



Dynamic Risk Model for the Medical Treatment of Graves' Hyperthyroidism according to Treatment Duration

Meihua Jin¹, Chae A Kim², Min Ji Jeon², Won Bae Kim², Tae Yong Kim², Won Gu Kim²

¹Division of Endocrinology and Metabolism, Department of Internal Medicine, Dankook University College of Medicine, Cheonan; ²Division of Endocrinology and Metabolism, Department of Internal Medicine, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Korea

Background: Changes in thyrotropin receptor antibody (TRAb) levels are associated with the clinical outcomes of Graves' hyperthyroidism. However, the effects of the patterns of TRAb changes on patient prognosis according to the treatment duration of anti-thyroid drugs (ATDs) are not well established.

Methods: In this retrospective cohort study, 1,235 patients with Graves' hyperthyroidism who were treated with ATDs for more than 12 months were included. Patients were divided into two groups according to treatment duration: group 1 (12–24 months) and group 2 (>24 months). Risk prediction models comprising age, sex, and either TRAb levels at ATD withdrawal (model A) or patterns of TRAb changes (model B) were compared.

Results: The median treatment duration in groups 1 ($n=667$, 54%) and 2 ($n=568$, 46%) was 17.3 and 37.1 months, respectively. The recurrence rate was significantly higher in group 2 (47.9%) than in group 1 (41.4%, $P=0.025$). Group 2 had significantly more goiter, thyroid eye disease, and fluctuating and smoldering type of TRAb pattern compared with group 1 (all $P<0.001$). The patterns of TRAb changes were an independent risk factor for recurrence after adjusting for other confounding factors in all patients, except in group 1. Integrated discrimination improvement and net reclassification improvement analyses showed that model B performed better than model A in all patients, except in group 1.

Conclusion: The dynamic risk model, including the patterns of TRAb changes, was more suitable for predicting prognosis in patients with Graves' hyperthyroidism who underwent longer ATD treatment duration.

Keywords: Graves hyperthyroidism; Antithyroid drug; Thyrotropin receptor antibody; Recurrence; Risk model

Received: 27 December 2023, Revised: 7 March 2024,
Accepted: 26 March 2024

Corresponding authors: Tae Yong Kim
Division of Endocrinology and Metabolism, Department of Internal Medicine,
Asan Medical Center, University of Ulsan College of Medicine, 88 Olympic-ro
43-gil, Songpa-gu, Seoul 05505, Korea
Tel: +82-2-3010-3249, Fax: +82-2-3010-6962, E-mail: tykim@amc.seoul.kr

Won Gu Kim
Division of Endocrinology and Metabolism, Department of Internal Medicine,
Asan Medical Center, University of Ulsan College of Medicine, 88 Olympic-ro
43-gil, Songpa-gu, Seoul 05505, Korea
Tel: +82-2-3010-5883, Fax: +82-2-3010-6962, E-mail: wongukim@amc.seoul.kr

Copyright © 2024 Korean Endocrine Society

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

INTRODUCTION

The recommended treatment duration for antithyroid drugs (ATDs) for patients with Graves' hyperthyroidism is approximately 12 to 18 months [1-3]. However, the optimal treatment duration of ATDs is still being debated. A previous meta-analysis showed that the remission rate did not differ significantly between patients treated for up to 18 months and those treated for over 18 months [3]. By contrast, randomized clinical trials and real-world studies have shown that a longer treatment duration is associated with a higher remission rate [4-6]. Considering the various risk factors of patients with Graves' hyperthyroidism, discussing clinical outcomes solely based on treatment duration is challenging. In real-world clinical practice, the treatment duration is usually adjusted by considering various clinical situations [7,8]. In a recent real-world study, surgery or radioactive iodine (RAI) therapy was provided as first-line treatment for patients with younger age, larger goiter size, smoking history, and higher thyrotropin receptor antibody (TRAb) titers [7]. Similarly, while individual risk of recurrence is considered in the selection of treatment modality, questions have been raised regarding the characteristics of patients treated with ATDs for longer durations than those indicated by the guidelines.

The effects of the patterns of TRAb changes on the prognosis of patients with Graves' hyperthyroidism have been described in several studies [7,9-11]. Compared to those with smooth disappearance type of TRAb pattern, patients with fluctuating and smoldering types of TRAb patterns were treated for a longer duration; nevertheless, they exhibited a higher recurrence rate [7]. Furthermore, dynamic changes in TRAb levels, younger age, and male sex were identified as independent risk factors associated with recurrence-free survival (RFS) [7]. The patterns of TRAb changes in patients with Graves' hyperthyroidism according to treatment duration may have varied prognostic implications; however, little is known to date.

Regarding risk models for predicting recurrence, a prospective study proposed a predictive score called the Graves' Recurrent Events After Therapy (GREAT) score, which includes the initial TRAb titer [12]. The GREAT score could predict recurrence before ATD treatment. In addition, negative TRAb levels before ATD treatment discontinuation and dynamic changes in TRAb levels are important factors for predicting prognosis at the time of ATD withdrawal. Therefore, prediction models comprising either TRAb levels at ATD withdrawal or patterns of TRAb changes can be proposed.

The present study has several central aims. First, it aimed to

understand the clinical features of patients with Graves' hyperthyroidism who were treated for longer durations. Second, it evaluated the effects of the patterns of TRAb changes according to treatment duration. Finally, it suggested risk prediction models for patients with Graves' hyperthyroidism and compared the prediction accuracy of these models in patients with different treatment durations.

