

Independent Pretreatment Predictors of Graves' Disease Outcome

Dalia Daukšienė¹, Albertas Daukša², Narseta Mickuvienė³

¹*Institute of Endocrinology, Medical Academy, Lithuanian University of Health Sciences,*

²*Institute for Digestive Research, Medical Academy, Lithuanian University of Health Sciences,*

³*Behavioral Medicine Institute, Lithuanian University of Health Sciences, Lithuania*

Key Words: Graves' disease; goiter; relapse; thyroid antibodies.

Summary. *Background and Objective.* Current treatment options of Graves' disease (GD) are often unsatisfactory. This study aimed at determining the independent baseline predictors of medical treatment failure in GD.

Material and Methods. A retrospective study of 194 patients with GD was carried out. According to the disease outcome, patients were divided into groups. The remission group included the patients who achieved long-term remission after initial antithyroid drug (ATD) treatment with no relapse (group 1) or after 2 or 3 courses of ATD therapy (group 2). The treatment failure group included the patients who underwent thyroid ablation due to relapse (group 3) or without ATD withdrawal (group 4).

Results. A family history of thyroid disorders was associated with greater odds of failure ($P=0.046$). Higher thyrotropin receptor antibodies (TRAb) levels and a larger goiter size (grade 2/grade 3) at the onset of the disease were both independently associated with a greater likelihood of failure. The initial TRAb concentration of 30.2 U/L and the TRAb concentration of 12.97 U/L at the end of ATD therapy were found to be the best cutoff values predicting the treatment failure. A hypoechogenic thyroid after ATD therapy, but not before therapy, increased the likelihood of failure by nearly 7.5 times ($P<0.001$).

Conclusions. Higher TRAb levels and a larger goiter size at the onset of the disease were found to be the independent predictors of medical treatment failure in GD.

Introduction

Autoimmune thyroid diseases (AITDs) are the most prevalent autoimmune endocrine disorders, affecting approximately 2%–5% of the general population (1). Genetic susceptibility in combination with environmental triggers, such as smoking, viral/bacterial infection, and chemicals, is thought to initiate the autoimmune response against thyroid antigens (2). In patients younger than 40 years, Graves' disease (GD) is the most common cause of hyperthyroidism, which occurs more frequently in women than men (3). Autoimmunity against a thyroid-stimulating hormone receptor (TSHR) is the main pathogenetic element of GD. Antibodies against the TSH receptor (TRAb) stimulate the growth and the function of the thyroid follicular cells leading to the excessive production of thyroid hormones and goiter formation (4). TRAb detection is often used for the differential diagnosis of GD in order to differentiate it from other causes of hyperthyroidism, for the follow-up of patients with GD during and after antithyroid

drug (ATD) therapy, and may predict the outcome of the disease (5). Currently, the main treatment strategies for GD include antithyroid drugs, thyroid ablation with radioiodine (RAI), or surgery. Therapy with ATD is the first choice treatment in Europe (6). The modalities of ATD treatment are considerably different among countries. A long-term therapy of about 12–18 months is usually administered, which requires careful monitoring of patients for side effects, such as rash, joint pain, hepatitis, and agranulocytosis (2). After the withdrawal of ATD therapy, the relapse rate is very high (30%–60%), and most of the patients need further treatment (6). Radioactive iodine is a safe and effective therapy, but sometimes it is associated with a difficult decision to give a high enough dose for the treatment to be effective and to minimize the chance of recurrence and at the same time to reduce the risk of hypothyroidism (7). In patients with moderate-to-severe thyroid eye disease, radioiodine may worsen pre-existing ophthalmopathy. Rare but major complications include hypoparathyroidism and thyrotoxic crisis (2, 8). Radioiodine therapy is sometimes associated with a long latency period before its effect is seen (9). Most patients achieve an euthyroid or a hypothyroid state after a

Correspondence to D. Daukšienė, Institute of Endocrinology, Medical Academy, Lithuanian University of Health Sciences, Eivenių 2, 50161 Kaunas, Lithuania
E-mail: dalia.dauksiene@gmail.com

single dose of RAI, and approximately 10%–30% of patients require more than one dose of RAI (10). Thyroid surgery is related with high hospitalization costs and the risk of surgery-related complications such as the lesions of the parathyroid glands and recurrent laryngeal nerves (11).

As the response to the treatment of GD is unpredictable, it is often problematic to choose the optimal treatment approach. An endocrinologist must weigh up the risks and benefits of each treatment option. The identification of prognostic factors may help select patients who have high recurrence risk after ATD therapy and recommend them early thyroid ablation in order to avoid long, useless, and potentially harmful ATD therapy.

The aim of our study was to determine independent baseline predictors of medical treatment failure in patients with Graves' disease.

