read out. TP 1 to 4 were selected for ROC analysis since ATD drug therapy is usually withdrawn within this time period.

▶ **Figure 3** (black dotted line) and ▶ **Table 2** (left part) show results of ROC analysis for the TSH-R-Ab Elecsys assay for the most clinically relevant TP 1–4. For all analyzed timepoints the area under the curve were between 0.74–0.8 (all significant). Relapse could be predicted with 90 % specificity at a cut off 13.3 IU/l at TP 1, 11.3 IU/l at TP 2, 8.3 IU/l at TP 3, and 4.6 IU/l at TP 4. Predictions could only be made for about half of the patients since sensitivity at 90 % specificity was 52 % (TP1), 59 % (TP2), 49 % (TP3), and 51 % (TP4), respectively. The likelihood ratios were between 4.7–9.2.

▶ **Figure 3** (triangles grey line) and ▶ **Table 2** (middle part) ROC analysis for the IMMULITE assay for the most clinically relevant TPs 1–4. For all analyzed timepoints the area under the curve were between 0.73–0.83 (all significant). Relapse could be predicted with 90 % specificity at a cut off 4.9 IU/l at TP 1, 5.3 IU/l at TP 2, 2.9 IU/l



▶ **Fig. 1** Percentage of measured positive TRAb levels during the course of Graves' hyperthyroidism. Within the first four months (Mo) after the beginning of antithyroid drug therapy TRAb can be measured in almost all patients (100 % with Thyretain Assay, 95 % IMMULITE bridge assay and 98 % TSH-R-Ab Elecsys binding assay). During the course of ATD treatment percentage of measured positive TRAb levels decreased. After two years positive TRAb levels can be still measured with the Thyretain bioassay in more in 85 % of the patients while the other test systems are less sensitive: IMMULITE bridge assay (64 %) and TSH-R-Ab Elecsys (63 %).



▶ **Fig. 2** Course of antibody levels during the course of ATD therapy measured with TSH-R-Ab Elecsys Assay **a**, IMMULITE brigde assay **b**, and Thyretain assay **c** for the remission and the relapse group. Cut off level for normal controls for Elecsys is 1.75 IU/l **a**, for IMMULITE 0.55 IU/l **b**, and for Thyretain 140SSR % **c**.

at TP 3, and 2.1 IU/l at TP 4. Predictions could only be made for more than half of the patients since sensitivity at 90 % specificity was 62 % (TP1), 62 % (TP2), 69 % (TP3), and 55 % (TP4), respectively. The likelihood ratios were between 5.6–8.2.

▶ **Figure 3** (squares black line) and ▶ **Table 2** (right part) show results of ROC analysis for the TSI Thyretain assay for the most clinically relevant TPs 1–4. For all analyzed TPs the area under the curve were between 0.67–0.78 (all significant). Relapse could be predicted with 90 % specificity at a cut off 794 % at TP 1, 687 % at TP 2, 650 % at TP 3, and 580 % at TP 4. At TP 1 predictions could only be made for a very small group of patients (21 %) with very high stimulating activity. At all other analyzed TP s predictions could be made for about half of the patients since sensitivity at 90 % specificity was 42 % (TP2), 52 % (TP3), 52 % (TP4), respectively, which was comparable to the other assays. The likelihood ratios were between 3.3–4.4 lower than in the other assay systems. ▶ **Table 3** shows the positive and negative predictive values for all cut off levels. As it can be seen the PPV is always high whereas the NPV is low.

Relapse and remission rates in relationship to severity of GO

Following ATD treatment, 35 patients (69 %) with mild course of GO went into thyroid disease remission, while only 3 patients (5 %) with a severe course of GO achieved remission ($p < 0.0001$). Contradictory, amongst patients with a severe GO course, 63 (95 %) needed definitive therapy in comparison to only 16 patients with mild course of GO (31 %). Therefore, manifestation of a severe GO is a strong predictor for not going into thyroid remission.