METHODS

Study subjects

Patients with Graves' hyperthyroidism who underwent ATD treatment as described previously were included in this retrospective cohort study [7]. A total of 1,235 patients who were newly diagnosed with Graves' hyperthyroidism and completed their first course of ATD treatment were included. Patients were treated with ATDs for at least 12 months and followed up for more than 12 months after treatment discontinuation. They were then divided into two groups according to the treatment duration: group 1 (12–24 months) and group 2 (>24 months). This study was approved by the Institutional Review Board of Asan Medical Center, Seoul, Korea (No. 2021-1121). Informed consent by the patients was waived due to a retrospective nature of our study.

Clinical definitions and treatment protocol

Graves' hyperthyroidism was diagnosed on the basis of a thyroid function test (TFT) with elevated TRAb levels and the findings of a thyroid scan. Patients' medical records were used to collect clinical features, including smoking status, goiter, and thyroid eye disease (TED). Smoking status was classified as nonsmoker, ex-smoker, or current smoker. Goiter was diagnosed at the time of initial physical examination by a physician and classified according to the World Health Organization goiter classification system (grade 0, no goiter; grade 1, thyroid palpable but not visible; and grade 2, thyroid visible with the neck in the normal position) [13]. TED was classified on the basis of disease severity: mild, moderate, or severe [14].

Depending on the severity of hyperthyroidism, physicians started with ATDs at initial diagnosis (methimazole [MMI], 15–30 mg/day; carbimazole [CMZ], 20–40 mg/day; or propylthiouracil [PTU], 100–400 mg/day) and used the dose titration method [15-17]. TFT and TRAb levels were measured at initial diagnosis and every 2 to 3 months during the treatment period. Before ATD discontinuation, minimum maintenance dose therapy (MMDT; MMI, 2.5 mg/day; CMZ, 5 mg/day; and PTU, 25

mg/day) was performed in most patients [16,18]. ATD treatment was discontinued when serum free thyroxine (fT4) and thyroid-stimulating hormone (TSH) levels were within the normal range. Regarding TRAb levels, most patients discontinued ATD treatment when TRAb levels were negative; nevertheless, ATD treatment was discontinued even if TRAb levels were positive when the ATD treatment period was sufficiently long in some cases.

A detailed description of the changes in TRAb levels during ATD treatment has been provided in a previous study [7]. TRAb patterns were categorized as smooth disappearance, fluctuating, and smoldering types. In the smooth disappearance type of TRAb pattern, TRAb levels decreased smoothly to become negative before ATD discontinuation. In the fluctuating type of TRAb pattern, TRAb titers fluctuated from positive to negative during ATD treatment. In the smoldering type of TRAb pattern, TRAb titers consistently remained positive during the course of treatment. Remission was defined when the euthyroid status was maintained for more than 12 months after ATD discontinuation [1,17], whereas recurrence was defined as persistent thyrotoxicosis (excluding transient thyrotoxicosis) during follow-up after ATD discontinuation [1,16]. At recurrence, TRAb levels were also measured. RFS was defined as the time from the date of ATD discontinuation until the date of recurrence or the last follow-up.

The risk scores and prediction models in patients with Grave's hyperthyroidism

First, risk factors associated with RFS were evaluated using a Cox proportional hazards model. Second, two risk prediction models were constructed using the factors in the multivariate analysis. Independent and nonindependent risk factors associated with RFS presented different scores. The two models comprised the same risk factors; however, model A included TRAb levels at ATD withdrawal, whereas model B included the patterns of TRAb changes.

Laboratory test

Competitive thyrotropin-binding inhibitory immunoglobulin (TBII) assay was performed using the B·R·A·H·M·S TRAK human radioimmunoassay (RIA; B·R·A·H·M·S GmbH, Hennigsdorf/Berlin, Germany) to measure TRAb levels [7,16]. TBII titers ≥ 1.5 IU/L were considered positive. Serum TSH levels were measured using the TSH-CTK-3 kit (RIA; DiaSorin S. p.A., Saluggia, Italy) with a reference range of 0.4 to 4.5 mIU/L [19]. Meanwhile, serum fT4 levels were measured using the fT4

RIA (Immunotech, Prague, Czech Republic) with a reference range of 0.80 to 1.90 ng/dL [7,17].

Statistical analysis

Statistical analyses were performed using R version 3.4.4 (R Foundation for Statistical Computing; www.R-project.org). Continuous and categorical variables were presented as median (interquartile range [IQR]) and number (percentage), respectively. The Wilcoxon rank-sum and chi-square tests were used for comparisons between groups. A Cox proportional hazards model was used to evaluate the prognostic factors associated with RFS and presented as a forest plot. The relative risk for RFS was presented as a hazard ratio (HR) with 95% confidence interval (CI). Kaplan-Meier curves were used to obtain the RFS, and their significance were determined using the log-rank test. The proportion of variation explained (PVE) was calculated to determine the predictive value of the dynamic risk models in patients with Graves' hyperthyroidism. The PVE (%) ranges from 0 to 100, with larger numbers indicating a more accurate predictive model for discriminating the outcome. Moreover, integrated discrimination improvement (IDI) and net reclassification improvement (NRI) were used to compare the models' prediction accuracy. Statistical significance was considered at $P < 0.05$.