Material and Methods

Patients. This study was designed as a retrospective review of 194 adult patients aged ≥ 18 years with newly diagnosed Graves' disease, who were referred to the Clinic of Endocrinology, Hospital of Lithuanian University of Health Sciences, in Kaunas between 2002 and 2007. The diagnosis of GD was based on commonly accepted clinical and laboratory criteria: clinical symptoms of hyperthyroidism, diffuse goiter on ultrasound, increased concentrations of free thyroxine (FT_4) and suppressed thyrotropin (TSH) concentrations, and increased serum concentrations of TRAb. Patients who received a minimum of 12-month initial ATD treatment were included in the study. All data were retrieved from medical documents and recorded from the time of GD diagnosis to thyroid ablation or at least 12 months of follow-up after the discontinuation of

ATD treatment. The following parameters were recorded into the database at the onset of the disease for each patient: age, gender, goiter size by palpation, concentrations of FT_4 and TRAb, and thyroid echogenicity. The TRAb concentration at the end of ATD treatment was recorded, and an ultrasound examination of the thyroid after ATD treatment were performed. Information about a family history of thyroid disorders (including euthyroid, nodular goiter, and unspecified thyroid disorders) and duration of ATD treatment was also gathered from medical documents. Complete remission was characterized by the disappearance of signs and symptoms of thyrotoxicosis and the normalization of serum FT_4 and TSH concentrations. Remission of at least 12-month duration was regarded as long-term remission. Relapse was defined as the reappearance of the signs and symptoms of thyrotoxicosis, an increased serum concentration of FT_4 , and a suppressed concentration of TSH during the first year after the withdrawal of ATD therapy.

According to the outcome of the disease, the patients were divided into the remission and treatment failure groups. The remission group included patients who achieved long-term (at least 12 months) complete remission after the withdrawal of ATD therapy: the patients who achieved long-term complete remission after initial ATD treatment with no relapses were assigned to the group 1, and the patients who achieved long-term complete remission after 2 or 3 courses of ATD therapy, to the group 2. The patients from the group 2 relapsed (1 or 2 times) within 12 months after the withdrawal of initial ATD therapy and restarted ATD treatment until long-term (at least 12 months) complete remission was achieved. The treatment failure group included the patients who underwent thyroid abla-

Table 1. The Demographic and Clinical Characteristics of Patients

Characteristic	Remission Group (n=66)		Treatment Failure Group (n=128)		Total (n=194)
	Group 1 (n=44)	Group 2 (n=22)	Group 3 (n=44)	Group 4 (n=84)	
Male/female, n (%)	13 (19.7)/53 (80.3)		25 (19.5)/103 (80.5)		38 (19.6)/156 (80.4)
	11 (25.0)/33 (75.0)	2 (9.1)/20 (90.9)	7 (15.9)/37 (84.1)	18 (21.4)/66 (78.6)	
Age, mean (SD), years	42.9 (13.2)		42.2 (11.4)		42.4 (12.0)
	44.2 (12.4)	40.2 (14.6)	41.4 (13.5)	42.6 (10.2)	
Initial ATD therapy, mean (SD), months	18.0 (10.0)		23.6 (16.6)*		21.7 (14.9)
	19.2 (11.6)	15.6 (5.1)	19.1 (10.4)	26.3 (18.9)	
Total ATD therapy, mean (SD), months	23.4 (13.1)		30.3 (19.2)*		28.0 (17.6)
	19.2 (11.6)	31.9 (11.9)	38.1 (17.5)	26.3 (18.9)	

Group 1, patients who achieved long-term complete remission after initial ATD treatment with no relapses; group 2, patients who achieved long-term complete remission after 2 or 3 courses of ATD therapy; group 3, patients who underwent thyroid ablation due to relapses; and group 4, patients who underwent thyroid ablation without the withdrawal of ATD therapy.

ATD, antithyroid drugs.

* $P=0.02$ as compared to the remission group.

tion (surgery or radioiodine therapy): the patients who underwent thyroid ablation due to relapses were assigned to the group 3, and the patients who underwent thyroid ablation without the withdrawal of ATD therapy and had no steady episode of euthyroidism, to the group 4 (Table 1).

The goiter size was assessed by a clinical examination according to the criteria recommended by the World Health Organization (1960): grade 0, no goiter; grade 1A, palpable but not visible goiter; grade 1B, goiter palpable and visible only when the neck is fully extended; grade 2, goiter visible with the neck in a normal position; grade 3, very large goiter visible at a distance (12).

An ultrasound examination of the thyroid was performed with a high frequency real-time linear transducer (12 MHz). Echogenicity of the thyroid was described as normal (normoechogenic), decreased (hypoechoic), or increased (hyperechogenic) in relation to other structures. Typically, a normal thyroid gland shows higher echogenicity as compared with the prethyroid muscles and similar echogenicity as compared with the submandibular glands. The thyroid parenchyma is considered as hypoechoic when its echogenicity is lower than that of the submandibular glands or similar to that of the sternomastoid muscle (13). An ultrasound examination of the thyroid was performed in all patients at the time of diagnosis and after ATD therapy.

