The Prediction Model Using Thyroid Stimulating Immunoglobulin Bioassay For Relapse of Graves’ Disease

Han-Sang Baek, Jaejun Lee, Chai-Ho Jeong, Jeongmin Lee, Jeonghoon Ha, Kwanhoon Jo, Min Hee Kim, Jae Hyoung Cho, Moo Il Kang, and Dong-Jun Lim

1Division of Endocrinology and Metabolism, Department of Internal Medicine, Seoul St. Mary’s Hospital, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea
2Division of Endocrinology and Metabolism, Department of Internal Medicine, Uijeongbu St. Mary’s Hospital, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea
3Division of Endocrinology and Metabolism, Department of Internal Medicine, Incheon St. Mary’s Hospital, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea
4Division of Endocrinology and Metabolism, Department of Internal Medicine, Eunpyeong St. Mary’s Hospital, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea
5Department of Internal Medicines, Armed Forces Goyang Hospital, Goyang, Republic of Korea

© The Author(s) 2022. Published by Oxford University Press on behalf of the Endocrine Society.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs licence (https://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial reproduction and distribution of the work, in any medium, provided the original work is not altered or transformed in any way, and that the work is properly cited. For commercial re-use, please contact journals.permissions@oup.com.
Correspondence: Dong-Jun Lim, MD, PhD

Division of Endocrinology and Metabolism, Department of Internal Medicine, Seoul St.
Mary’s Hospital, College of Medicine, The Catholic University of Korea, 222, Banpo-daero,
Seocho-gu, Seoul, 06591, Republic of Korea
E-mail: ldj6026@catholic.ac.kr

ORCiD numbers:
Dong-jun Lim: https://orcid.org/0000-0003-0995-6482
Han-sang Baek: https://orcid.org/0000-0002-6896-9024

Disclosure Statement: The authors have nothing to disclose
Abstract

Objective: Thyroid-stimulating immunoglobulin (TSI) bioassay has a better ability to predict the relapse rate of Graves’ disease (GD) than the TSH binding inhibitory immunoglobulin (TBII) method in terms of measuring the TSH receptor antibody. However, the optimal TSI bioassay cut-off for predicting relapse after antithyroid drug (ATD) withdrawal is not well evaluated.

Methods: This retrospective study enrolled GD patients who had been treated with ATD and obtained their TSI bioassay below 140% from January 2010 to December 2019 in a referral hospital.

Results: Among 219 study subjects, 86 patients (39.3%) experienced relapse. The TSI bioassay value of 66.5% significantly predicted the relapse of GD (p=0.049). The group with TSI bioassay value of above 66.5% were expected to show 23.8% at two and five years from ATD withdrawal, otherwise TSI below 66.5 group with 12.7% relapse rates based on Kaplan-Meier curves analysis. The TSI bioassay showed a good predictive ability to relapse GD in the female group (p=0.041) but did not in the male group (p=0.573). The risk scoring based on the nomogram with risk factors for GD relapse, which was constructed to overcome the limitation, increased the predictive ability of GD relapse by 11.5% than the use of the TSI bioassay alone.

Conclusions: The cut-off value of the TSI bioassay to predict GD relapse should be lower than that for diagnosing GD. However, as the single-use of the TSI bioassay has limitations, a nomogram with multiple risk factors including TSI bioassay could be helpful to predict GD relapse.

Key Words: Graves’ disease; Immunoglobulins, Thyroid-Stimulating, Recurrence; Nomograms
Introduction

Graves’ disease (GD) is one of the most well-known autoimmune thyroid diseases (1). It is well known that the thyrotropin receptor antibody (TSH-R-Ab) plays an important role in the pathogenesis of GD by causing thyroid stimulation and inducing hyperthyroidism (2). However, this TSH-R-Ab has a different action from the TSH receptor; stimulation or blocking (1-3). Stimulating TSH-R-Ab activates the cyclic adenosine monophosphate (c-AMP) pathway to stimulate the TSH receptor, thus inducing thyroid growth and increasing thyroid hormone production (2, 4). On the other hand, blocking TSH-R-Ab acts as an antagonist to the TSH receptor (2, 3).

There are two assays for TSH-R-Ab detection; the competitive thyrotropin-binding inhibitory immunoglobulin (TBII) assay and the thyroid stimulatory immunoglobulin (TSI) bioassay (1). Immunoglobulins that inhibit the binding of radiolabeled TSH to the TSH receptor could be detected by the TBII assay (5, 6). The problem is that this assay measures thyroid-blocking immunoglobulins, as well as TSIs [6]. On the contrary, the TSI bioassay could differentiate between stimulating TSH-R-Ab and blocking TSH-R-Ab (7, 8). The TSI bioassay can measure the c-AMP produced when TSI stimulates the TSH receptor (5). Although the TBII assay has limitations, TBII offers an accurate diagnosis of GD, and the TSI bioassay is predictive of extrathyroidal manifestations. (6, 9, 10).

For the treatment of GD, there are three options; surgery, radioactive iodine treatment (RAI), or antithyroid drug (ATD) (4, 10). While surgery or RAI treats GD by destroying thyroid tissue, ATD inhibits the synthesis of thyroid hormone to treat GD without destroying the thyroid structure. This is an advantage of ATD and a limitation simultaneously; the relapse from remaining thyroid tissue is always a concern (11). According to previous studies, the relapse rate after ATD withdrawal almost approached 50% (12, 13). In addition, many clinical factors such as sex of men, younger age, smoking, severe hyperthyroidism,
large goiter, and orbitopathy are associated with a high relapse rate (14). In addition, there is debate about ATD use during pregnancy because there could be harmful for embryonic development (15).

Furthermore, TSH-R-Ab levels showed a good ability to predict relapse and disease course in previous studies (16, 17). In these studies, the TBII assay was used to measure the TSH-R-Abs. Because it measures both stimulating and blocking antibodies, the TSI bioassay method appeared to be more accurate in predicting the course of disease (18, 19). Kwon H et al. showed that the TSI bioassay could better predict relapse after withdrawal from ATD (20). However, they did not measure two assays (TBII and TSI bioassay) simultaneously in one person and used a predetermined cut-off point of the TSI bioassay derived from the diagnosis of Graves’ disease, not based on the prognosis of Graves’ disease. Because they only used the positivity of the assay without quantitative measurement, the exact cut-off value to predict relapse was difficult to find.