RESULTS

Clinical features of patients with Graves' hyperthyroidism according to treatment duration

Table 1 shows the clinical features of patients with Graves' hyperthyroidism according to treatment duration. The median age of all patients was 45.9 years, and 31.2% were male. The number of patients in groups 1 and 2 was 667 (54%) and 568 (46%), respectively. There were no significant differences in terms of age, sex, and smoking status between the two groups. Goiter and TED were frequently found in group 2 than in group 1 ($P < 0.001$ and $P < 0.001$, respectively). Moreover, group 2 had higher TRAb titers at initial diagnosis and at the end of treatment compared with group 1 ($P < 0.001$ and $P < 0.001$, respectively). More patients in group 1 discontinued ATD treatment when their TRAb levels were negative compared with those in group 2 (79.2% vs. 68.5%, $P < 0.001$). The patterns of TRAb changes were significantly different between the two groups ($P < 0.001$). Group 2 showed more fluctuating (32.6%) and smoldering (23.4%) types of TRAb patterns, whereas group 1 showed more smooth disappearance (75.9%) type of TRAb pattern. The median time required for TRAb normalization was

Table 1. Clinical Features of Patients with Graves' Hyperthyroidism according to the Treatment Duration of Antithyroid Drugs

Variable	Total (<i>n</i> =1,235)	Group 1 (<i>n</i> =667) (12–24 mo)	Group 2 (<i>n</i> =568) (>24 mo)	<i>P</i> value
Age, yr	45.9 (35.2–55.0)	47.2 (36.4–55.0)	44.7 (33.7–55.1)	0.056
Male sex	385 (31.2)	203 (30.4)	182 (32.0)	0.585
Smoking				
None	1,172 (94.9)	631 (94.6)	541 (95.2)	0.702
Ex- or current	63 (5.1)	36 (5.4)	27 (4.8)	
Goiter ^a				
Grade 0	556 (45.0)	330 (49.5)	226 (39.8)	0.001
Grade 1–2	679 (55.0)	337 (50.5)	342 (60.2)	
TED				
None	1,073 (86.9)	606 (90.9)	467 (82.2)	<0.001
Mild to severe	162 (13.1)	61 (9.1)	101 (17.8)	
ATD				
MMI	688 (55.7)	399 (59.8)	289 (50.9)	0.006
CMZ	461 (37.3)	228 (34.2)	233 (41.0)	
PTU	86 (7.0)	40 (6.0)	46 (8.1)	
Initial dose of ATD (MMI), mg ^b	13.3 (10.0–20.0)	13.3 (10.0–20.0)	13.3 (10.0–20.0)	0.571
Initial fT4, ng/dL	2.5 (2.0–3.4)	2.5 (1.0–3.2)	2.6 (1.9–3.5)	0.686
Initial TRAb, IU/L	8.9 (4.7–17.5)	6.9 (3.9–11.7)	13.3 (6.7–26.8)	<0.001
TRAb level at withdrawal, IU/L	1.1 (0.6–1.6)	1.0 (0.6–1.5)	1.2 (0.7–2.0)	<0.001
TRAb at withdrawal (negative)	917 (74.3)	528 (79.2)	389 (68.5)	<0.001
Pattern of TRAb changes				
Smooth disappearance	756 (61.2)	506 (75.9)	250 (44.0)	<0.001
Fluctuating	241 (19.5)	56 (8.4)	185 (32.6)	
Smoldering	238 (19.3)	105 (15.7)	133 (23.4)	
TRAb normalization, mo ^c	13.5 (7.2–22.2)	9.3 (5.7–14.4)	23.8 (14.2–36.8)	<0.001
MMDT (yes)	1,024 (82.9)	551 (82.6)	473 (83.3)	0.815
Duration of MMDT, mo	10.1 (6.2–15.2)	9.0 (5.8–11.8)	13.8 (8.3–22.2)	<0.001
TSH normalization, mo	5.3 (3.2–10.9)	4.6 (3.0–6.9)	7.8 (3.7–17.1)	<0.001
Duration of ATD, mo	23.0 (17.0–35.5)	17.3 (14.9–20.0)	37.1 (29.1–52.3)	<0.001
Recurrence	548 (44.4)	276 (41.4)	272 (47.9)	0.025

Values are expressed as median (interquartile range) or number (%).

TED, thyroid eye disease; ATD, antithyroid drug; MMI, methimazole; CMZ, carbimazole; PTU, propylthiouracil; fT4, free thyroxine; TRAb, thyrotropin receptor antibody; MMDT, minimum maintenance dose therapy.

^aWorld Health Organization goiter classification system: grade 0, no goiter; grade 1, thyroid palpable but not visible; and grade 2, thyroid visible with the neck in the normal position; ^bA 7.5-mg CMZ and 100-mg PTU dose is equivalent to 5 mg of MMI; ^cThe duration of TRAb normalization was measured in patients with smooth remission and fluctuating types of TRAb patterns.

shorter in group 1 (9.3 months) than in group 2 (23.8 months, $P<0.001$). Most patients underwent MMDT, and group 2 had a longer MMDT duration than group 1 ($P<0.001$). Although no significant difference was observed in the initial level of fT4 and initial dose of ATD between the two groups ($P=0.686$ and $P=0.571$, respectively), the median time required for the normaliza-

tion of TSH levels was significantly longer in group 2 (7.8 months) than in group 1 (4.6 months) ($P<0.001$). Group 2 had a median treatment duration of 37.1 months (IQR, 29.1 to 52.3), which was significantly longer than that of group 1 (17.3 months [IQR, 14.9 to 20.0]) ($P<0.001$). However, group 2 had a higher recurrence rate (47.9%) compared with group 1 (41.4%, $P<$