The serum TSH concentration (reference range, 0.17–4.05 mIU/L) was measured by an immunoradiometric assay (IRMA); the FT₄ concentration (reference range, 11.5–23.0 pmol/L) was determined by a radioimmunoassay (RIA) using commercially available kits (Immunotech, Beckman Coulter, Prague, Czech Republic) with a sensitivity of 0.025 mIU/L for TSH and 0.4 pmol/L for FT₄. TSH receptor antibodies (reference range, <9 U/L) were measured by a radioreceptor assay (RIAZEN TSH-R-Ab, ZenTech S.A., Belgium) with a sensitivity of 0.8 U/L. The intra-assay and interassay coefficients of variation were 3.7% and 5.7% for TSH, 6.7% and 6.5% for FT₄, 9.7% and 9.2% for TRAb, respectively.

Statistical Analysis. Descriptive analysis was used for demographic, clinical, and laboratory variables. One sample Kolmogorov-Smirnov test was used to test the normality of data distribution, and data are presented as mean and standard deviation. The chi-square (χ^2) test was used for the analysis of data when variables were categorical. For continuous data, the Mann-Whitney test was used. One-way analysis of variance (ANOVA) was used to compare means among more than 3 groups. Odds ratios (ORs) with 95% confidence interval (CI) were calculated by logistic regression to estimate associations between different variables and the outcome

of GD. For bivariate correlations, Spearman r correlation was employed. A receiver-operating characteristic (ROC) curve analysis was conducted to assess the prognostic value of TRAb in predicting the outcome of GD and determine an appropriate cutoff value. P values lower than 0.05 were considered as statistically significant. All statistical analyses were performed with the SPSS software (version 17.0).

Results

The data of 194 patients with Graves' disease were included in the study. The demographic and clinical characteristics of the patients are summarized in Table 1.

The peak incidence of Graves' disease was observed in the fifth decade of life for both the genders. More than three-quarters of patients were females, giving a male-to-female ratio of 1:4. The mean age of women was lower than that of men (41.4 [SD, 11.7] vs. 46.5 [12.5] years, $P=0.022$). At the onset of the disease, the thyroid eye disease was documented in 46% of the patients with the similar proportions of patients being in the remission and treatment failure groups. There were no significant differences in the mean age, sex, and mean concentrations of FT₄ among all the groups at the beginning of the treatment. Men and women had the same outcome after ATD treatment with a remission rate of 34%.

Age at the onset of the disease did not correlate with the TRAb concentrations before treatment ($P=0.248$), but showed a weak negative correlation with the initial FT₄ concentrations ($r=-0.249$, $P=0.046$), although this relationship was lost after adjustment for gender ($P=0.101$).

For the analysis, the patients were divided into 2 age groups: younger than 40 years at presentation and 40 years and older. Age at the onset of the disease did not predict the outcome of ATD treatment. Younger patients had the same remission rate as older ones ($P=0.79$).

There were 18.6% of the patients with GD (30 women and 6 men) who had a family history of thyroid disorders (including euthyroid, nodular goiter, and unspecified thyroid disorders). Similar proportions of men and women had a family history (19.2% and 15.8%, respectively, $P=0.625$). A family history of thyroid disorders was associated with a poor outcome of GD. The patients who failed ATD treatment had a family history of thyroid disorders more frequently than patients who achieved long-term remission (22.7% vs. 10.6%, $P=0.041$). The patients with a family history were 2.5 times more likely to fail to respond to medical treatment (OR 2.5; 95% CI, 1.02–5.99; $P=0.046$).

The mean concentration of TRAb before and at the end of ATD therapy was similar between

Table 2. Goiter Size and Laboratory Characteristics of Patients

Characteristics	Group 1 (n=44)	Group 2 (n=22)	Group 3 (n=44)	Group 4 (n=84)	Remission Group (1+2) (n=66)	Treatment Failure Group (3+4) (n=128)
Pretreatment goiter size, n (%)						
Grade 1 (n=31)	15 (34)	4 (18)	5 (11)	7 (8)	19 (29)	12 (9)
Grade 2 (n=126)	26 (59)	17 (77)	30 (68)	53 (63)	43 (65)	83 (65)
Grade 3 (n=37)	3 (7)	1 (5)	9 (21)	24 (29)	4 (6)	33 (26)
Pretreatment FT ₄ , mean (SD), pmol/L	58.6 (22.9)	55.0 (21.3)	59.2 (20.4)	60.3 (25.5)	57.3 (22.2)	59.9 (23.8)
Pretreatment TRAb, mean (SD), U/L	32.2 (30.8)	29.2 (22.2)	60.9 (75.4)	89.5 (88.1)	31.2 (28.1)	79.7 (84.8)*
TRAb at the end of ATD therapy, mean (SD), U/L	5.0 (3.9)	5.9 (4.4)	24.8 (22.8)	66.6 (78.0)	5.4 (4.0)	50.8 (66.0)*

Group 1, patients who achieved long-term complete remission after initial ATD treatment with no relapses; group 2, patients who achieved long-term complete remission after 2 or 3 courses of ATD therapy; group 3, patients who underwent thyroid ablation due to relapses; and group 4, patients who underwent thyroid ablation without the withdrawal of ATD therapy.