Although the TSI bioassay has a better ability to predict relapse of GD, it is not known whether the TSI bioassay cut-off value for diagnosing GD and predicting relapse is the same. Therefore, in this study, we tried to achieve the optimal TSI bioassay cut-off value to predict relapse after withdrawal from ATD in patients with the results of two assays. Furthermore, we tried to make a prediction model with confounding factors for the relapse of GD.

Methods

This study adhered to the tenets of the Declaration of Helsinki and was approved by the institutional review board of the Seoul St. Mary’s hospital (Seoul, Korea) (KCR21RASI0731).
We tried to find out patients whose physician stopped ATD after checking that the TSI bioassay showed below 140, which is the TSI bioassay cut-off value to diagnose GD in our hospital (21). For this reason, among 1,831 patients who had TSI bioassay results and underwent ATD therapy between 2010 and 2019, 1,160 were excluded because the value of the TSI bioassay was greater than 140. Remained 671 medical records were reviewed. Patients who did not have follow-up data (102) or whose start date for ATD is unknown (98) and whose final diagnosis is not GD (66) were excluded. Patients who never stopped ATD or whose period from TSI bioassay measurement to ATD stop was longer than 3 months were excluded (158). So, only the patients with the TSI bioassay result less than 3 months before ATD stop were included for the analysis. A patient was excluded due to the lack of data from the TBII assay. Patients who received thyroid surgery or radioactive iodine treatment were also excluded (3). All patients had both the TSI bioassay result and the TBII result. The patient selection flow chart is summarized in Figure 1.

The ATD withdrawal was defined as the patient did not take ATD at least for three months. Patients who stopped ATD were divided into relapse group (R) and nonrelapse group (N-R). Relapse was defined as biochemical or clinical hyperthyroidism leading the clinician to re-start ATD; free T4 was higher than the upper normal limit, or TSH was lower than the lower normal limit and newly developed GD symptoms such as palpitation and sweating. If the patient did not show a relapse of the disease at least one year after stopping ATD, it was considered non-relapse even if there were no data after the last follow-up data (22-24). Patients who stopped ATD for less than three months in the relapse group (5) and patients who had a shorter follow-up duration than one year in the non-relapse group (19) were excluded. Finally, 219 remaining patients (158 women and 61 males) were analyzed.
Graves orbitopathy was defined when the study subjects had a medical history of the diagnosis of it by the ophthalmologist. If there was a prescription for levothyroxine during ATD treatment, it was considered concomitant T4 replacement.

**Laboratory test**

Free T4 and TSH were performed in two ways: 1) using the BECKMAN immunoradiometric assay kit (IRMA) (Immunotech, Prague, Czech Republic), and 2) using the ADIVA Centaur electrochemiluminescence immunoassay kit (ECLIA) (Siemens Healthcare Diagnostic Inc. USA). The normal ranges were as follows: TSH of 0.55 to 4.78 uIU/mL in ECLIA and 0.17 to 4.05 IU/mL in IMRA, free T4 of 0.89 to 1.76 ng/dL in ECLIA and of 0.89 to 1.79 ng/dL in IRMA. (RRID: AB_2895179, AB_2895183 in ECLIA and AB_2895185, AB_2895187 in IMRA)

TRAb was measured in two ways: First, the Elecsys / Cobas electrochemiluminescence immunoassay kit (Roche Diagnostics, Mannheim, Germany) is the third generation TBII assay. This assay measures the inhibition of binding of the labeled monoclonal antibody clone M22 to the TSH receptor with a positive value greater than 1.75 IU/L. (RRID:AB_2801453)

Second, TRAK Human radioimmunoassay (RAI) kit (BRAHMS Thermo Scientific, Henningsdorf, Germany) is the second-generation TBII assay. Detection is based on the ability of TRAb to prevent the binding of labeled TSH to the TSH receptor with a positive value greater than 1.5 IU/L.

The TSI titer was measured by the Thyretain™ TSI reporter bioassay (Diagnostic Hybrids, Inc., Athens, OH, USA). The Thyretain kit is based on Chinese hamster ovary cells (CHO) transfected with chimeric TSHR, which has amino acids 262-335 substituted with 73 amino acids from the rat luteinizing hormone receptor (LH) (Mc4) (25, 26). Mc4 was
designed to limit the effect of TBI that exists coincidentally with TSI in up to 25% of patients with GD, which can interfere with TSI measurements. The results of the TSI bioassay are reported as specimen-to-reference ratio percentages (SRR%), calculated as follows: SRR% = (mean TSI specimen / mean TSI reference) × 100. A specimen was considered positive if SRR was ≥140%.

**Statistical analysis**

The two groups of patients (those who experienced relapse and those who did not experience relapse) were compared using the t-test or the chi-square test. To obtain the optimal cut-off value for the TSI bioassay to predict relapse of GD, receiver operating characteristic (ROC) curve analysis was performed. To obtain the odds ratio, Pearson’s chi-square test was used. Kaplan-Meier curves were used to obtain the overall relapse rate using the log-rank *p*-value.

The logistic regression model was used to adjust for confounders. Statistical significance was set at *p* ≤ 0.05. SPSS® v.24 (IBM Corp., New York, NY; formerly SPSS Inc., Chicago, IL, USA) was used for statistical analyses. The R software version 4.1.1 (R Project for Statistical Computing, Vienna, Austria) was used to build the nomogram. Graphs were produced using Prism version 8.02 (GraphPad Software Inc., La Jolla, CA, USA). For graphical improvement of nomogram, we used open web-site made by Dr. Jeehyoung Kim (Department of orthopedic surgery, Seoul Sacred Heart General Hospital, Seoul, Korea), available at https://tinyurl.com/Cox-Logistic-Nomogram

**Nomogram**

Nomograms graphically represent a complex mathematical formula [28]. In medical research, the nomogram identifies risk factors for a specific event and generates a prognostic model. Each factor is assigned a score in a nomogram-based on the estimated regression coefficients.
calculated from a complex statistical model such as logistic regression or the Cox
proportional hazard model. The more important the factor, the higher the score and the most
important factor is assigned 100 points. In addition, the length of the lines is proportional to
the impact in the model [29].