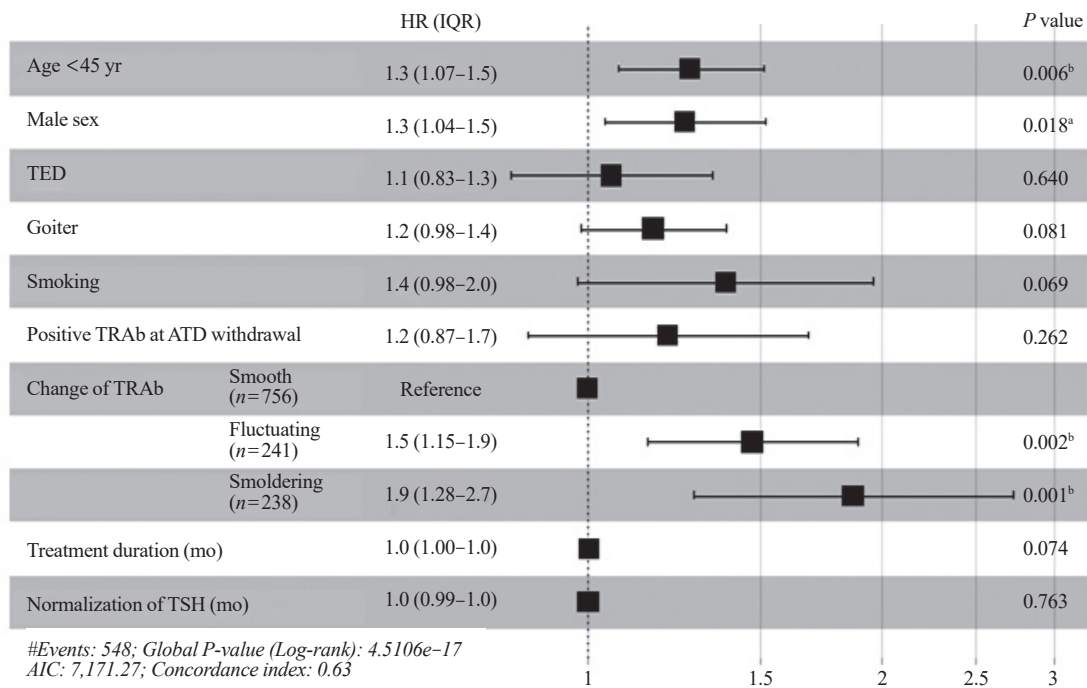


Fig. 1. A forest plot of the Cox proportional hazards model for recurrence-free survival in patients with Graves' hyperthyroidism is shown. HR, hazard ratio; IQR, interquartile range; TED, thyroid eye disease; TRAb, thyrotropin receptor antibody; ATD, antithyroid drug; TSH, thyroid-stimulating hormone; AIC, Akaike information criterion. ^a $P < 0.05$; ^b $P < 0.01$.

Table 2. Hazard Ratios of TRAb Patterns for Recurrence-Free Survival Using Multivariate Analysis

Models	Total (n=1,235)		Group 1 (n=667) 12–24 mo		Group 2 (n=568) >24 mo	
	HR (95% CI)	P value	HR (95% CI)	P value	HR (95% CI)	P value
Unadjusted						
Smooth disappearance	1	-	1	-	1	-
Fluctuating	1.75 (1.42–2.17)	<0.001	1.71 (1.15–2.51)	<0.001	1.72 (1.29–2.30)	<0.001
Smoldering	2.26 (1.84–2.76)	<0.001	1.82 (1.36–2.43)	<0.001	2.62 (1.94–3.52)	<0.001
Adjusted model 1 ^a						
Smooth disappearance	1	-	1	-	1	-
Fluctuating	1.72 (1.39–2.12)	<0.001	1.78 (1.21–2.62)	0.004	1.72 (1.29–2.30)	<0.001
Smoldering	2.34 (1.91–2.86)	<0.001	1.90 (1.42–2.55)	<0.001	2.73 (2.02–3.68)	<0.001
Adjusted model 2 ^b						
Smooth disappearance	1	-	1	-	1	-
Fluctuating	1.69 (1.36–2.09)	<0.001	1.73 (1.17–2.57)	0.006	1.72 (1.29–2.30)	<0.001
Smoldering	2.33 (1.90–2.85)	<0.001	1.88 (1.40–2.53)	<0.001	2.74 (2.03–3.70)	<0.001
Adjusted model 3 ^c						
Smooth disappearance	1	-	1	-	1	-
Fluctuating	1.48 (1.15–1.89)	0.002	1.41 (0.89–2.25)	0.146	1.65 (1.20–2.26)	0.002
Smoldering	1.90 (1.30–2.78)	<0.001	1.30 (0.73–2.30)	0.366	2.69 (1.57–4.60)	<0.001

TRAb, thyrotropin receptor antibody; HR, hazard ratio; CI, confidence interval.

^aAdjusted model 1 was adjusted for age and sex; ^bAdjusted model 2 was adjusted for age, sex, goiter, thyroid eye disease (TED), and smoking status;

^cAdjusted model 3 was adjusted for age, sex, goiter, TED, smoking status, positive TRAb levels at antithyroid drug withdrawal, and treatment duration.

0.025). Moreover, group 2 had a significantly lower RFS compared with group 1 ($P<0.001$) (Supplemental Fig. S1).

Risk factors associated with RFS in patients with Graves' hyperthyroidism

The results of univariate analysis showed that younger age (<45 years), male sex, TED, goiter, smoking status, positive TRAb levels at ATD withdrawal, patterns of TRAb changes, treatment duration, and time required for TSH normalization were risk factors associated with RFS in 1,235 patients (Supplemental Table S1). The cutoff values for age, initial fT4, and TRAb were determined on the basis of the median values of all patients. Fig. 1 shows a forest plot for RFS obtained through multivariate analysis. As shown in the figure, age <45 years (HR, 1.3; 95% CI, 1.07 to 1.5; $P=0.006$), male sex (HR, 1.3; 95% CI, 1.04 to 1.5; $P=0.018$), and patterns of TRAb changes (fluctuating and smoldering type; HR, 1.5; 95% CI, 1.15 to 1.9; $P=0.002$; and HR, 1.9; 95% CI, 1.28 to 2.7; $P=0.001$, respectively) were independent risk factors associated with RFS in all patients. Moreover, the results of univariate analysis showed that TED, goiter, smoking, and TRAb levels at ATD withdrawal were associated with RFS (Supplemental Table S1), but they were not independent risk factors in multivariate analysis (Fig. 1).