Serial TRAb analysis and group assignment

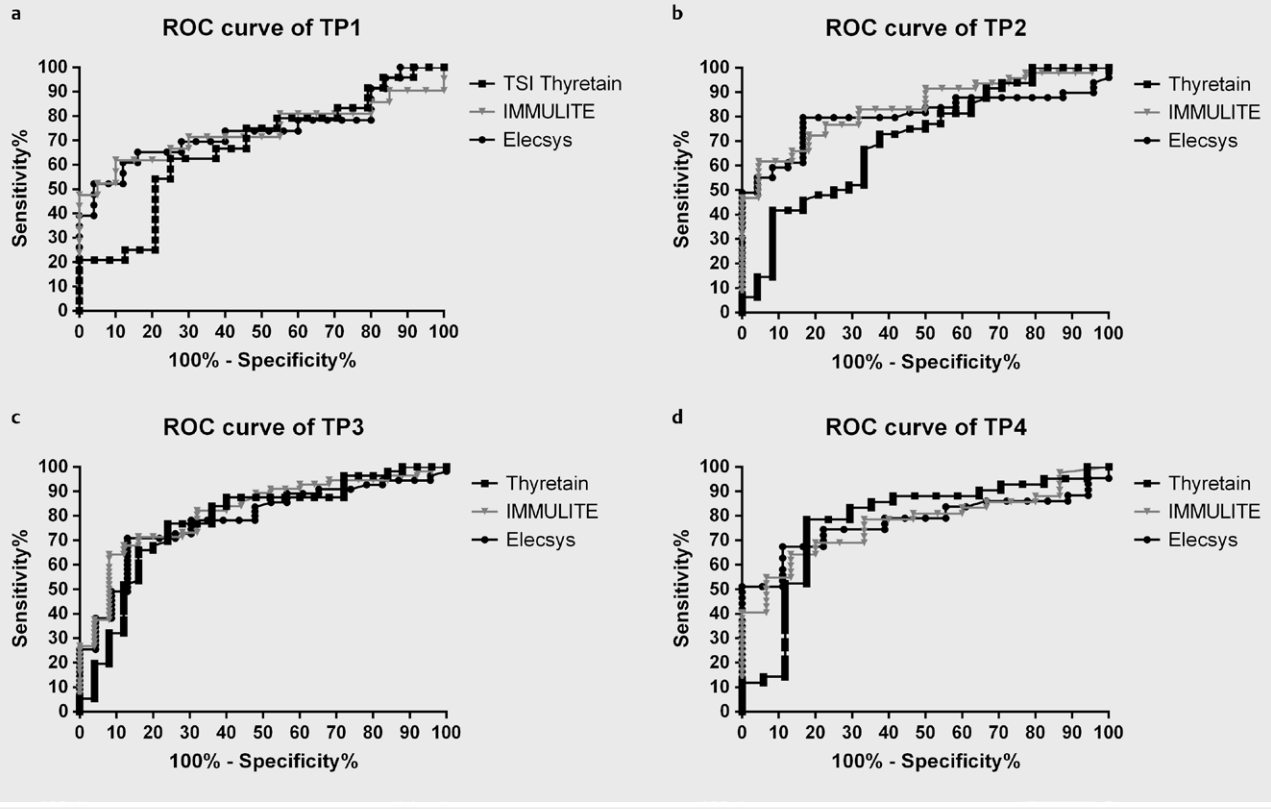
To analyze if patients switch between the groups at different timepoints, as shown in the heat map results in ▶ **Fig. 4**, patients were assigned into probable relapse according to TRAb above the cut off (green colour in the heat map) or not probable to relapse according to TRAb below the cut off (blue colour in the heat map). Green colour were almost exclusively only found in the group of patients were the clinical observation revealed a relapse. That means that once a patient is in the risk zone during serial measurements he/she will probably relapse. On the other hand, blue colour was found in both groups where the clinical observation revealed a relapse or a remission. As stated already, low TRAb titers are associated with both courses of hyperthyroidism. However the serial analysis revealed remission patients almost never have TRAb levels in the risk area. With serial TRAb analysis relapse patients could be identified in 61 % with Roche Elecsys, in 70 % with TSI IMMULITE and in 55 % with the Thyretain Assay (▶ **Fig. 4**).

Discussion

This analysis of serial TSH receptor autoantibody measurements in serum of GD/GO biobank is the most comprehensive comparative report on the clinical utility of assay technologies. The predictability of relapse vs remission of hyperthyroidism is 50–60 % relatively independent of the anti-TSH-R assay technology. For patients who test within the risk range (> cut-off) the recommendation is to continue ATD therapy or to decide definitive treatment of the thyroid. Within the first months of ATD therapy all assays have comparable high diagnostic sensitivity, however Thyretain bioassay shows the highest diagnostic sensitivity later in the course of hyperthyroidism. It turned out that serial TRAb measuremnts during the course of ATD therapy improve diagnostic significance.