Results

Clinical characteristics of study subjects with relapse (Table 1)

Among 219 study subjects, 86 patients (39.3%) experienced relapse. Male sex and
methimazole use showed more frequent relapse than female sex and carbimazole use.
Younger age and lower TSH levels by IRMA at the end of ATD were also predictive factors
for GD relapse despite not being statistically significant. On the other hand, serum levels of
the TBII and TSI bioassay at the time of withdrawal from ATD between the R and N-R
groups were not significantly different. The presence of Graves’ orbitopathy also showed no
difference between the two groups. In N-R group (N=133), the median period from ATD stop to
last follow-up was 23 months (interquartile range [IQR], 15.5–44): 24 patients were followed one year
(they stop the visit), 48 patients were followed two years and 61 patients followed more than two
years. In R group (N=86), the median period from ATD stop to relapse was 11.5 months
(IQR, 5.75–23): 43 patients experienced relapse within one year after the ATD withdrawal, 23
patients did within two years, and the rest 20 patients experienced the relapse after two years
after the ATD withdrawal. (Fig.2). 15.8% of the N-R group and 17.4% of the R group had a
history of one cycle of ATD treatment, which is not statistically significant. Most of the study
subjects were treated with methimazole (MZ) (68.4% in the N-R group and 87.2% in the R
group). 12 in the N-R group and six in the R group changed to ATD during treatment. The
main reason for the change in ATD was carbimazole (CM) from MZ for further reduction in
ATD dose (9), PTU for pregnancy issue (6) and minor adverse events such as skin reaction
after MZ. In both groups, ATD was discontinued when the thyroid function test showed a euthyroid state. In this cohort, the TSI bioassay was the main indicator of withdrawal of ATD. So, when even TBII was positive, ATD was discontinued. In the N-R group, 24 cases (18.0%) showed positive TBII and in the R group, 20 cases (23.3%) showed positive TBII ($p=0.443$)

Clinical values of the TBII assay and the TSI bioassay to predict relapse of Graves’ disease

We wanted to predict GD relapse according to TBII positivity and each TSI bioassay cut-off value. (Table 2) Positive TBII levels showed an odds ratio of 1.376 without statistical significance ($p=0.347$). The cut-off value of the positive TSI bioassay was calculated in three ways; the mean value of the TSI bioassay in the study population (66.87%), the median value of the TSI bioassay in the study population (66%), and the final value was obtained from the ROC curve (66.5%). In ROC curve analysis, above the cut-off value (66.5%) to predict the relapse of GD showed 57.0% sensitivity, 60.9% specificity, and AUC was 0.557. The cut-off value from the mean of the TSI bioassay values and the ROC curve showed statistical significance with the odds ratio with the relapse of the disease. ($p=0.010$ in both groups).

We built a logistic regression model to find the risk factors for relapse of GD. (Table 3). Male sex and TSI bioassay value greater than 66.5% significantly increase the risk of relapse. (HR 2.476, $p=0.007$ in male sex, HR 1.992, $p=0.022$ in the TSI group above 66.5% group). Younger age increased the risk of GD relapse with borderline significance (HR 0.980, $p$-value = 0.052). The TBII positivity or the duration of ATD usage did not show a relationship with relapse. Graves orbitopathy increases the risk of relapse without statistical significance. Concomitant replacement of thyroid hormone reduced the risk without statistical significance. In Kaplan-Meier curve analysis, the TSI bioassay value of 66.5%
significantly predicted GD relapse (p=0.049) (Fig. 3). The TSI bioassays above 66.5% group and below 66.5% at the time of ATD stop were expected to show a relapse rate of 23.8% and 12.7%, respectively, at two years after the ATD stop (p=0.033). At five years after stopping ATD, it was expected that the TSI bioassay above 66.5% group would show a relapse rate of 40.3%, and the TSI bioassay below 66.5% group would show a relapse rate of 24.3% (p=0.017).

**TSI bioassay cut-off value to predict relapse of GD according to sex.**

Because male sex significantly increased the risk of GD relapse, we separately analyzed data by sex. (Table 3 and Fig 4). The TBII assay positivity and the TSI bioassay values did not show differences between the two sexes. In the male group, the optimal cut-off value of the TSI bioassay to predict relapse was the same as 66.5%, obtained from the whole study cohort. In ROC curve analysis, with optimal value, the TSI bioassay in the male group showed 58.8% sensitivity, 66.7% specificity, and AUC 0.557.

However, in the female group, the optimal cut-off value of the TSI bioassay for the relapse was 59.5%, less than that of the entire study cohort. In ROC curve analysis, with optimal value, the TSI bioassay in the female group showed 71.2% sensitivity, 46.2% specificity and AUC 0.560. The TSI bioassay showed better predictive ability in the female group (p=0.041) (Fig. 4A) than in the male group, in which the TSI bioassay was unable to predict the relapse of GD significantly (p=0.573) (Fig. 4B)

**The ATD regimen at withdrawal and the duration of ATD use were associated with the risk of relapse.**

In the logistic regression model, the ATD regimen at the withdrawal of the drug appears to be a possible factor for relapse of the disease. Carbimazole use in ATD reduces the
risk of relapse. (HR 0.248, p=0.004). Because PTU was used only for female patients and the number of patients who used each medication was different, we did a propensity score matching MZ and CM usage. After adjusting the sex and age variation for matching, TBII with the RAI assay (3.0 ± 5.2 vs 0.9 ± 0.7, p=0.040) and duration of ATD (22.6 ± 12.2 vs 36.1 ± 26.3 p=0.006) showed a significant difference between two groups of medications. (Table 4) The difference was not shown before matching. Moreover, TBII positivity also showed the difference between MZ and CM use groups (p=0.050). Simultaneous replacement of T4 was more common in the MZ group before (34.9% vs 15.8% p=0.036) and after (42.1% vs 15.8% p=0.023) propensity score matching but TSI bioassay did not show differences between the two groups (p = 0.79). Before and after the matching, the dose of MZ (3.0 ±1.3 mg) was higher than that of CM (2.3 ±1.6 mg) when ATD stop significantly.

**Nomogram construction for risk scoring for relapse of GD and prediction of relapse.**

The single-use of the TSI bioassay to predict relapse had limitations. Therefore, we constructed a nomogram to show the probability of GD relapse. The nomogram graphically represents the numerical relationships between the risk of relapse and four main risk factors (age, sex, TBII positivity and TSI bioassay positivity with a 66.5% cut-off value). For each risk factor, the assigned points were 100 for men, 62.6 for a younger age, 22.8 for TBII assay positivity and 67.2 for TSI bioassay positivity. (Fig. 5A) The points for each risk factor were assigned from the logistic regression model. For analysis, the age factor was divided by age 46, which was the median age of the study cohort. The higher the point in each factor, the more important factor in relapse. For example, assume that 31 years old (62.6) female (0) with negative TBII assay (0) and positive TSI bioassay (67.2). The total score for her is 129.8. (Fig. 5B) Her risk of relapse is 47%. To verify the discrimination of the nomogram, we performed a ROC curve analysis. We gained the risk scoring form constructed nomogram
for each of our patients in the study cohort. An optimal cut-off value (111.4) showed a sensitivity of 57%, specificity of 70.7%, and AUC 0.672. It was better able to predict GD relapse compared to the TSI bioassay single usage. The cumulative hazard curve between the two groups (risk score greater than 111.4 and risk score smaller than 111.4) showed a significant difference (p<0.01). (Fig. 6). At 2 years from ATD stop, the risk score below 111.4 group was expected to show an 11.5% relapse rate, and the risk score above 111.4 group was expected to show a 27.3% relapse rate (p=0.004).

