Archives of Endocrinology and Metabolism
ISSN 2359-4292
OFFICIAL JOURNAL OF THE BRAZILIAN SOCIETY OF ENDOCRINOLOGY AND METABOLISM

original

Predictive factors for remission in pediatric Graves' disease treated with antithyroid drugs: a retrospective study

Ana João Fernandes[1,*]
https://orcid.org/0009-0004-8175-8426

João Pedro Gomes[2,*]
https://orcid.org/0009-0007-1086-4506

Rita Aldeia da Silva[1]
https://orcid.org/0009-0005-6468-125X

Mariana Portela[1]
https://orcid.org/0000-0002-3297-9431

Sofia Martins[1]
https://orcid.org/0000-0003-3197-1320

Olinda Marques[3]
https://orcid.org/0000-0002-9612-0589

Ana Antunes[1]
https://orcid.org/0000-0002-6277-6037

Maria Miguel Gomes[1,2]
https://orcid.org/0000-0002-1728-9387

[1] Departamento de Pediatria, Unidade Local de Saúde de Braga, Braga, Portugal
[2] Escola de Medicina, Universidade do Minho, Braga, Portugal
[3] Departamento de Endocrinologia, Unidade Local de Saúde de Braga, Braga, Portugal

*Ana João Fernandes and João Pedro Gomes contributed equally to this work and share first authorship.

ABSTRACT

Objective: Pediatric Graves' disease (PGD) is a rare autoimmune disorder. Predictors of remission after antithyroid drug (ATD) therapy have been described, but their applicability in pediatric populations remains inconsistent. This study aimed to identify clinical and biochemical factors associated with remission in PGD. **Subjects and methods:** Retrospective observational study of children diagnosed with Graves' disease at a tertiary hospital in Portugal (2001–2021), all initially treated with ATD. Demographic, clinical, and biochemical data at diagnosis, as well as treatment characteristics, were analyzed. The primary outcome was "In Remission". Secondary outcomes included "Remission Experienced" and "Disease Not Active ≥1 Month". Comparative analyses, logistic regression, and receiver operating characteristic (ROC) curves were applied. **Results:** Among 45,000 pediatric patients, 36 had PGD (prevalence 0.08%). Median age at diagnosis was 13.7 years; 81% were female, and 78% were pubertal. Methimazole was prescribed in 81%. Mean initial treatment duration (ITD) was 35 months, and 47% of patients underwent definitive therapy. At data collection, 18% of patients were in remission. Longer ITD (>24 months) was associated with remission; the optimal cutoff ITD was 51 months. Smaller thyroid volume and TRAb levels at diagnosis correlated with "Disease Not Active ≥1 Month". Thyroid volume ≥2.5-fold the upper limit of normal (ULN) and TRAb ≥7.055-fold the ULN reduced the likelihood of disease inactivity. **Conclusion:** Prolonged ATD therapy, smaller thyroid volume, and lower TRAb levels at diagnosis were associated with favorable outcomes in PGD. The optimal ITD (≥51 months) supports extended ATD therapy. Larger multicenter studies are warranted to confirm these results.

Descriptors: Hyperthyroidism; Graves disease; pediatric endocrinology; antithyroid agents; remission

INTRODUCTION

Pediatric hyperthyroidism is rare, with an estimated incidence of 0.9 per 100,000 in children under the age of 15 years and a clear female predominance. Graves' disease (GD) accounts for over 95% of cases (1–4). Its pathogenesis involves genetic, environmental, and immune factors mediated by thyrotropin receptor antibodies (TRAb) that stimulate thyroid hormone overproduction and may coexist with other autoimmune conditions, such as type 1 diabetes or celiac disease (2,3,5). A positive family history in a first-degree relative is observed in ~15% of cases, and the risk of the disease is higher in individuals with Down syndrome or under immunomodulatory therapy (2,3).

Clinically, pediatric GD (PGD) presents with typical symptoms of hyperthyroidism, *i.e.*, tachycardia, palpitations, hyperphagia, tremors, diarrhea, heat intolerance, anxiety, and insomnia. Over 50% of patients have diffuse goiter, and mild ophthalmopathy may occur. Growth velocity and bone maturation can also be accelerated (1–3). The diagnosis relies on suppressed thyroid-stimulating hormone (TSH) levels, elevated

Received on Sep/3/2025
Accepted on Mar/24/2026
DOI: 10.20945/2359-4292-2026-0089

Correspondence to:
Ana João Fernandes
ana.joao.fernandes@ulsb.min-saude.pt

Associated editor:
Sonir R. Antonini
https://orcid.org/0000-0003-4778-8803

This is an open-access article distributed under the terms of the Creative Commons Attribution License

Copyright© AE&M all rights reserved.

free triiodothyronine (FT3) and free thyroxine (FT4) levels, and positive TRAb, which are highly specific for GD (1,2). Thyroid ultrasound, though not required for diagnosis, can aid prognostic assessment, while thyroid scintigraphy is not routinely indicated (1,2). Antithyroid drugs (ATDs) remain first-line therapy, with radioiodine (^{131}I) and thyroidectomy as second-line options (1,3,6,7). In children, however, the optimal therapeutic approach remains controversial (1,4–10).