Patterns of TRAb changes associated with RFS in groups 1 and 2

The effects of the patterns of TRAb changes on RFS were evaluated in each group and both groups combined (Table 2). When other risk factors were not adjusted, patients with fluctuating and smoldering types of TRAb patterns exhibited a significantly lower RFS than those with a smooth disappearance type of TRAb pattern in each group and both groups combined (all $P=0.001$). In group 1, TRAb patterns were still a significant factor when adjusting for age, sex, goiter, TED, and smoking status; however, they became nonsignificant when adjusting for positive TRAb levels at ATD withdrawal and treatment duration as described in adjusted model 3 (fluctuating type: HR, 1.41; 95% CI, 0.89 to 2.25; $P=0.146$; smoldering type: HR, 1.30; 95% CI, 0.73 to 2.30; $P=0.366$). However, TRAb patterns were significant risk factors after adjusting for all factors in both groups combined (fluctuating type: HR, 1.48; 95% CI, 1.15 to 1.89; $P=0.002$; smoldering type: HR, 1.90; 95% CI, 1.30 to 2.78; $P<0.001$) and group 2 (fluctuating type: HR, 1.65; 95% CI, 1.20 to 2.26; $P=0.002$; smoldering type: HR, 2.69; 95% CI, 1.57 to 4.60; $P<0.001$).

Table 3. Risk Score and Model for Predicting the Recurrence of Graves' Hyperthyroidism

Risk factor	Score	
	Model A	Model B
Age, yr		
≥45	0	0
<45	1	1
Sex		
Female	0	0
Male	1	1
Goiter		
Grade 0	0	0
Grade 1–2	0.5	0.5
TED		
None	0	0
Mild to severe	0.5	0.5
Smoking		
None	0	0
Ex- or current	0.5	0.5
TRAb at ATD withdrawal		
Negative	0	-
Positive	0.5	
Pattern of TRAb changes		
Smooth disappearance	-	0
Fluctuating		1
Smoldering		2
Total	4 points	5.5 points
Class	Class I: 0–1.0 points	
	Class II: 1.5–2.5 points	
	Class III: 3.0–5.5 points	

TED, thyroid eye disease; TRAb, thyrotropin receptor antibody; ATD, antithyroid drug.

Dynamic risk score and model for the recurrence of Graves' hyperthyroidism

As described in the Methods section, multivariate analysis was used to present scores, and two recurrence risk prediction models were proposed (Table 3). Independent risk factors including age, sex, and patterns of TRAb changes were expressed as a score of 1 or 2. Nonindependent risk factors including goiter, TED, smoking, and TRAb levels at ATD withdrawal were expressed as a score of 0.5. Model A comprised age, sex, goiter, TED, smoking, and simple TRAb levels before ATD withdrawal, whereas model B comprised dynamic TRAb changes (Table 3). The maximum total scores for models A and B were 4 and

Table 4. Risk Models in Patients with Graves' Hyperthyroidism according to Treatment Duration

	Models	Class	No. (%)	HR (95% CI)	<i>P</i> value	PVE, %	IDI* (95% CI) <i>P</i> value	NRI* (95% CI) <i>P</i> value
Total (group 1+group2) (<i>n</i> =1,235)	Model A	Class I	609 (49.3)	1	-	2.9		
		Class II	566 (45.8)	1.60 (1.34–1.91)	0.001		-	-
		Class III	60 (4.9)	2.23 (1.56–3.19)	<0.001			
	Model B	Class I	451 (36.5)	1	-	6.4	0.034 (0.017 to 0.051)	0.214 (0.131 to 0.268)
		Class II	567 (45.9)	1.69 (1.38–2.07)	<0.001		<0.001	<0.001
		Class III	217 (17.6)	2.95 (2.34–3.72)	<0.001			
Group 1 (12–24 mo) (<i>n</i> =667)	Model A	Class I	359 (53.8)	1	-	3.0		
		Class II	281 (42.1)	1.64 (1.29–2.10)	<0.001		-	-
		Class III	27 (4.0)	2.40 (1.40–4.11)	0.001			
	Model B	Class I	305 (45.7)	1	-	4.5	0.013 (–0.003 to 0.033)	0.111 (–0.038 to 0.021)
		Class II	279 (41.8)	1.64 (1.27–2.15)	<0.001		0.126	0.060
		Class III	83 (12.4)	2.56 (1.82–3.58)	<0.001			
Group 2 (>24 mo) (<i>n</i> =568)	Model A	Class I	250 (44.0)	1	-	2.1		
		Class II	228 (40.1)	1.49 (1.14–1.94)	0.004		-	-
		Class III	90 (15.8)	1.64 (1.18–2.30)	0.003			
	Model B	Class I	146 (25.7)	1	-	5.9	0.033 (0.007 to 0.059)	0.163 (0.064 to 0.246)
		Class II	328 (57.7)	1.84 (1.32–2.57)	<0.001		0.013	0.013
		Class III	94 (16.5)	3.14 (2.14–4.61)	<0.001			

HR, hazard ratio; CI, confidence interval; PVE, proportion of variation explained; IDI, integrated discrimination improvement; NRI, net reclassification improvement.

*The IDI and NRI of model B for predicting recurrence compared with model A at 24 months after the discontinuation of antithyroid drugs.

5.5 points, respectively. Patients were classified into three classes: I (score 0–1.0), II (score 1.5–2.5), and III (score 3.0–5.5).