Discussion
The finding of our study is summarized in figure 7. In our study, a large proportion (39.3%) of the patients experienced a relapse of GD, although their TSI bioassay values showed below the cut-off (140%) originally defined to diagnose GD. In addition, serum levels of TSI bioassay at the time of withdrawal from ATD between the R and N-R groups were not significantly different. In this context, one could assume that the cut-off point for TSI to diagnose GD and stop ATD should be different.

Kwon H. et al. showed that TSI bioassay at withdrawal from ATD could predict relapse of GD, but the TBII assay could not (20). However, the number of patients evaluated in their study was relatively small (TSI bioassay, n=35 and TBII assay, n=39), and both the TSI bioassay and TBII bioassay were not measured simultaneously within one person. Liu K et al. showed that the TSI bioassay could be used to predict response to methimazole treatment. However, they used the same TSI cut-off value like that to diagnose GD (18). The prospective trial of Kahaly GJ et al. showed that the TSI bioassay was positive at week 24 after stopping ATD in the group of nonresponders to methimazole or patients who experience relapse (19). However, they also used the TSI bioassay cut-off value to diagnose GD.

Unlike previous studies, our study, through ROC curve analysis, unveiled the newly calculated TSI bioassay cut-off value to discriminate the relapse of GD was 66.5% with AUC
0.557. With this cut-off value, the TSI bioassay above 66.5% group showed a relapse rate of 23.8% and the TSI bioassay below 66.5% group were expected to show a relapse rate of 12.7% at two years from the ATD stop in the Kaplan-Meier curves model. Two groups were expected to show 40.3% and 24.3% relapse rates, respectively, at five years from the ATD stop. According to an observational study, the negative TBII assay was associated with a 58% risk of relapse four years after discontinuation of ATD (16). Considering this result, the negative TSI bioassay with a cut-off value of 66.5% is expected to reduce the risk of relapse by about 20%.

However, the single-use of the TSI bioassay to predict the relapse of GD had some limitations. In addition, despite of lower TSI cut-off value, the AUC (0.557) of TSI-bioassay to predict the relapse of GD in ROC curve analysis was similar to that (0.62) from one observational pilot study with small population (27). Especially, the gender factor of male subject had a large contribution to increasing relapse. The sex of men significantly increased the risk of relapse. (HR 2.569, p=0.005). Meta-analysis or large population studies indicated male factor is a risk for relapse for GD (24, 28). One previous study revealed that estrogen is associated with B cell hyperactivity, causing severe autoimmune disease, and (29) Chailurkit LO et al. showed that higher circulating estradiol is related to thyroid autoimmunity in men (30). Furthermore, Ishido N et al suggested skewed X chromosome inactivation was likely related to the prognosis of GD (31).

Other risk factors for relapse that should be discussed are age at diagnosis and the drug regimen at the end of ATD. Younger age increased the risk of GD relapse with borderline significance. Bano A et al. showed an association between increased TBII and increased risk of relapse in younger patients (32). In the large population study in Japan, the TBII values have decreased with advancing age, which may explain the high probability of relapse at younger age in our study.
In our study, the use of CM at the time of ATD decreased the risk of relapse. (HR 0.248, p=0.004). Equivalent doses of ATD are 40 mg of CM, 30 mg of MZ, and 400 mg of PTU (4). Although CM has a lower potency than MZ, it would be possible to use CM longer than other ATDs. After matching the propensity score with sex and age, the use of ATD in the CM group showed a significantly longer duration. In randomized clinical trials, the long-term treatment with ATD increases remission rates (33). The proportion of TBII positivity, concomitant use of T4, and actual dosage was higher in MZ group. It maybe because physicians might tend to prescribe MZ more frequently than CM or PTU due to the difference in drug potency or might wait for the negative conversion of TBII with long-term use of CM. Concomitant use of T4 was probably because of the high efficacy of MZ and subsequent hypothyroidism.

From our results, the single usage of the TSI bioassay to predict GD relapse was difficult. In this context, we build the nomogram with multiple factors that could influence the risk of relapse. The logistic regression's score was assigned to each risk factor. When verified with the study cohort, the risk scoring gained by the constructed nomogram showed a good ability to predict the relapse of GD. With the optimal cut-off value (111.4), the nomogram-based risk scoring increased the ability to predict relapse by 11.5% compared to the single-use of TSI.

Previously, Vos XG et al. in 2016 developed a prediction model for the recurrence risk of GD (GREAT score, Graves’ Recurrent Events After Therapy), based on clinical data (age, goiter, serum free T4, and serum TBII level) as well as genetic predisposition of untreated 178 patient in a prospective, multicenter, observational study. (34) Although the TSI bioassay was not included in this model, the GREAT score was validated in a large retrospective observational study with 741 patients and concluded it might help treatment selection in GD patients(35). Secondly, a clinical severity score (CSS) was developed in 2018 from an
observational study of 387 consecutive, newly diagnosed GD patients, and showed similar predictability to GREAT score, although it only included clinical parameters such as goiter, orbitopathy and fT4 but did not include TBII (36, 37). These two models at the time of diagnosis can predict the relapse after one course of ATD treatment so it may be helpful at initial stage to decide whether to start ATD or to move on to definite treatments such as thyroidectomy or radioiodine therapy. On the other hand, as our prognostic model was generated based on study subjects on ATD just before stopping ATD, it may guide whether to stop ATD or to continue ATC in a long-term manner.