FT₄, free thyroxine; ATD, antithyroid drugs; TRAb, thyrotropin receptor autoantibodies.

* $P < 0.001$ as compared to the remission group.

the groups 1 and 2 ($P > 0.05$) and was significantly higher in the groups 3 and 4 than in the groups 1 and 2 and also were higher in the group 4 than in the group 3 ($P < 0.05$). The mean concentrations of initial TRAb and TRAb at the end of ATD therapy were significantly higher in the treatment failure group than in the remission group (79.7 [SD, 84.8] vs. 31.2 [SD, 28.1] U/L, $P < 0.001$; and 50.8 [SD, 66.0] vs. 5.4 [SD, 4.0] U/L, $P < 0.001$) (Table 2).

The area under the ROC curve (AUC) of TRAb levels before treatment was 0.76 (95% CI, 0.7–0.82; $P < 0.001$), and the TRAb concentration of 30.2 U/L was found to be the best cutoff value. The initial TRAb concentrations above 30.2 U/L identified patients who had never achieved long-term remission with a sensitivity of 74% and a specificity of 68% (positive predictive value, 82%; negative predictive value, 58%). The AUC of TRAb levels at the end of ATD therapy was 0.87 (95% CI, 0.78–0.93; $P < 0.001$). The difference in the AUC values between initial TRAb levels and TRAb levels at the end of ATD therapy was significant ($P = 0.011$). The TRAb concentrations at the end of ATD therapy above 12.97 U/L (as the best cutoff value) identified patients who had never achieved long-term remission with a sensitivity of 74% and a specificity of 97% (positive predictive value, 98%; negative predictive value, 71%). The calculated predicted probability of treatment failure with logistic regression analysis, based on a combination of initial TRAb levels and TRAb levels at the end of ATD therapy, had an AUC of 0.88 (95% CI, 0.79–0.94; $P < 0.001$), which was significantly higher than that of initial TRAb levels alone ($P = 0.0017$), but was not significantly different from that of TRAb levels at the end of ATD therapy ($P = 0.679$) (Fig. 1).

The goiter size was not associated with the serum FT₄ concentration at the onset of Graves' disease

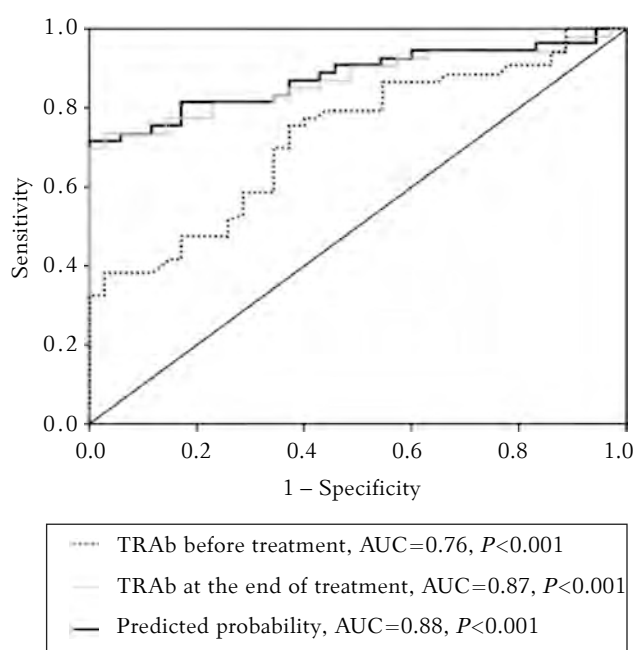


Fig. 1. Receiver operating characteristic curves comparing serum TRAb levels (U/L) at different time and the combined predicted probability.

AUC, area under the curve.

($P = 0.325$). A significant relationship was observed between the serum TRAb concentration at the time of diagnosis and the goiter size: the patients who had grade 3 goiter had higher mean concentrations of initial TRAb comparing with those who had grade 1 goiter (81.2 [SD, 89.6] vs. 53.0 [SD, 83.2] U/L, respectively; $P = 0.005$). The pretreatment goiter size was associated with the failure of ATD treatment. The presence of grade 2 goiter (OR, 3.1; 95% CI, 1.4–6.9; $P = 0.007$) and especially grade 3 goiter (OR, 13.1; 95% CI, 3.7–46.3; $P < 0.001$) as compared with grade 1 goiter increased the likelihood of failure.