The predictive value of each model was evaluated in both groups (Table 4, Fig. 2). In all patients, the RFS of classes II and III was significantly lower than that of class I in both models (Fig. 2A, B). On the basis of model B, the HRs of classes II and III were 1.69 (95% CI, 1.38 to 2.07; $P<0.001$) and 2.95 (95% CI, 2.34 to 3.72; $P<0.001$), respectively. Model B had a higher PVE (6.4%) than model A (2.9%) in all patients. In the IDI and NRI analyses at 24 months, model B performed better than model A ($P<0.001$ and $P<0.001$, respectively). In the subgroup analysis of group 1, the RFS in both models is shown in Fig. 2C, D. Model B had a slightly higher PVE (4.5%) than model A (3.0%), but the IDI and NRI assessed at 24 months were not significant in group 1. In the subgroup analysis of group 2, the RFS of classes II and III was significantly lower than that of class I in both models (Fig. 2E, F). On the basis of models A and B, the HRs of class III were 1.84 (95% CI, 1.32 to 2.57) and 3.14 (95% CI, 2.14 to 4.61), respectively. Model B had a higher PVE (5.9%) than model A (2.1%). In terms of IDI and NRI, model B performed

better than model A ($P=0.013$ and $P=0.013$, respectively). Moreover, the IDIs and NRIs assessed at 12 and 36 months showed the same results in both groups combined, group 1 alone, and group 2 alone (data not shown).

DISCUSSION

In the present study, the median treatment duration of group 1 was 17.3 months, which is similar to that indicated in American and European guidelines (12 to 18 months) [1,2]. The patients in group 2 were found to have more risk factors and had a higher recurrence rate despite being treated with ATDs for a longer period of time (median, 37 months). The patterns of TRAb changes were an independent risk factor for predicting recurrence after adjusting for other confounding factors in group 2. However, it was not an independent risk factor in group 1. These findings suggested that the patterns of TRAb changes during follow-up should be considered as an important factor in deciding whether to continue ATDs or apply definite treatment for patients who underwent ATD treatment for more than 24

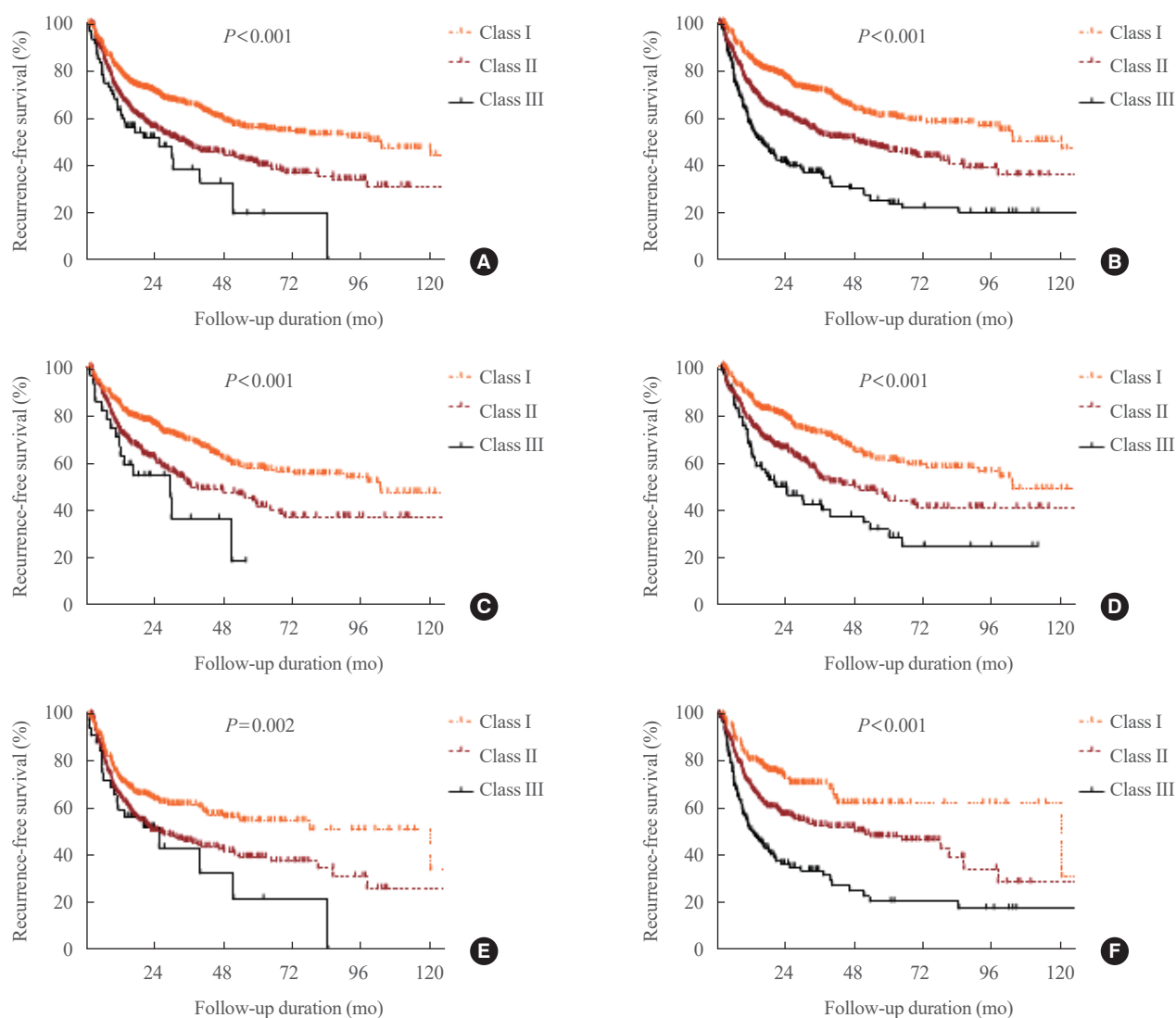


Fig. 2. A comparison of the models for the recurrence of Graves' hyperthyroidism according to treatment duration is shown. (A) It shows the total number of patients using model A. (B) It shows the total number of patients using model B. (C) It shows group 1 patients using model A. (D) It shows group 1 patients using model B. (E) It shows group 2 patients using model A. (F) It shows group 2 patients using model B.

months. Meanwhile, TRAb levels can be considered for ATD withdrawal in patients who can discontinue ATD treatment within 24 months.