Our study has some limitations. First, it is a retrospective study and only included patients whose TSI bioassay value was less than 140. Therefore, it is hard to say that our study cohort represents all patients with GD. Second, although we defined a minimum follow-up period of remission as one year, as the previous studies defined, it could be relatively short. Several patients did not visit the hospital after one year of follow-up from the ATD stop because they thought themselves to be in remission, but possibly they might show relapse and visit another hospital for a check-up. Third, the smoking status or the presence of goiter, which could be a possible factor associated with relapse of GD, could not be included in our study. Finally, we were not able to compare the TSI values after ATD stop to the baseline TSI values, because of a lack of data. In Korea, the TSI bioassay can be covered by the insurance only in specific cases; (a) when the diagnosis of GD is unclear because the TBII is negative or not enough high, (b) to check the probability of GD relapse before ATD withdrawal, and (c) in patients with Graves’ orbitopathy or in the 3rd trimester in pregnancy. The diagnosis of GD was based on clinical history and TBII results. In our study population, only six patients had baseline TSI values: two patients in case (a), three patients in case (c), and one patient without insurance coverage. In the near future, a prognostic model with incorporation of baseline TSI, which was not measured in our study, may better predict a relapse of GD. We also could not
evaluate free T3, total T4, total T3, ratio of T4/T3, thyroglobulin and thyroid peroxidase because of insurance issue.

In conclusion, the TSI bioassay cut-off point to predict GD relapse should be lower than that to diagnose GD. The male factor and the medication regimen at the end of ATD could be attributed to the risk of relapse. Nomogram with multiple predictive factors such as age, sex, TBII, and TSI can predict GD relapse more effectively.
Acknowledgements

We appreciate the help for statistical support to refine the nomogram by Dr. Jeehyoung Kim (Department of Orthopedic Surgery, Seoul Sacred Heart General Hospital, Seoul, Korea).

Author Declaration

A portion of this study was presented in abstract form at the SICEM 2021 (Seoul International Congress of Endocrinology and Metabolism 2021) in Busan, Korea.

Ethics approval and consent to participate

The study adhered to the tenets of the Declaration of Helsinki and was approved by the institutional review board of Seoul St. Mary’s hospital (KCR21RASI0731). Permission to use hospital data was granted by the institutional review board of Seoul St. Mary’s hospital.

Consent for publication

Due to the retrospective nature of the study, the requirement to obtain informed consent was waived by the institutional review board of St. Mary’s hospital.

Competing interests

The authors declare that they have no competing interests.

Funding

None.
Authors’ contributions

Han-Sang Baek and Dong-Jun Lima mainly designed the study. Han-Sang Baek mainly wrote the manuscript. Dong-Jun Lim supervised the study and is corresponding author.

Jaejun Lee contributed to data analysis. All authors contributed to drafting or revising the article, gave final approval of the version to be published, agreed to the submitted journal, and agree to be accountable for all aspects of the work.

Availability of data and materials

The datasets generated and/or analysed during the current study are not publicly available due to personal data protection legislation but are available from the first author or corresponding author on reasonable request.
References

1. Kamath C, A. AM, Premawardhana LD. The role of thyrotrophin receptor antibody assays in graves' disease. *J Thyroid Res.* 2012;2012:525936.

2. Morshed SA, Davies TF. Graves' Disease Mechanisms: The Role of Stimulating, Blocking, and Cleavage Region TSH Receptor Antibodies. *Horm Metab Res.* 2015;47(10):727-734.

3. Giuliani C, Saji M, Bucci I, Napolitano G. Bioassays for TSH Receptor Autoantibodies, from FRTL-5 Cells to TSH Receptor-LH/CG Receptor Chimeras: The Contribution of Leonard D. Kohn. *Front Endocrinol (Lausanne).* 2016;7:103.

4. Badiu Corin MD. WILLIAMS TEXTBOOK OF ENDOCRINOLOGY. *Acta Endocrinol (Buchar).* 2019;15(3):416.

5. Barbesino G, Tomer Y. Clinical review: Clinical utility of TSH receptor antibodies. *J Clin Endocrinol Metab.* 2013;98(6):2247-2255.

6. Tozzoli R, Bagnasco M, Giavarina D, Bizzaro N. TSH receptor autoantibody immunoassay in patients with Graves' disease: improvement of diagnostic accuracy over different generations of methods. Systematic review and meta-analysis. *Autoimmun Rev.* 2012;12(2):107-113.

7. Diana T, Wüster C, Kanitz M, Kahaly GJ. Highly variable sensitivity of five binding and two bio-assays for TSH-receptor antibodies. *J Endocrinol Invest.* 2016;39(10):1159-1165.

8. 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(5):243-249.

9. Scappaticcio L, Trimboli P, Keller F, Imperiali M, Piccardo A, Giovanella L. Diagnostic testing for Graves' or non-Graves' hyperthyroidism: A comparison of two
thyrotropin receptor antibody immunoassays with thyroid scintigraphy and ultrasonography. *Clin Endocrinol (Oxf).* 2020;92(2):169-178.

10. Kahaly GJ. Management of Graves Thyroidal and Extrathyroidal Disease: An Update. *J Clin Endocrinol Metab.* 2020;105(12):3704-3720.

11. Chung HK. Medical treatment of Graves' disease. *Int J Thyroidol.* 2019;12(2):79-84.

12. Cooper DS. Antithyroid drugs. *N Engl J Med.* 2005;352(9):905-917.

13. Feldt-Rasmussen U, H. S, Carayon P. Meta-analysis evaluation of the impact of thyrotropin receptor antibodies on long term remission after medical therapy of Graves' disease. *J Clin Endocrinol Metab.* 1994;78(1):98-102.

14. Wiersinga WM. Graves' Disease: Can It Be Cured? *Endocrinol Metab (Seoul).* 2019;34(1):29-38.

15. Van Herck SL, Geysens S Fau - Bald E, Bald E Fau - Chwatko G, et al. Maternal transfer of methimazole and effects on thyroid hormone availability in embryonic tissues. (1479-6805 (Electronic)).

16. Tun NNZ, Beckett G, Zammit NN, Strachan MWJ, Seckl JR, Gibb FW. Thyrotropin Receptor Antibody Levels at Diagnosis and After Thionamide Course Predict Graves' Disease Relapse. *Thyroid.* 2016;26(8):1004-1009.

17. 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(9):1047-1054.

18. Liu K, Fu Y, Li T, et al. Clinical efficacy of thyroid-stimulating immunoglobulin detection for diagnosing Graves' disease and predictors of responsiveness to methimazole. *Clin Biochem.* 2021;97:34-40.

19. Kahaly GJ, Diana T, Kanitz M, Frommer L, Olivo PD. Prospective Trial of Functional
Thyrotropin Receptor Antibodies in Graves Disease. *J Clin Endocrinol Metab.* 2020;105(4):e1006-e1014.