From the beginning of ATD treatment, all the patients received methimazole at a daily dose of 30 mg. Some patients were treated with thiamazole alone (57.7%) by a dose titration regimen; others (42.3%) were supplemented by levothyroxine, continuing low-dose antithyroid drugs after initial high-dose ATD therapy (low-dose block-replace treatment). The remission group and the treatment failure group did not differ significantly according to the different drug regimens ($P=0.52$).

The duration of the first course of ATDs was 21.7 (SD, 14.9) months, and the total duration of ATD therapy was 28.0 (SD, 17.6) months. No significant difference in the initial duration of ATD treatment comparing the groups 1, 2, and 3 was observed ($P=0.408$). The duration of total ATD therapy was similar comparing the group 2 and the group 3 ($P=0.27$); no significant differences were found comparing other groups ($P<0.05$). The duration of the first course of ATD therapy and the total duration of ATD therapy were significantly shorter in the remission than the treatment failure group ($P=0.02$) (Table 1). The patients who underwent thyroid ablation due to relapse (group 3) had early relapse (within 6 months) after ATD withdrawal more frequently than those who achieved a long-term remission (group 2) ($P=0.041$).

At baseline, 172 patients (88.7%) had a hypoechogenic thyroid and 22 patients (11.3%) had a normoechogenic thyroid. After ATD therapy, a hypoechogenic thyroid was observed in 146 patients (75.3%), and a normoechogenic thyroid in 48 patients (24.7%). The remission and treatment failure

groups had the same frequency of hypoechogenic thyroid at the onset of the disease ($P=0.093$). The patients with a hypoechogenic thyroid after ATD therapy failed the ATD therapy more frequently than those with a normoechogenic thyroid (77.4% and 31.3%, respectively; $P<0.001$) (Fig. 2). The presence of a hypoechogenic thyroid after ATD therapy increased the risk of failure by more than 7.5 times (OR, 7.53; 95% CI, 3.7–15.5; $P<0.001$) with a sensitivity of 88% and a specificity of 50% (positive predictive value, 77%; negative predictive value, 73%).

Multiple logistic regression analysis was performed to select the independent baseline prognostic parameters. The variables that were chosen to enter them in the analysis were those with significant univariate associations with ATD treatment failure in this study, as well as those found in other studies to be significant predictors of failure (sex, age at diagnosis [less than 40 years], family history of thyroid disorders, goiter size at the onset of the disease, initial TRAb levels, thyroid echogenicity before treatment). The analysis revealed that higher TRAb levels and a larger goiter size (grade 2/grade 3) at the onset of Graves' disease were independent prognostic factors predicting the failure of medical treatment (Table 3).

Table 3. Multiple Logistic Regression Analysis of Baseline Prognostic Factors Associated With Failure of Antithyroid Drug Treatment

Predictor	OR (95% CI)	P
Age <40 years	0.6 (0.3–1.3)	0.182
Family history of thyroid diseases	2.5 (0.9–6.7)	0.067
Sex (male)	1.3 (0.6–3.2)	0.535
Goiter size		
Grade 2	3.2 (1.3–8.2)	0.015
Grade 3	13.9 (3.3–57.6)	<0.001
TRAb levels	1.03 (1.01–1.04)	<0.001
Hypoechogenic thyroid	1.5 (0.6–4.2)	0.419

OR, odds ratio; CI, confidence interval;

TRAb, thyrotropin receptor autoantibodies.

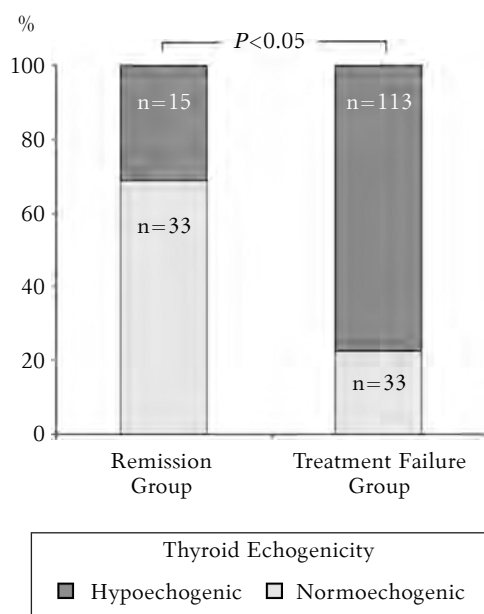


Fig. 2. Distribution of the patients by thyroid echogenicity after ATD treatment in the remission and treatment failure groups

Discussion

Conservative therapy with ATD is the first choice treatment of Graves' disease in Europe. Antithyroid drugs are effective for an acute control of the disease, but approximately 30% to 60% of patients relapse within 1 year of ATD discontinuation (6, 14). In this study, the remission rate of 34% after ATD treatment was documented. This finding is in accordance with those of other European studies, where remission rates after ATD treatment ranged between 30% and 60% (3, 15, 16). As the response to the treatment of GD is unpredictable, it is often problematic to choose the optimal treatment approach. The reliable predictors of relapse

after ATD treatment would greatly improve patient management by facilitating the identification of patients who require long-term ATD treatment, early surgery, or radioiodine therapy.