Patient characteristics

Since patients were recruited in a tertiary referral center in the department of ophthalmology all patients derived from a more severely afflicted cohort – all have orbitopathy. Therefore, the remission group was smaller than the relapse group. In the literature, remission rates are higher with approximately 50 % in cohorts irrespective of the presence of GO [2–4, 24, 25]. This is a limitation of the study and it is suggested to perform a similar study in a primary referral center for patients with Graves' disease. In addition, the follow-up after cessation of ATD therapy is relatively short in comparison to other studies with several years of follow up. Impor-



▶ **Fig. 3** ROC analysis of TSH-R-Ab Elecsys Assay, IMMULITE brigde assay and Thyretain assay at TP 1 **a**, TP2 **b**, TP3 (T3**c**) and TP 4 **d** after beginning of ATD therapy and 12 after beginning of ATD treatment, patients (n = 38) with remission of hyperthyroidism (specifity cohort) and patients (n = 79) relapse (sensitivity cohort). The optimal cut off limit was calculated at for a specificity for relapse of 90 %.

tantly, other studies have shown a clear difference of remission rates between 1–2 and 4 years for follow up (1 year up to 70 %, 2 years about 50 %, 4 years 30 %) [24, 25].

However, since the center has a very close cooperation to the regional referring thyroid specialists, patients with mild GO and remission are referred to our center as well and therefore the cohort seems representative enough to perform such a study.

Correlation of measurements of TRAb levels with the three different assays

A very strong correlation between level of TRAb measured by the automated binding-assays TSH-R-Ab Elecsys and IMMULITE was found. The two binding assays gave similar results despite potential differences in binding sites and detection. The IMMULITE uses a chimeric TSHR, which has been postulated by the company to bind only stimulating TSHR autoantibodies (which is however controversial). A significant proportion of false positive and false negative values were detected by the IMMULITE in comparison to the Thyretain cAMP Bioassay. So it has to be assumed that the IMMULITE assay is not 100 % precise in measuring stimulating TRAb and blocking TBAb may also be detected. The weaker correlation between each of the binding assays, the TSH-R-Ab Elecsys or IMMULITE, and the cell-based bioassay TSI Thyretain supports this presumption. The Thyretain bioassay and the IMMULITE uses similar (if not identical) chimeric TSHR constructs to enable specific recognition of stimulating antibodies. However, the Thyretain bioassay can be used also to detect blocking antibodies indicating an overlap in the binding sites of stimulating and blocking antibodies [14]. It is a limitation of our study that blocking antibodies has not been measured. It appears that assays the two automated binding assays, TSH-R-Ab Elecsys and IMMULITE show in generally a stronger correlation than assays based on different methods like binding assay vs. cell-based bioassay. As a limitation of our study, another cell-based bioassay to correlate with the TSI Thyretain bioassay has not been included in the study.

Percentage of positivity of antibody levels during the course of GD/GO

Comparable diagnostic accuracy has been shown for all three applied types of assay technology in cohorts with untreated patients with Graves' disease [26]. The 3rd generation binding assays uses the monoclonal TRAb M22 instead of TSH. Sensitivity and specificity of 98 and 99 % have been described in untreated patients. This improvement was based on the binding of M22 to TSHR which does not dissociate in contrast to labelled TSH [27]. The assay applying bridge technology has been published to have a comparable sensitivity of 99.8 % and a specificity of 99.1 % [22]. The bioassay even reaches 100 % sensitivity in dilutions [28].

Positivity of antibody measurements persist for all assay systems over time but in the highest proportion for the TSI Thyretain

▲ **Table 2** Cut-off values predicting a relapse of hyperthyroidism (specificity 90 %) after ATD therapy calculated for TRAb measurements with the TSH-R-Ab Elecsys, the IMMULITE bridge assay, and the TSI Thyretain bioassay at four different timepoints during the course of ATD therapy.