ATDs are increasingly being used as the initial treatment modality worldwide [20-22]. Therefore, it is reasonable to assume that the clinical spectrum of patients treated with ATDs has expanded compared with earlier times. Because the treatment period can be unknowingly adjusted considering the clinical situations by physicians in real-world practice, the clinical characteristics of patients who are receiving ATD treatment longer than the indicated guidelines need to be investigated. In the present study, approximately 46% of patients were treated with ATDs

for more than 24 months. Previously, patients who underwent surgery or RAI therapy as first-line treatment ($n=192$) were younger, were smokers, had a larger goiter size, and had a higher TRAb titer compared with the 1,235 patients treated with ATDs as first-line treatment. In a similar context, the present study suggested that longer treatment duration was associated with goiter size, TED, and more fluctuating and smoldering types of TRAb patterns. Notably, group 2 had significantly longer time required for the normalization of serum TSH levels (median, 7.8 months) and MMDT duration (13.8 months) compared with group 1 (median, 4.6 and 9.0 months, respectively). Group 2 also had a higher recurrence rate even with longer

treatment duration compared with group 1 with fewer risk factors. This finding was inconsistent with that of a previous Korean study, which reported a lower recurrence rate after longer ATD treatment [6]. It would be appropriate to tailor the treatment duration according to individual risk factors rather than insisting on a uniform or longer treatment period.

The association between changes in TRAb levels and treatment duration is difficult to interpret. The patterns of TRAb changes were an important risk factor in both groups before adjusting for other risk factors. However, they were independent risk factors associated with the recurrence of Graves' hyperthyroidism after adjusting for other confounding factors such as TRAb at ATD withdrawal and treatment duration only in group 2. This finding might be explained by the different proportions of the three TRAb patterns in the two groups. Group 2 had more fluctuating (32.6%) and smoldering types (23.4%) of TRAb changes than group 1 (8.4% and 15.7%, respectively). In addition, if patients with relatively fewer risk factors (group 1) were treated for a sufficient period, the prognostic effect of patterns of TRAb changes on recurrence might be minimized.

To assess individual risk factors for the recurrence of Graves' hyperthyroidism, Vos et al. [12] introduced the GREAT score by combining age, goiter, initial fT4 level, and initial TRAb titer to predict the recurrence of Graves' hyperthyroidism before ATD treatment. In a recent Korean study, a prediction model using a thyroid-stimulating immunoglobulin bioassay along with age, sex, and TBII positivity was proposed [23]. However, these models are helpful at the early stage of ATD treatment and do not reflect the course or the last stage of treatment. In the present study, a relatively simple model that can be easily used in clinical practice was proposed. Furthermore, one strength of this model is that dynamic changes in TRAb during ATD treatment were included. The prediction model of this study can be used by performing a simple scoring system at the time of considering ATD discontinuation in patients who have been treated for more than 12 months. Based on the predictive values of models in the two groups, model B, which includes TRAb changes, may be more useful in patients with a relatively longer treatment duration. Meanwhile, model A, which includes TRAb levels at ATD withdrawal, may be more used in patients with low risk factors and who were treated with a relatively short duration. Future multicenter and prospective studies with a large cohort are needed to verify the utility of these models.

Guidelines suggest that patients should be counseled regarding treatment alternatives, including another course of ATD, RAI therapy, or surgery, upon recurrence after a course of ATD

treatment [1]. However, when and in which patients should be switched to another treatment modality during ATD treatment remain controversial. In the present study, among 568 patients in group 2, 94 (16.5%) were classified as class III on the basis of model B. These patients showed significantly lower RFS than other patients after ATD discontinuation (Fig. 2F). Therefore, these patients should maintain ATD treatment or be switched to a definite treatment such as RAI therapy or surgery rather than discontinuing ATD. The prediction model, including the patterns of TRAb changes, will be helpful in selecting patients at high risk of recurrence even if they are treated for more than 24 months.

This study has some limitations. First, this was a single-center, retrospective study, and selection bias might exist. In addition, TRAb levels may be affected by various factors such as systemic steroids for treating TED. However, patients in group 2 who had more severe TED showed more smoldering type of TRAb changes compared with group 1. Therefore, the effect of steroids on TRAb might be relatively minimal. Second, the changes in TRAb and treatment duration have a complex relationship, and interaction effects might exist. The dynamic changes in TRAb can be assessed more precisely with medical advances, which could influence the treatment duration. The strength of the present study is that it accurately reflected the complexity of real-world data and proposed a relatively simple risk prediction model.

In conclusion, the patterns of TRAb changes have a greater predictive value in patients with Graves' hyperthyroidism who underwent a longer duration of treatment. A predictive model that includes changes in TRAb levels is useful in patients with Graves' hyperthyroidism. For patients who underwent ATD treatment for more than 24 months, the prediction model might be helpful in selecting patients who require ATD maintenance or need to be switched to another treatment modality.

CONFLICTS OF INTEREST

No potential conflict of interest relevant to this article was reported.

ACKNOWLEDGMENTS

This study was supported by a grant (2019-374) from the Asan Institute for Life Sciences, Asan Institute for Life Sciences, Asan Medical Center, Seoul, Korea.

AUTHOR CONTRIBUTIONS

Conception or design: T.Y.K., W.G.K. Acquisition, analysis, or interpretation of data: M.J., C.A.K., M.J.J., W.B.K., W.G.K. Drafting the work or revising: M.J., W.G.K. Final approval of the manuscript: T.Y.K., W.G.K.