The annual incidence of GD is around 21 cases per 100 000 population in Sweden (17) and around 38 cases per 100 000 in the United States (18). GD occurs more frequently (4 to 6 times) in women (9), and this may be related to the influence of estrogens on the immune system, particularly the B cell repertoire (19). In our study, more than three-quarters of the patients were women, giving a male-to-female ratio of 1:4. Some authors have found that men with GD are less likely to enter remission after ATD therapy (3, 20), but we, in agreement with other authors (21), did not find any effect of sex on long-term remission; men and women had the same outcome after ATD therapy. However, a small sample size should be taken into account.

Age-related differences in the clinical presentation of Graves' hyperthyroidism have also been reported (22), with the severity of hyperthyroidism and prevalence of antibodies shown to decrease with advancing age (23). Allahabadia et al. (3) reported that younger patients (younger than 40 years) were more likely to fail to respond to ATD treatment. In our study, younger patients aged less than 40 years and older patients had the same frequency of treatment failure. These findings agree with those of our previous smaller study (24).

Several other factors have been postulated to indicate a poor prognosis in terms of remission rates after ATD treatment, including a large goiter size (20) and biochemical severity of thyroid dysfunction (25). The effect of goiter size on the remission rate has been reported in many studies (16, 20), but some researchers have failed to confirm this association (21, 26). Differences in goiter size assessment may account for such discrepancies. Despite the limitations of the assessment of goiter size by a clinical examination, rather than ultrasound, which is more precise in diagnosing goiter, our results showed a significant association between a larger goiter size and failure of medical treatment. The present finding showing that larger (grade 2/grade 3) goiter independently predicts the failure of ATD treatment is plausible, suggesting that the goiter size may be a significant marker of the severity of the autoimmune process. The severity of biochemical hyperthyroidism has been suggested to indicate a poor prognosis after medical treatment (21, 26). Our data could not confirm this hypothesis: the FT_4 level at the onset of the disease did not differ between the remission and treatment failure groups. An association between grade 3 goiter and higher TRAb concentrations may be explained by a stronger stimulatory effect of higher TRAb levels on

the growth of the thyroid gland.

Early evidence that GD has a hereditary component comes from familial studies. The strongest evidence for genetic susceptibility to GD comes from twin studies (27). In our study, a family history of thyroid disorders was observed in 18.6% of the patients, contrary to reports by other authors, where the frequency of around 50% has been observed (28). A low frequency of family history in our patients was most probably because of a retrospective study design. Family data have some limitations because the design of this study was retrospective, and patients reported a family history of all types of thyroid diseases, including euthyroid, nodular goiter and unspecified thyroid disorders without medical confirmation. Although a family history of thyroid disorders failed to reach statistical significance as a predictor in the multiple logistic regression analysis, it showed a significant univariate association with a poor outcome of ATD treatment. These findings provide some insight into the complex interactions among genetic factors, which account for a major part of disease susceptibility.

Conflicting data exists regarding the effect of the duration of ATD therapy on the relapse of GD. Abraham et al. (14) summarized trials examining the effect of duration and mode of ATD therapy on relapse rates and reported that using the titration regimen, 12 months was superior to 6 months, but there was no benefit in extending treatment beyond 18 months. The addition of thyroxine with continued low-dose antithyroid therapy after initial ATD therapy does not appear to provide any benefit in terms of recurrence of hyperthyroidism (14). In our study, all the patients were treated with ATD for a minimum period of 12 months. The initial duration of ATD therapy was significantly shorter in the remission group than the treatment failure group, suggesting that a longer duration of initial treatment did not have influence on treatment outcome according to our previous findings (24). The total duration of ATD treatment was longer in the patients who underwent thyroid ablation comparing with the patients in the remission group; obviously, this result is due to the fact that subjects with more relapses had to restart the therapy more times. Similar results were reported by Cappelli et al., who investigated the impact of the duration of ATD therapy on the GD outcome (21). In our study, a combined administration of thyroxine and an antithyroid drug after initial high-dose antithyroid therapy did not have a positive influence on treatment outcome. These findings are in agreement with the results of the study by Abraham et al. and our previous findings (14, 24).

Ultrasonography is the most widely used imaging method for the diagnosis and follow-up of thyroid

disorders. Graves' disease leads to thyroid enlargement and reduction of tissue echogenicity (29). Zingrillo et al. (30) reported that thyroid hypoechogenicity before treatment could not be used as a predictor of relapse of GD. They pointed out that antithyroid drugs may alter the follicular structure and influence changes in thyroid hypoechogenicity. The absence of thyroid hypoechogenicity after ATD treatment is a favorable prognostic sign of remission, because it directly reflects an inflammatory status of the thyroid gland. Schiemann et al. (29) analyzed thyroid echogenicity on standardized gray-scale ultrasonography and reported that ultrasonic hypoechogenicity was closely correlated with the levels of TSH and TRAb. Patients in clinical and laboratory remission show significantly higher thyroid echogenicity than patients who need active antithyroid treatment (29). In our study, thyroid echogenicity before treatment did not predict the outcome of medical treatment, and this is in agreement with the study by Zingrillo et al. (30). However, a hypoechogenic thyroid after ATD therapy was associated with the higher probability of medical treatment failure, and this is in line with other studies (29, 30).