	TSH-R-Ab Elecsys				IMMULITE				TSI Thyretain			
Timepoint	**Cut off level (IU/l)**	**Sensitivity (%) (95 % CI)**	**Area under the curve**	**Likelyhood ratio**	**Cut off level (IU/l)**	**Sensitivity (%) (95 % CI)**	**Area under the curve**	**Likelyhood ratio**	**Cut off level (IU/l)**	**Sensitivity (%) (95 % CI)**	**Area under the curve**	**Likely-hood ratio**
TP1: 1–4 months	13.3	52 (31–74)	0.74 (p = 0.004)	6.5	4.9	62 (38–82)	0.73 (p = 0.009)	6.2	794	21 (7–42)	0.67 (p = 0.04)	2.5
TP2: 5–8 months	11.3	59 (44–73)	0.80 (p < 0.0001)	4.7	5.3	62 (46–75)	0.83 (p < 0.0001)	6.7	687	42 (27–57)	0.71 (p = 0.0045)	3.3
TP3: 9–12 months	8.3	49 (35–63)	0.78 (p < 0.0001)	5.6	2.9	69 (54–79)	0.82 (p < 0.0001)	5.6	650	52 (38–65)	0.78 (p < 0.0001)	4.3
TP4: 13–16 months	4.6	51 (35–67)	0.77 (p = 0.0008)	9.2	2.1	55 (39–70)	0.77 (p < 0.0019)	8.2	580	52 (34–66)	0.78 (p = 0.0009)	4.4

bioassay with positivity of 90 % 13–16 months after the beginning of ATD therapy. More research has to be done in elucidating the relation between the biologic activity of the TRAb and the clinical course of thyroid and eye disease. First results delivered the sample dilution study of Diana et al 2017 showed higher detection sensitivity for the TSAb bioassay, and an antibody mixture study demonstrated exclusive specificity of the bioassays over all automated and ELISA binding assays [29].

Levels of antibodies during the course of GD

TRAB levels decreased over the course of ATD therapy which is in accordance with the literature [30]. However, especially in the measurements with the TSI Thyretain assay it can be clearly seen that there is much more persistence of TSAb levels in the relapse group in comparison to the remission group. However also in the remission group individual values do not decline into the normal range. This probably relates to greater sensitivity of the cell-based cAMP bioassay over the binding assays for the detection of thyroid stimulating immunoglobulins. Persisting positive [PDAE] TSI measured in a CHO-cell based, cAMP-dependent bioassays were also seen in the GD no relapse group of Struja et al. 2019 [18]. Persisting TSI indicate that thyroid autoimmunity is "simmering" also in remission patients and therefore, patients with euthyroid status but testing positive on Thyretain must possibly monitored at shorter intervals than patients with negative levels.

Cut off levels for predictions of relapse during the course of ATD therapy

When using variable TRAb tresholds, the new assay technologies are as useful as the second generation TRAb assay for assessment of relapse risk for hyperthyroidism after a course of ATD therapy. The third generation binding assay and the bridge assay reached a slightly higher sensitivity than the 2nd generation human TRAb assay [3]. It turned out that the binding inhibition assay technology is as efficient as the bridge and the bioassay. TRAb measurements with bridge and binding inhibition assay allow even more sensitive predictions in the early phase (TP1 and TP2).