20. Kwon H, Kim WG, Jang EK, et al. Usefulness of Measuring Thyroid Stimulating Antibody at the Time of Antithyroid Drug Withdrawal for Predicting Relapse of Graves Disease. *Endocrinol Metab (Seoul).* 2016;31(2):300-310.

21. Lytton SD, Li Y, Olivo PD, Kohn LD, Kahaly GJ. Novel chimeric thyroid-stimulating hormone-receptor bioassay for thyroid-stimulating immunoglobulins. *Clin Exp Immunol.* 2010;162(3):438-446.

22. Wang SY, Wang CT, Tien KJ, Chang CC, Liu TH. Thyroid-stimulating hormone receptor antibodies during follow-up as remission markers in childhood-onset Graves' disease treated with antithyroid drugs. *Kaohsiung J Med Sci.* 2020;36(4):281-286.

23. Ross DS, Burch HB, Cooper DS, et al. 2016 American Thyroid Association Guidelines for Diagnosis and Management of Hyperthyroidism and Other Causes of Thyrotoxicosis. *Thyroid.* 2016;26(10):1343-1421.

24. Suzuki N, Noh JY, Yoshimura R, et al. Does Age or Sex Relate to Severity or Treatment Prognosis in Graves' Disease? *Thyroid.* 2021;31(9):1409-1415.

25. Li Y, Kim J, Diana T, Klasen R, Olivo PD, Kahaly GJ. A novel bioassay for anti-thyrotrophin receptor autoantibodies detects both thyroid-blocking and stimulating activity. *Clin Exp Immunol.* 2013;173(3):390-397.

26. Watson PF, Ajjan RA, Phipps J, Metcalfe R, Weetman AP. A new chemiluminescent assay for the rapid detection of thyroid stimulating antibodies in Graves' disease. *Clin Endocrinol (Oxf).* 1998;49(5):577-581.

27. 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(1):38.
28. Zuhur SS, Yildiz I, Altuntas Y, et al. The effect of gender on response to antithyroid drugs and risk of relapse after discontinuation of the antithyroid drugs in patients with Graves' hyperthyroidism: a multicentre study. *Endokrynol Pol.* 2020;71(3):207-212.

29. Da Silva JA. Sex hormones, glucocorticoids and autoimmunity: facts and hypotheses. *Ann Rheum Dis.* 1995;54(1):6-16.

30. Chailurkit LO, Aekplakorn W, Ongphiphadhanakul B. The relationship between circulating estradiol and thyroid autoimmunity in males. *Eur J Endocrinol.* 2013;170(1):63-67.

31. Ishido N, Inoue N, Watanabe M, Hidaka Y, Iwatani Y. The relationship between skewed X chromosome inactivation and the prognosis of Graves' and Hashimoto's diseases. *Thyroid.* 2015;25(2):256-261.

32. Bano A, Gan E, Addison C, et al. Age May Influence the Impact of TRAbs on Thyroid Function and Relapse-Risk in Patients With Graves Disease. *J Clin Endocrinol Metab.* 2019;104(5):1378-1385.

33. Azizi F, Amouzegar A, Tohidi M, et al. Increased Remission Rates After Long-Term Methimazole Therapy in Patients with Graves' Disease: Results of a Randomized Clinical Trial. *Thyroid.* 2019;29(9):1192-1200.

34. Vos XG, Endert E, Zwinderman AH, Tijssen JGP, Wiersinga WM. Predicting the Risk of Recurrence Before the Start of Antithyroid Drug Therapy in Patients With Graves' Hyperthyroidism. *Journal of Clinical Endocrinology & Metabolism.* 2016;101(4):1381-1389.

35. Struja T, Kaeslin M, Boesiger F, et al. External validation of the GREAT score to predict relapse risk in Graves' disease: results from a multicenter, retrospective study with 741 patients. *Eur J Endocrinol.* 2017;176(4):413-419.

36. Bartalena L, Masiello E, Magri F, et al. The phenotype of newly diagnosed Graves'
disease in Italy in recent years is milder than in the past: results of a large observational longitudinal study. *Journal of Endocrinological Investigation.* 2016;39(12):1445-1451.

37. 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(12):1425-1432.
Table Legends

Table 1.
Baseline characteristics of 219 patients who stopped ATD after checking that their TSI bioassay was below 140

Table 2.
Odds ratio to relapse of Graves’ disease according to each cut-off value of TSI and TBII

Table 3.
Logistic regression analysis of risk factors for relapse of Graves’ disease relapse

Table 4.
Clinical characteristics according to medication at withdrawal and propensity score correlated with age and sex results
Figure legends

**Figure 1.** Flow chart of study subjects. Initially, 671 patients who had a history of prescription of ATD and TSI bioassay less than 140%. The final 219 subjects were analyzed after exclusion 452 subjects with each exclusion criteria.

**Figure 2.** Study population. In N-R group (N=133), 24 patients were followed one year (they stop the visit), 48 patients were followed two years and 61 patients followed more than two years. In R group (N=86), 43 patients experienced relapse within one year after the ATD withdrawal, 23 patients did within two years, and the rest 20 patients experienced the relapse after two years after the ATD withdrawal.

**Figure 3.** Cumulative relapse rate curve according to the TSI cut-off with 66.5%. The TSI value of 66.5% significantly predicted the relapse of GD (p=0.049).

**Figure 4.** Cumulative Relapse Rate Curve According to Sex.

A. The TSI bioassay showed better predictive ability in the female group (p=0.041) 
B. TSI bioassay was unable to predict GD relapse significantly in the male group (p=0.573)

**Figure 5.** Nomogram construction for risk scoring to predict relapse of GD.

A. Based on logistic regression, its score was assigned to each risk factor. (100 for males, 62.6 for younger age, 22.8 for TBII assay positivity and 67.2 for TSI bioassay positivity with cut-off 66.5). Pr (event) represents the probability of the event. 
B. An example of applying the nomogram to risk scoring with 31 years of age (62.6) female (0) with negative TBII assay (0) and positive TSI assay (67.2). The total score for her is 129.8. (62.6+0+0+67.2) and her relapse risk is 47%.
**Figure 6.** Cumulative relapse rate curve according to the risk score calculated from the constructed nomogram. The cumulative hazard curve between the two groups (risk score greater than 111.4 and risk score smaller than 111.4) showed a significant difference (p<0.01).