Autoimmunity against the thyroid-stimulating hormone receptor is the main pathogenetic element of GD, and TRAb detection is used for diagnosis of GD. The presence of TRAb in the serum of patients treated with ATD may reflect the activity of the ongoing disease. The role of TRAb as a predictor of outcome after ATD treatment is still controversial. There are many factors such as a design of studies (prospective vs. retrospective), laboratory tests for the evaluation of TRAb (thyroid-stimulating vs. thyroid-blocking antibodies), and time of TRAb measurement (before vs. after treatment) that complicate the overall interpretation of this issue. In agreement with many other studies, our study revealed that TRAb levels evaluated at the onset of the disease and at the end of ATD therapy were associated with a disease outcome (5, 16, 20, 21). Higher levels of TRAb at the onset of the disease were independently associated with the failure of ATD treatment; although the ROC curve analysis revealed that TRAb levels at the end of ATD therapy provided a higher accuracy in predicting medical treatment failure. Because of

the retrospective design of this study, some laboratory parameters were measured in less than 50% of the patients. Although some of these parameters, including TRAb levels at the end of ATD therapy, were significant predictors of medical treatment failure in the univariate analysis, they could not be included in the multivariate analysis without excessively restricting the sample size.

An additional limitation of the study is that patients were followed up for only 1 year after the discontinuation of antithyroid medications. It is possible that some patients experienced the relapse of the disease after 1 year; however, the data from previous studies indicated that the frequency of such late relapses was lower. We also were unable to follow ophthalmopathy progression in patients with GD because of the absence of standardized clinical assessments of eye disease severity in medical documents (e.g., NOSPECS classification). Some patients with ophthalmopathy required therapy with glucocorticoids; other patients due to a good control of thyroid function with ATD did not need further treatment (therapy with glucocorticoids or retrobulbar irradiation). Because of a retrospective design of this study, patients with ophthalmopathy were treated with glucocorticoids at various doses and for different treatment periods. We could not follow the changes in TRAb concentrations and determine the influence of therapy with glucocorticoids on the outcome of the disease.

Conclusions

This study demonstrates that thyroid-stimulating hormone receptor antibodies and the goiter size at the onset of the disease are independent baseline predictors of medical treatment failure in patients with Graves' disease. The patients with higher levels of thyroid-stimulating hormone receptor antibodies and a larger goiter size at the onset of the disease were unlikely to achieve long-term remission. For the patients who are less likely to achieve long-term remission, the consideration of alternative treatment options (surgery or radioiodine therapy) may be advisable.

Statement of Conflict of Interest

The authors state no conflict of interest.

References

- Huber A, Menconi F, Corathers S, Jacobson EM, Tomer Y. Joint genetic susceptibility to type 1 diabetes and autoimmune thyroiditis: from epidemiology to mechanisms. *Endocr Rev* 2008;29:697-725.
- Weetman AP. Graves' disease. *N Engl J Med* 2000;343:1236-48.
- Allahabadi A, Daykin J, Holder RL, Sheppard MC, Gough SC, Franklyn JA. Age and gender predict the outcome of treatment for Graves' hyperthyroidism. *J Clin Endocrinol Metab* 2000;85:1038-42.
- Muehlberg T, Gilbert JA, Rao PV, McGregor AM, Banga JP. Dynamics of thyroid-stimulating and -blocking antibodies to the thyrotropin receptor in a murine model of Graves' disease. *Endocrinology* 2004;145:1539-45.
- Schott M, Scherbaum WA, Morgenthaler NG. Thyrotropin receptor autoantibodies in Graves' disease. *Trends Endocrinol Metab* 2005;16:243-8.
- Mao XM, Li HQ, Li Q, Li DM, Xie XJ, Yin GP, et al. Prevention of relapse of Graves' disease by treatment with an intrathyroid injection of dexamethasone. *J Clin Endocrinol Metab* 2009;94:4984-91.