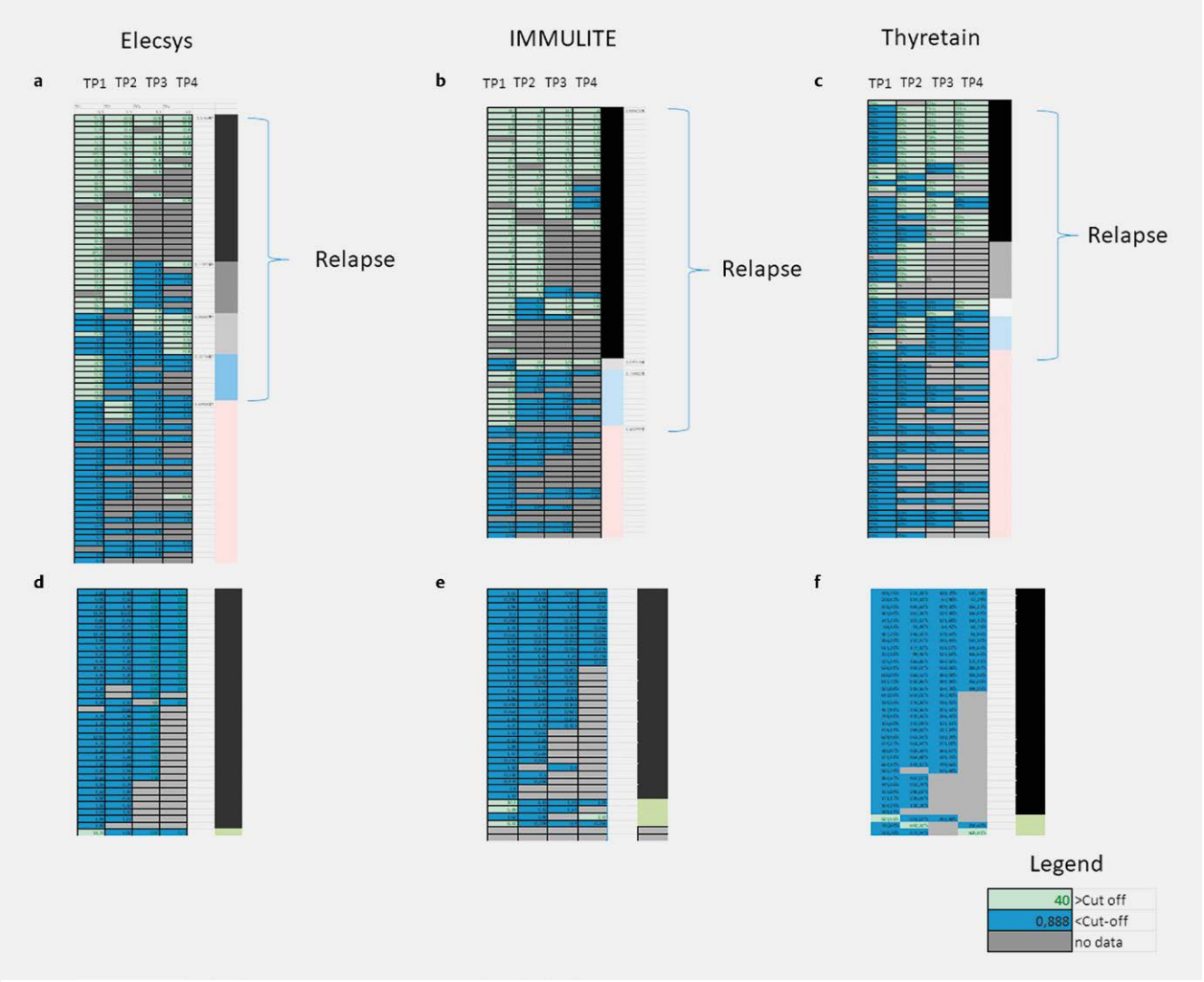
Hwang et al. published a comparison of the predictive ability of the binding inhibition assay using M22 and the Mc4 using bioassay. Similar area under the curves (AUCs) between Mc4-TSAb [AUC = 0.79 (95 % CI 0.69–0.89)] and M22-TRAb [AUC = 0.69 (95 % CI 0.58–0.81)] were published [31].

With dilutions of the serum samples it may be possible that relapse and remission rates of patients could be better differentiated since affinity and avidity of certain patients' antibodies contribute to more severe disease [32]. However, dilution studies are rather not practicable in daily routine diagnostic laboratories. Additionally, the aim of the original study was to provide cut offs for commercially applicable test systems.

Comparison to prior cut off levels achieved with the second generation TRAb assays revealed cut off levels almost in the same range [3–6, 24] for the third generation binding inhibition assay. Cut off levels for the TSI Thyretain and IMMULITE were not available in the literature yet and can now be used for clinical routine. During the early treatment phase (TP 1–4 and 5–8 months) Elecsys Anti-TSHR assay and IMMULITE Bridge assay are slightly more useful than the bioassay (higher sensitivity values) bioassay. At the timepoints

▶ **Table 3** Positive and negative predictive values for all three assays at four different timepoints during the course of ATD therapy.