**Figure 7.** Graphic abstract of our study. TSI, thyroid stimulating immunoglobulin; ATD; anti-thyroid drugs; TBII; inhibitory immunoglobulin binding to thyrotropin; HR, hazard ratio; GD, graves’ disease.
Table 1. Baseline characteristics of 219 patients who stopped ATD after checking that their TSI bioassay was below 140

|                                | Total (N=219) | No Relapse (N=133) | Relapse (N=86) | P-value |
|--------------------------------|---------------|--------------------|----------------|---------|
| Age                            | 45.3 ± 13.3   | 46.7 ± 13.2        | 43.2 ± 13.3    | 0.055   |
| Female sex                     | 158 (72.1%)   | 106 (79.7%)        | 52 (60.5%)     | 0.011   |
| Graves orbitopathy             | 22 (10.0%)    | 12 (9.0%)          | 10 (11.6%)     | 0.692   |
| Median period from ATD stop to relapse (R) or last follow-up (N-R) (month) | 19 (IQR, 11-33) | 23 (IQR, 15.5-44) | 11.5 (IQR, 5.75-23) | <0.001 |
| History of GD treatment        | 36 (16.4%)    | 21 (15.8%)         | 15 (17.4%)     | 0.892   |
| ATD regimen when stopping      |               |                    |                | 0.002   |
| Methimazole                    | 166 (75.8%)   | 91 (68.4%)         | 75 (87.2%)     |         |
| Carbimazole                    | 38 (17.4%)    | 32 (24.1%)         | 6 (7.0%)       |         |
| PTU                            | 15 (6.8%)     | 10 (7.5%)          | 5 (5.8%)       |         |
| Switch during treatment        | 18 (8.2%)     | 12 (9.0%)          | 6 (7.0%)       | 0.775   |
| Duration of ATD usage          | 31.0 ± 28.5   | 31.0 ± 29.5        | 31.0 ± 27.0    | 0.997   |
| Concomitant replacement of T4 during ATD | 67 (30.6%) | 44 (33.1%)       | 23 (26.7%)     | 0.399   |
| Thyroid function tests         |               |                    |                |         |
| fT4 at ATD stop                |               |                    |                |         |
| ECLIA (0.89-1.76 ng/dL)        | 1.2 ± 0.2     | 1.2 ± 0.2          | 1.2 ± 0.2      | 0.139   |
| IRMA (0.89-1.79 ng/dL)         | 1.3 ± 0.2     | 1.4 ± 0.2          | 1.3 ± 0.2      | 0.106   |
| TSH at the ATD stop            |               |                    |                |         |
| ECLIA (0.55-4.78 uIU/mL)       | 2.3 ± 1.7     | 2.2 ± 1.5          | 2.5 ± 2.1      | 0.482   |
| IRMA (0.17-4.05 uIU/L)         | 2.5 ± 1.8     | 2.8 ± 1.7          | 2.2 ± 1.8      | 0.054   |
| fT4 in relapse                 |               |                    |                |         |
| ECLIA (0.89-1.76 ng/dL)        | 2.6 ± 1.3     |                    |                |         |
| IRMA (0.89-1.79 ng/dL)         | 2.5 ± 0.7     |                    |                |         |
TSH at relapse

| Method         | 0.55-4.78 uIU/mL | 0.1 ± 0.4 |
|----------------|------------------|-----------|
| ECLIA (0.17-4.05 uIU/L) | 0.1 ± 0.1 |

Thyroid autoantibodies

**TBII**

| Method         | 1.0 ± 1.1 | 1.0 ± 1.1 | 1.2 ± 1.0 | 0.372 |
|----------------|-----------|-----------|-----------|-------|
| ECLIA (> 1.75 IU/L) | 3.5 ± 6.5 | 4.1 ± 7.0 | 2.8 ± 5.7 | 0.270 |
| RAI (>1.5 IU/L)   | 44 (20.1%) | 24 (18.0%) | 20 (23.3%) | 0.443 |
| TSI (%)          | 68.9 ± 28.0 | 67.3 ± 29.0 | 71.4 ± 26.5 | 0.290 |

GD, Graves’ disease; ATD, antithyroid drug; PTU, propylthiouracil; fT4, free T4; ECLIA, electrochemiluminescence immunoassay; IRMA, immunoradiometric assay; TSH, thyrotropin; TBII, inhibitory immunoglobulin binding to thyrotropin; TSI, thyroid-stimulating immunoglobulin; R, relapse group; N-R, no relapse group
Table 2. Odds ratio to relapse of Graves’ disease according to each cut-off value of TSI and TBII

|                          | Odds ratio | 95% CI       | P-value |
|--------------------------|------------|--------------|---------|
| TBII positivity          | 1.376      | 0.706-2.683  | 0.347   |
| Cut-off based on the mean of TSI (66.87) | 2.063      | 1.189-3.579  | 0.010   |
| Cut-off based on the median of TSI (66) | 1.690      | 0.977-2.923  | 0.060   |
| Cut-off based on ROC curve (66.5) | 2.063      | 1.189-3.79   | 0.010   |

TSI, thyroid-stimulating immunoglobulin; TBII, thyrotropin-binding inhibitory immunoglobulin.
Table 3. Logistic regression analysis of risk factors for relapse of Graves’ disease relapse

|                      | Univariate          |          | Multivariate        |          |
|----------------------|---------------------|----------|---------------------|----------|
|                      | HR                  | 95% CI   | P-value             | HR       | 95% CI   | P-value             |
| Age                  | 0.980               | 0.960-1.001 | 0.056   | 0.980               | 0.957-1.003 | 0.052   |
| Male to female       | 2.567               | 1.403-4.698 | 0.002   | 2.476               | 1.277-4.803 | 0.007   |
| Graves’ orbitopathy  | 0.327               | 0.547-3.221 | 0.532   | 1.046               | 0.388-1.046 | 0.0929  |
| ATD regimen\(^a\)    |                     |          |                     |          |
| Carbimazole          | 0.228               | 0.090-0.573 | 0.002   | 0.234               | 0.089-0.615 | 0.003   |
| PTU                  | 0.607               | 0.199-1.852 | 0.380   | 0.647               | 0.197-2.124 | 0.473   |
| Duration of ATD usage| 1.000               | 0.991-1.010 | 0.997   | 1.003               | 0.992-1.013 | 0.630   |
| Concomitant T4       | 0.738               | 0.406-1.344 | 0.321   | 0.714               | 0.370-1.379 | 0.316   |
| replacement during ATD|                    |          |                     |          |
| TBII positivity      | 1.376               | 0.706-2.683 | 0.348   | 1.186               | 0.568-2.615 | 0.611   |
| TSI positivity based on ROC curve | 2.063 | 1.189-3.579 | 0.010   | 1.992               | 1.095-3.623 | 0.022   |