7. Weetman AP. Controversy in thyroid disease. *J R Coll Physicians Lond* 2000;34:374-80.
8. Schussler-Fiorenza CM, Bruns CM, Chen H. The surgical management of Graves' disease. *J Surg Res* 2006;15:133:207-14.
9. Stalberg P, Svensson A, Hessman O, Akerstrom G, Hellman P. Surgical treatment of Graves' disease: evidence-based approach. *World J Surg* 2008;32:1269-77.
10. Alexander EK, Larsen PR. High dose of (131)I therapy for the treatment of hyperthyroidism caused by Graves' disease. *J Clin Endocrinol Metab* 2002;87:1073-7.
11. Osmolski A, Frenkiel Z, Osmolski R. [Complications in surgical treatment of thyroid diseases]. *Otolaryngol Pol* 2006;60:165-70.
12. Delange F, Bastani S, BenMiloud M. Definitions of endemic goiter and cretinism, classification of goiter size and severity of endemias, and survey techniques. In: Dunn JT, Pretell EA, Daza CH, Viteri FE, editors. *Towards the eradication of endemic goiter, cretinism, and iodine deficiency*. Washington, DC: Pan American Health Organization; 1986. p. 373-6.
13. Ying M, Brook F, Ahuja A, Metreweli C. The value of thyroid parenchymal echogenicity as an indicator of pathology using the sternomastoid muscle for comparison. *Ultrasound Med Biol* 1998;24:1097-105.
14. Abraham P, Avenell A, McGeoch SC, Clark LF, Bevan JS. Antithyroid drug regimen for treating Graves' hyperthyroidism. *Cochrane Database Syst Rev* 2010;1:CD003420.
15. Anagnostis P, Adamidou F, Polyzos SA, Katargari S, Karathanasi E, Zouli C, et al. Predictors of long-term remission in patients with Graves' disease: a single center experience. *Endocrine* 2013;44:448-53.
16. Nedrebo BG, Holm PI, Uhlving S, Sorheim JI, Skeie S, Eide GE, et al. Predictors of outcome and comparison of different drug regimens for the prevention of relapse in patients with Graves' disease. *Eur J Endocrinol* 2002;147:583-9.
17. Abraham-Nordling M, Bystrom K, Torring O, Lantz M, Berg G, Calissendorff J, et al. Incidence of hyperthyroidism in Sweden. *Eur J Endocrinol* 2011;165:899-905.
18. McLeod DS, Cooper DS. The incidence and prevalence of thyroid autoimmunity. *Endocrine* 2012;42:252-65.
19. Da Silva JA. Sex hormones, glucocorticoids and autoimmunity: facts and hypotheses. *Ann Rheum Dis* 1995;54:6-16.
20. Vitti P, Rago T, Chiovato L, Pallini S, Santini F, Fiore E, et al. Clinical features of patients with Graves' disease undergoing remission after antithyroid drug treatment. *Thyroid* 1997;7:369-75.
21. Cappelli C, Gandossi E, Castellano M, Pizzocaro C, Agosti B, Delbarba A, et al. Prognostic value of thyrotropin receptor antibodies (TRAb) in Graves' disease: a 120 months prospective study. *Endocr J* 2007;54:713-20.
22. Nordyke RA, Gilbert FI Jr, Harada AS. Graves' disease. Influence of age on clinical findings. *Arch Intern Med* 1988;148:626-31.
23. Aizawa T, Ishihara M, Hashizume K, Takasu N, Yamada T. Age-related changes of thyroid function and immunologic abnormalities in patients with hyperthyroidism due to Graves' disease. *J Am Geriatr Soc* 1989;37:944-8.
24. Daukšienė D, Daukša A, Mickuvienė N. Greivso ligos medikamentinio gydymo nesėkmės prognozės žymenys. (Prognostic factors in predicting medical treatment failure of Graves disease.) *Lietuvos endokrinologija* 2011;19:69-74.
25. Weetman AP, Ratanachaiyavong S, Middleton GW, Love W, John R, Owen GM, et al. Prediction of outcome in Graves' disease after carbimazole treatment. *Q J Med* 1986;59:409-19.
26. Cinemre H, Bilir C, Gokosmanoglu F, Akdemir N, Erdogmus B, Buyukkaya R. Predictors of time to remission and treatment failure in patients with Graves' disease treated with propylthiouracil. *Clin Invest Med* 2009;32:E199-205.
27. Cappa M, Bizzarri C, Crea F. Autoimmune thyroid diseases in children. *J Thyroid Res* 2010;2011:675703.
28. Manji N, Carr-Smith JD, Boelaert K, Allahabadia A, Armitage M, Chatterjee VK, et al. Influences of age, gender, smoking, and family history on autoimmune thyroid disease phenotype. *J Clin Endocrinol Metab* 2006;91:4873-80.
29. Schiemann U, Gellner R, Riemann B, Schierbaum G, Menzel J, Domschke W, et al. Standardized grey scale ultrasonography in Graves' disease: correlation to autoimmune activity. *Eur J Endocrinol* 1999;141:332-6.
30. Zingrillo M, D'Aloiso L, Ghiggi MR, Di CA, Chiodini I, Torlontano M, et al. Thyroid hypoechogenicity after methimazole withdrawal in Graves' disease: a useful index for predicting recurrence? *Clin Endocrinol (Oxf)* 1996;45:201-6.

Received 18 September 2012, accepted 30 October 2013