	TSH-R-Ab Elecsys			IMMULITE			TSI Thyretain		
Timepoint	**Cut off level (IU/l)**	**PPV**	**NPV**	**Cut off level (IU/l)**			**Cut off level (IU/l)**		
TP1: 1–4 months	13.3	0.92	0.50	4.9	0.89	0.56	794	0.50	0.34
TP2: 5–8 months	11.3	0.96	0.50	5.3	0.97	0.54	687	0.88	0.45
TP3: 9–12 months	8.3	1.0	0.51	2.9	0.96	0.56	650	0.84	0.49
TP4: 13–16 months	4.6	1.0	0.53	2.1	0.87	0.50	580	0.84	0.44



▶ **Fig. 4** Consistency of the TRAb levels in the relapse group assessed with the different assays in each individual during TP1, TP2, TP3, and TP4 (green > cut-off, blue < cut-off, grey = no data available) **a** in 60 % of the cases is with Elecsys a prediction of a relapse possible. **b** In 70 % of the cases is a prediction of a relapse with Immulite possible. **c** In 55 % of the cases is a prediction of a relapse is with Thyretain possible. **d**–**f** Patients in the remission group showed in at least 80 % of the cases persistent low antibody levels.

9–12 and 13–16 months after the beginning of ATD therapy all assays deliver comparable results. TRAb assays demonstrate in this study slightly higher sensitivity to predict relapse at 90 % specificity-set point of clinical decision on therapy of thyroid. By virtue of the fact that patients are from Essen orbital centre mostly moderate and severe GO patients, the evaluation on the data for predicting relapse of hyperthyroidism, might be confounded by the underlying GO condition.

TRAb level measurement adds to other scoring systems of thyroid function for prognosis assessment like CSS (quantification of

Merseburg triad) and GREAT score (quantifying goiter size (by palpation), serum free thyroxine (FT4) and TRAb levels, and age at diagnosis) [7]. In comparison to CSS and GREAT, which are assessed at diagnosis–risk assessments with single TRAb measurements can be used during the whole course of ATD treatment. Due to the design of the study laboratory and clinical data before the start of antithyroid drug therapy were not available. Therefore, the add on effect of serial TRAB measurements to the GREAT or CSS score could not be evaluated.

Persisting high TRAb levels are an indicator for ongoing autoimmune stimulation. In the future thyroid specialists should consider immunosuppressive/immunomodulatory therapy not only for patients with severe GO but possibly also for patients which reach the cut off for relapse to improve the prognosis for ATD therapy, because it has been shown with the serial evaluation that once a patient reaches the relapse zone the probability to switch in the remission group is almost zero. Immunotherapeutic options have been recently tested to be effective in treating autoimmune hyperthyroidism as well in a GD mouse model [33].

Despite the considerations of immunosuppression definitive treatment of the thyroid is the other treatment decision to be made in patients with high TRAb levels and therefore low chance of remission. According to high TRAb levels if intended decision towards definitive treatment can be made quite early after the beginning of ATD treatment. Radioiodine therapy is contraindicated at least in patients with additional recently manifested GO [1]. The risks of early near total thyroidectomy [34] must be set in relation to the prognosis of long term ATD treatment [35]. However, since patients with high TRAb levels have usually more severe stages of orbitopathy these patients are likely to have a need for surgery for proptosis, squint and lid changes. For these patients the timing for treatment decisions for hyperthyroidism is very important. Therefore, the new cut off levels will be very useful for these patients to make the decisions either towards definitive treatment of the thyroid before ophthalmic rehabilitative surgery ore to maintain ATD therapy if TRAb levels are above the cut offs.

Conclusion

Identification of relapsing hyperthyroidism is possible with automated immunoassays and cell-based bioassay as early as 6 months after the beginning ATD therapy. Serial evaluation of group assignment revealed that once a patient reaches the relapse zone the probability to switch in the remission group is almost zero. However, since this is a retrospective study with also missing values this results should be confirmed in a prospective study.

Funding Information

This study was supported by the Junior Clinician Scientist Program of the University Medicine Essen Clinician Scientist Academy (UMEA).