ORCID

Meihua Jin <https://orcid.org/0000-0003-0196-631X>

Tae Yong Kim <http://orcid.org/0000-0003-4982-4441>

Won Gu Kim <http://orcid.org/0000-0002-8404-7759>

REFERENCES

- Ross DS, Burch HB, Cooper DS, Greenlee MC, Laurberg P, Maia AL, et al. 2016 American Thyroid Association guidelines for diagnosis and management of hyperthyroidism and other causes of thyrotoxicosis. *Thyroid* 2016;26:1343-421.
- Kahaly GJ, Bartalena L, Hegedus L, Leenhardt L, Poppe K, Pearce SH. 2018 European Thyroid Association guideline for the management of Graves' hyperthyroidism. *Eur Thyroid J* 2018;7:167-86.
- Abraham P, Avenell A, Park CM, Watson WA, Bevan JS. A systematic review of drug therapy for Graves' hyperthyroidism. *Eur J Endocrinol* 2005;153:489-98.
- Azizi F, Amouzegar A, Tohidi M, Hedayati M, Khalili D, Cheraghi L, et al. Increased remission rates after long-term methimazole therapy in patients with Graves' disease: results of a randomized clinical trial. *Thyroid* 2019;29:1192-200.
- Azizi F, Malboosbaf R. Long-term antithyroid drug treatment: a systematic review and meta-analysis. *Thyroid* 2017; 27:1223-31.
- Park SY, Kim BH, Kim M, Hong AR, Park J, Park H, et al. The longer the antithyroid drug is used, the lower the relapse rate in Graves' disease: a retrospective multicenter cohort study in Korea. *Endocrine* 2021;74:120-7.
- Jin M, Jang A, Kim CA, Kim TY, Kim WB, Shong YK, et al. Long-term follow-up result of antithyroid drug treatment of Graves' hyperthyroidism in a large cohort. *Eur Thyroid J* 2023;12:e220226.
- Chung JH. Antithyroid drug treatment in Graves' disease. *Endocrinol Metab (Seoul)* 2021;36:491-9.
- Bandai S, Okamura K, Fujikawa M, Sato K, Ikenoue H, Kitazono T. The long-term follow-up of patients with thionamide-treated Graves' hyperthyroidism. *Endocr J* 2019;66:535-45.
- Cooper DS. Long-term antithyroid drug treatment of patients with graves' disease. *Clin Thyroidol* 2019;31:230-3.
- Yu J, Baek HS, Jeong C, Jo K, Lee J, Ha J, et al. The early changes in thyroid-stimulating immunoglobulin bioassay over anti-thyroid drug treatment could predict prognosis of Graves' disease. *Endocrinol Metab (Seoul)* 2023;38:338-46.
- Vos XG, Endert E, Zwinderman AH, Tijssen JG, Wiersinga WM. Predicting the risk of recurrence before the start of antithyroid drug therapy in patients with Graves' hyperthyroidism. *J Clin Endocrinol Metab* 2016;101:1381-9.
- World Health Organization. Assessment of iodine deficiency disorders and monitoring their elimination: a guide for programme managers. 3rd ed. Geneva: WHO; 2007.
- Bartalena L, Baldeschi L, Dickinson AJ, Eckstein A, Kendall-Taylor P, Marcocci C, et al. Consensus statement of the European group on Graves' orbitopathy (EUGOGO) on management of Graves' orbitopathy. *Thyroid* 2008;18:333-46.
- Vaidya B, Wright A, Shuttleworth J, Donohoe M, Warren R, Brooke A, et al. Block & replace regime versus titration regime of antithyroid drugs for the treatment of Graves' disease: a retrospective observational study. *Clin Endocrinol (Oxf)* 2014;81:610-3.
- Park S, Song E, Oh HS, Kim M, Jeon MJ, Kim WG, et al. When should antithyroid drug therapy to reduce the relapse rate of hyperthyroidism in Graves' disease be discontinued? *Endocrine* 2019;65:348-56.
- Kwon H, Kim WG, Jang EK, Kim M, Park S, Jeon MJ, 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; 31:300-10.
- Konishi T, Okamoto Y, Ueda M, Fukuda Y, Harusato I, Tsukamoto Y, et al. Drug discontinuation after treatment with minimum maintenance dose of an antithyroid drug in Graves' disease: a retrospective study on effects of treatment duration with minimum maintenance dose on lasting remission. *Endocr J* 2011;58:95-100.
- Jeon MJ, Kim WG, Jang EK, Choi YM, Lee YM, Sung TY, et al. Thyroglobulin level in fine-needle aspirates for preoperative diagnosis of cervical lymph node metastasis in patients with papillary thyroid carcinoma: two different cutoff values according to serum thyroglobulin level. *Thyroid* 2015; 25:410-6.
- Ahn HY, Cho SW, Lee MY, Park YJ, Koo BS, Chang HS, et

- al. Prevalence, treatment status, and comorbidities of hyperthyroidism in Korea from 2003 to 2018: a nationwide population study. *Endocrinol Metab (Seoul)* 2023;38:436-44.
21. Burch HB, Burman KD, Cooper DS. A 2011 survey of clinical practice patterns in the management of Graves' disease. *J Clin Endocrinol Metab* 2012;97:4549-58.
22. Brito JP, Payne S, Singh Ospina N, Rodriguez-Gutierrez R, Maraka S, Sangaralingham LR, et al. Patterns of use, efficacy, and safety of treatment options for patients with Graves' disease: a nationwide population-based study. *Thyroid* 2020;30:357-64.
23. Baek HS, Lee J, Jeong CH, Lee J, Ha J, Jo K, et al. The prediction model using thyroid-stimulating immunoglobulin bioassay for relapse of Graves' disease. *J Endocr Soc* 2022;6:bvac023.