\(^a\)all compared to methimazole

HR, hazard ratio; ATD; PTU, propylthiouracil; TSI, thyroid-stimulating immunoglobulin; TBII, inhibitory immunoglobulin; ROC, receiver operating characteristic curve.
Table 4. Clinical characteristics according to medication at withdrawal and propensity score correlated with age and sex results

|                           | Before Propensity Score Matching | After Propensity Score Matching |
|---------------------------|----------------------------------|---------------------------------|
|                           | MZ  | CM  | P-value | standardized difference | MZ  | CM  | P-value | standardized difference |
|                           | N=166 | N=38 |         |                     | N=38 | N=38 |         |                     |
| Male sex                  | 53  | 8   | 0.261   | -0.267             | 8 (21.1%) | 8 (21.1%) | 1.000 | 0.000 |
|                           | (31.9%) | (21.1%) |     |                    |                             |                             |                   |
| Age                       | 45.5 ± 13.5 | 46.5 ± 13.4 | 0.703 | 0.069             | 46.2 ± 13.5 | 46.5 ± 13.4 | 0.919 | 0.024 |
|                           | 13.5 | 13.4 |                     |                     |                             |                             |                   |
| Duration of ATD usage     | 29.8 ± 29.5 | 36.1 ± 26.3 | 0.229 |                     | 22.6 ± 12.2 | 36.1 ± 26.3 | 0.006 |
|                           | (month) |                     |                     |                     |                             |                             |                   |
| Concomitant T4 replacement during ATD | 58  | 6   | 0.036   |                     | 16 (42.1%) | 6 (15.8%) | 0.023 |
|                           | (34.9%) | (15.8%) |     |                    |                             |                             |                   |
| Dose when ATD stop (mg)   | 3.1 ± 1.4 | 2.3 ± 1.6 | 0.003 |                     | 3.0 ± 1.3 | 2.3 ± 1.6 | 0.031 |
| Thyroid autoantibodies    |     |     |         |                     |                             |                             |                   |
| TBII                      |     |     |         |                     |                             |                             |                   |
| ECLIA                     | 1.2 ± 0.7 | 1.3 ± 0.4 | 0.009 | 1.1 ± 1.0 | 0.7 ± 0.4 | 0.327 |
| (>1.75IU/L)               |     |     |         |                     |                             |                             |                   |
| RAI                       | 3.2 ± 0.9 | 6.2 ± 0.7 | <0.001 | 3.0 ± 5.2 | 0.9 ± 0.7 | 0.040 |
| (>1.5IU/L)                |     |     |         |                     |                             |                             |                   |
|               | Count | Percentage | Count | Percentage | Count | Percentage |
|---------------|-------|------------|-------|------------|-------|------------|
| TBII positivity | 35    | 2          | 9 (23.7%) | 2 (5.3%)   | 0.040 |
|               |       |            |        |            |       |            |
| TSI           | 68.1 ± 70.7 ± | 70.6 ± 28.9 | 70.7 ± 31.5 | 0.979 |
| TSI positivity according to cut-off with 66.5 | 79 | 15 | 18 (47.4%) | 15 (39.5%) | 0.468 |
| Relapse of disease | 75 | 6 | 17 (44.7%) | 6 (15.8%) | 0.002 |

ATD; MZ, methimazole; CM, carbimazole; TBII, thyrotropin-binding inhibitory immunoglobulin; ECLIA, electrochemiluminescence immunoassay; IRMA, immunoradiometric assay; TSI, thyroid-stimulating immunoglobulin
Figure 1

Graves' disease patients who were stopped ATD after checking their TSI 40x40px

Exclusion (N=428)

1. No follow-up data (N=102)
2. The start point of ATD is unclear (N=98)
3. Final diagnosis was not Graves' disease (N=66)
4. Never stopped ATD or duration from ATD stop to TSI measurement > 3 months (N=158)
5. RAI or surgery (N=3)
6. No TBI data (N=1)

N=243

ATD stop duration < 3 months (N=5)

N=86

Follow-up duration < 1 year (N=10)

N=133

Finally analyzed (N=219)
Figure 2

Study population: patients whose TSI bioassay < 140%, so the physician stopped the ATD.

| Stop ATD | 1 year | 2 year | 3 year |
|----------|--------|--------|--------|
| N=24     |        |        |        |
| N=48     |        |        |        |
| N=61     |        |        |        |
| N=133    |        |        |        |
| N=43     |        |        |        |
| N=23     |        |        |        |
| N=86     |        |        |        |

No Relapse Group

Relapse Group

* All patients were observed at least 1 year after ATD stop
Figure 3

Cumulative Relapse Rate (%) vs. Months after ATD stop

- Green line: TSI < 66.5
- Red line: TSI ≥ 66.5

P = 0.049
Figure 4

A

Cumulative Relapse Rate (%)

0 50 100 150

Months after ATD stop

TSI < 59.5

TSI > 59.5

p = 0.041

B

Cumulative Relapse Rate (%)

0 50 100 150

Months after ATD stop

TSI < 66.5

TSI > 66.5

p = 0.573
Figure 5

A

Risk for Relapse of Graves' Disease

B

Risk for Relapse of Graves' Disease

Total-points-to-outcome nanoparameters:

Total-points-to-outcome nanoparameters:
Figure 7

The Prediction Model Using Thyroid Stimulating Immunoglobulin Bioassay For Relapse of Graves' Disease

Method
- Study population: patients whose TSI bioassay < 140%, so the physician stopped the ATD.

| Study Population | 1 year | 2 year | 3 year |
|------------------|--------|--------|--------|
| TSI ≤ 40         |        |        |        |
| TSI > 40         |        |        |        |

Result
- The Prediction Model Using TSI Bioassay For Relapse of Graves' Disease

Nomogram
- TSI 67.2 score (HR 2.007, with cut-off 66.5%)
- TBI 22.8 score (HR 1.166, positive: more relapse)
- Sex 100 score (HR 2.959, Male: more relapse)
- Age 62.6 score (HR 0.0978, older age: less relapse)

*Sensitivity 57%, Specificity 70.7%, Ability to predict of Graves' relapse 67.2%

Key findings
- The cut-off value of the TSI bioassay to predict GD relapse should be lower than that for diagnosing GD.
- Nomogram with multiple risk factors including TSI bioassay could be helpful to predict GD relapse.