Acknowledgements

We thank M. Horstmann (Department for molecular ophthalmology, Essen, Germany) and U. Kaiser (Department of ophthalmology, Essen, Germany) for blood specimen and data collection. We thank Ashley Brown, Jennifer Holliday, and Ron Lollar of Quidel Athens Ohio USA for their technical excellence in performing the Thyretain bioassay.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

[1] Kahaly GJ, Bartalena L, Hegedüs L et al. 2018; European Thyroid Association Guideline for the Management of Graves' Hyperthyroidism. Eur Thyroid J 2018: 167–186

[2] Orgiazzi J, Madec AM. Reduction of the risk of relapse after withdrawal of medical therapy for Graves' disease. Thyroid 2002; 12: 849–853

[3] Schott M, Morgenthaler NG, Fritzen R . et al. Levels of autoantibodies against human TSH receptor predict relapse of hyperthyroidism in Graves' disease. Horm Metab Res 2004; 36: 92–96

[4] Quadbeck B, Hoermann R, Roggenbuck U et al. Sensitive thyrotropin and thyrotropin-receptor antibody determinations one month after discontinuation of antithyroid drug treatment as predictors of relapse in Graves' disease. Thyroid 2005; 15: 1047–1054

[5] Carella C, Mazziotti G, Sorvillo F et al. Serum thyrotropin receptor antibodies concentrations in patients with Graves' disease before, at the end of methimazole treatment, and after drug withdrawal: Evidence that the activity of thyrotropin receptor antibody and/or thyroid response modify during the observation period. Thyroid 2006; 16: 295–302

[6] Eckstein AK, Lax H, Losch C et al. Patients with severe Graves' ophthalmopathy have a higher risk of relapsing hyperthyroidism and are unlikely to remain in remission. Clin Endocrinol (Oxf) 2007; 67: 607–612

[7] Masiello E, Veronesi G, Gallo D et al. Antithyroid drug treatment for Graves' disease: baseline predictive models of relapse after treatment for a patient-tailored management. J Endocrinol Invest 2018; 41: 1425–1432

[8] Zimmermann-Belsing T, Nygaard B, Rasmussen AK et al. Use of the 2nd generation TRAK human assay did not improve prediction of relapse after antithyroid medical therapy of Graves' disease. Eur J Endocrinol 2002; 146: 173–177

[9] McIver B, Rae P, Beckett G et al. Lack of effect of thyroxine in patients with Graves' hyperthyroidism who are treated with an antithyroid drug. N Engl J Med 1996; 334: 220–224

[10] Schott M, Eckstein A, Willenberg HS et al. Improved prediction of relapse of Graves' thyrotoxicosis by combined determination of TSH receptor and thyroperoxidase antibodies. Horm Metab Res 2007; 39: 56–61

[11] Schott M, Hermsen D, Broecker-Preuss M, Casati M et al. Clinical value of the first automated TSH receptor autoantibody assay for the diagnosis of Graves' disease (GD): An international multicentre trial. Clin Endocrinol (Oxf) 2009; 71: 566–573

[12] 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

[13] Allelein S, Ehlers M, Goretzki S et al. Clinical evaluation of the first automated assay for the detection of stimulating TSH receptor autoantibodies. Horm Metab Res 2016; 48: 795–801

[14] Allelein S, Diana T, Ehlers M et al. Comparison of a bridge immunoassay with two bioassays for thyrotropin receptor antibody detection and differentiation. Horm Metab Res 2019; 51: 341–346

[15] Lytton S, Li Y, Olivo P et al. Novel chimeric thyroid-stimulating hormone-receptor bioassay for thyroid-stimulating immunoglobulin. Clin Exp Immunol 2010; 162: 438–446

[16] Giuliani C, Cerrone D, Harii N et al. A TSHr-LH/CGr chimera that measures functional TSAb in Graves' disease. J Clin Endocrinol Metab 2012; 97: E1106–E1115

[17] Giuliani C, Cerrone D, Harii N, Thornton M et al. A TSHR-LH/CGR chimera that measures functional thyroid-stimulating autoantibodies (TSAb) can predict remission or recurrence in Graves' patients undergoing antithyroid drug (ATD) treatment. J Clin Endocrinol Metab 2012; 97: E1080–E1087

[18] Struja T, Jutzi R, Imahorn N et al. Comparison of Five TSH-receptor antibody assays in Graves' disease: results from an observational pilot study. BMC Endocr Disord 2019; 19: 38

[19] Autilio C, Morelli R, Locantore P et al. Stimulating TSH receptor autoantibodies immunoassay: Analytical evaluation and clinical performance in Graves' disease. Ann Clin Biochem 2018; 55: 172–177

[20] Smith BR, Bolton J, Young S et al. A new assay for thyrotropin receptor autoantibodies. Thyroid 2004; 14: 830–835

[21] Hermsen D, Broecker-Preuss M, Casati M et al. Technical evaluation of the first fully automated assay for the detection of TSH receptor autoantibodies. Clin Chim Acta 2009; 401: 84–89

[22] Frank CU, Braeth S, Dietrich JW et al. Bridge Technology with TSH receptor chimera for sensitive direct detection of TSH receptor antibodies causing Graves' disease: Analytical and Clinical Evaluation. Horm Metab Res 2015; 47: 880–888

[23] Tozzoli R, D'Aurizio F, Villalta D et al. Evaluation of the first fully automated immunoassay method for the measurement of stimulating TSH receptor autoantibodies in Graves' disease. Clin Chem Lab Med 2017; 55: 58–64

[24] Tun NN, Beckett G, Zammitt NN et al. Thyrotropin receptor antibody levels at diagnosis and after thionamide course predict Graves' disease relapse. Thyroid 2016; 26: 1004–1009

[25] Massart C, Gibassier J, d'Herbomez M. Clinical value of M22-based assays for TSH-receptor antibody (TRAb) in the follow-up of antithyroid drug treated Graves' disease: Comparison with the second generation human TRAb assay. Clin Chim Acta 2009; 407: 62–66

[26] Diana T, Wuster C, Kanitz M et al. Highly variable sensitivity of five binding and two bio-assays for TSH-receptor antibodies. J Endocrinol Invest 2016; 39: 1159–1165

[27] Schott M, Feldkamp J, Bathan C et al. Detecting TSH-receptor antibodies with the recombinant TBII assay: Technical and clinical evaluation. Horm Metab Res 2000; 32: 429–435

[28] Diana T, Kanitz M, Lehmann M et al. Standardization of a bioassay for thyrotropin receptor stimulating autoantibodies. Thyroid 2015; 25: 169–175

[29] Diana T, Krause J, Olivo PD et al. Prevalence and clinical relevance of thyroid stimulating hormone receptor-blocking antibodies in autoimmune thyroid disease. Clin Exp Immunol 2017; 189: 304–309

[30] Laurberg P, Wallin G, Tallstedt L et al. TSH-receptor autoimmunity in Graves' disease after therapy with anti-thyroid drugs, surgery, or radioiodine: A 5-year prospective randomized study. Eur J Endocrinol 2008; 158: 69–75

[31] Hwang S, Shin DY, Song MK et al. High cut-off value of a chimeric TSH receptor (Mc4)-based bioassay may improve prediction of relapse in Graves' disease for 12 months. Endocrine 2015; 48: 89–95

[32] Diana T, Wüster C, Olivo PD et al. Performance and specificity of 6 immunoassays for TSH receptor antibodies: A multicenter study. Eur Thyroid J 2017; 6: 243–249

[33] Plohn S, Hose M, Schluter A et al. Fingolimod improves the outcome of experimental Graves' Disease and Associated orbitopathy by modulating the autoimmune response to the thyroid-stimulating hormone receptor. Thyroid : Official Journal of the American Thyroid Association 2019; 29: 1286–1301

[34] Meyer Zu Horste M, Pateronis K, Walz MK et al. The effect of early thyroidectomy on the course of active Graves' Orbitopathy (GO): A retrospective case study. Horm Metab Res 2016; 48: 433–439

[35] Elbers L, Mourits M, Wiersinga W. Outcome of very long-term treatment with antithyroid drugs in Graves' hyperthyroidism associated with Graves' orbitopathy. Thyroid 2011; 21: 279–283

[36] Vitti P, Rago T, Chiovato L et al. Clinical features of patients with Graves' disease undergoing remission after antithyroid drug treatment. Thyroid 1997; 7: 369–375

[37] Eckstein AK, Plicht M, Lax H et al. Thyrotropin receptor autoantibodies are independent risk factors for Graves' ophthalmopathy and help to predict severity and outcome of the disease. J Clin Endocrinol Metab 2006; 91: 3464–3470

[38] Perros P, Crombie AL, Matthews JN et al. Age and gender influence the severity of thyroid-associated ophthalmopathy: A study of 101 patients attending a combined thyroid-eye clinic. Clin Endocrinol (Oxf) 1993; 38: 367–372