Thionamides, such as methimazole, carbimazole, and propylthiouracil (PTU), inhibit thyroid hormone synthesis, but PTU is avoided in children due to hepatotoxicity (2,3,10). After 2 years of ATD therapy, remission rates are low (20%–30%), though they increase with longer treatment duration, reaching rates of 19.6%, 34.1%, 43.5%, and 50.6% after 3, 5, 7, and 10 years, respectively (1,3,4,11,12).

Since 2016, the American Thyroid Association (ATA) has acknowledged the potential benefit of extending ATD therapy for GD beyond 2 years (7). In contrast, the 2022 European Thyroid Association (ETA) guideline recommends an initial treatment duration (ITD) of at least 3 years (3). However, prolonged ATD therapy poses challenges, including the occurrence of adverse effects, requirement for frequent thyroid function monitoring, and poor adherence, especially among adolescents (4,5,10,13).

Adverse effects of ATDs, although generally uncommon, must be considered when evaluating prolonged therapy. Minor reactions such as rash, pruritus, or mild gastrointestinal intolerance occur in approximately 5%–25% of pediatric patients, while severe complications like agranulocytosis or hepatotoxicity are rare, with estimated incidences of 0.3%–0.7%. Definitive treatments remain controversial in children due to the potential oncogenic risk of ^{131}I, and higher surgical complication rates than in adults (1–4). These considerations highlight the importance of balancing the risks and benefits of prolonged medical therapy versus early definitive treatment, and of identifying reliable clinical and biochemical predictors of response to ATD that can guide individualized management in PGD (10,14).

While international studies have explored age, sex, TRAb levels, thyroid volume, and treatment duration as prognostic factors in PGD, results have been inconsistent (1,4,6,8–16). In Portugal, research is scarce, with only one multicenter study published to date (6). Further investigation is needed to clarify prognostic markers and align national practice with international evidence.

In this context, the present study aimed to evaluate clinical and biochemical factors associated with remission in PGD following ATD therapy in patients diagnosed and followed at a Portuguese tertiary hospital between 2001 and 2021.

SUBJECTS AND METHODS

The study protocol was approved by the Ethics Committee and the Data Protection Office of Braga Hospital. Ethical guidelines and good clinical practice standards were followed, and data confidentiality was maintained throughout the study period.

This retrospective, observational, analytical, longitudinal study included children and adolescents diagnosed with GD at Braga Hospital, a Portuguese tertiary care center, between January 2001 and December 2021. All patients received initial ATD treatment and were followed at the Pediatrics Department and, when applicable, the Pediatric Endocrinology Unit.

A non-random convenience sample was selected. The inclusion criteria were a first consultation within the study period, a GD diagnosis with ATD as the initial therapy, and a follow-up ≥ 6 months. The exclusion criteria included neonatal GD, age > 18 years at diagnosis, and loss to follow-up.

A GD diagnosis required suppressed serum TSH levels, elevated FT3 and/or FT4 levels, and positive TRAb. In some cases, FT3 measurement was unavailable at diagnosis; in these instances, the diagnosis was based on TSH and FT4 levels, along with TRAb results. Consequently, FT3 was excluded from statistical analyses due to missing data.

Treatment followed a dose-titration regimen aiming for euthyroidism (thyroid hormones within reference ranges, followed by TSH). ATD withdrawal was considered after ≥ 24 months of stable euthyroidism on minimal doses (as low as 2.5 mg of methimazole on alternate days) and TRAb levels near or below the upper limit of normal (ULN). Remission was defined as

Copyright© AE&M all rights reserved.

≥ 12 months of euthyroidism after ATD discontinuation, and relapse as recurrent biochemical hyperthyroidism with positive TRAb after remission.

Throughout the 20-year study period, patients were managed exclusively using a dose-titration regimen. A block-and-replace approach was not employed at any time. Despite changes in ATD preference over time — namely, a shift toward reduced PTU use and increased methimazole use — and modifications to laboratory assay methodologies, clinical decision-making regarding treatment adjustment and ATD withdrawal followed consistent criteria.

Body mass index (BMI) — calculated as weight divided by the square of height (kg/m²) and expressed as a standard deviation (SD) score (SDS) by age and sex — was classified as underweight, normal weight, overweight, or obese (17,18). According to the World Health Organization (WHO) growth reference, for children under 5 years of age, underweight was defined as a BMI-for-age < −2 SD, normal weight as −2 ≤ SD ≤ +2, overweight as +2 < SD ≤ +3, and obesity as SD > +3. For children aged 5 years or older, underweight was defined as BMI-for-age < −2 SD, normal weight as −2 ≤ SD ≤ +1, overweight as +1 < SD ≤ +2, and obesity as SD > +2. Pubertal status was assessed at diagnosis, based on Tanner staging documented at the first endocrinology evaluation, classifying patients as prepubertal or pubertal.

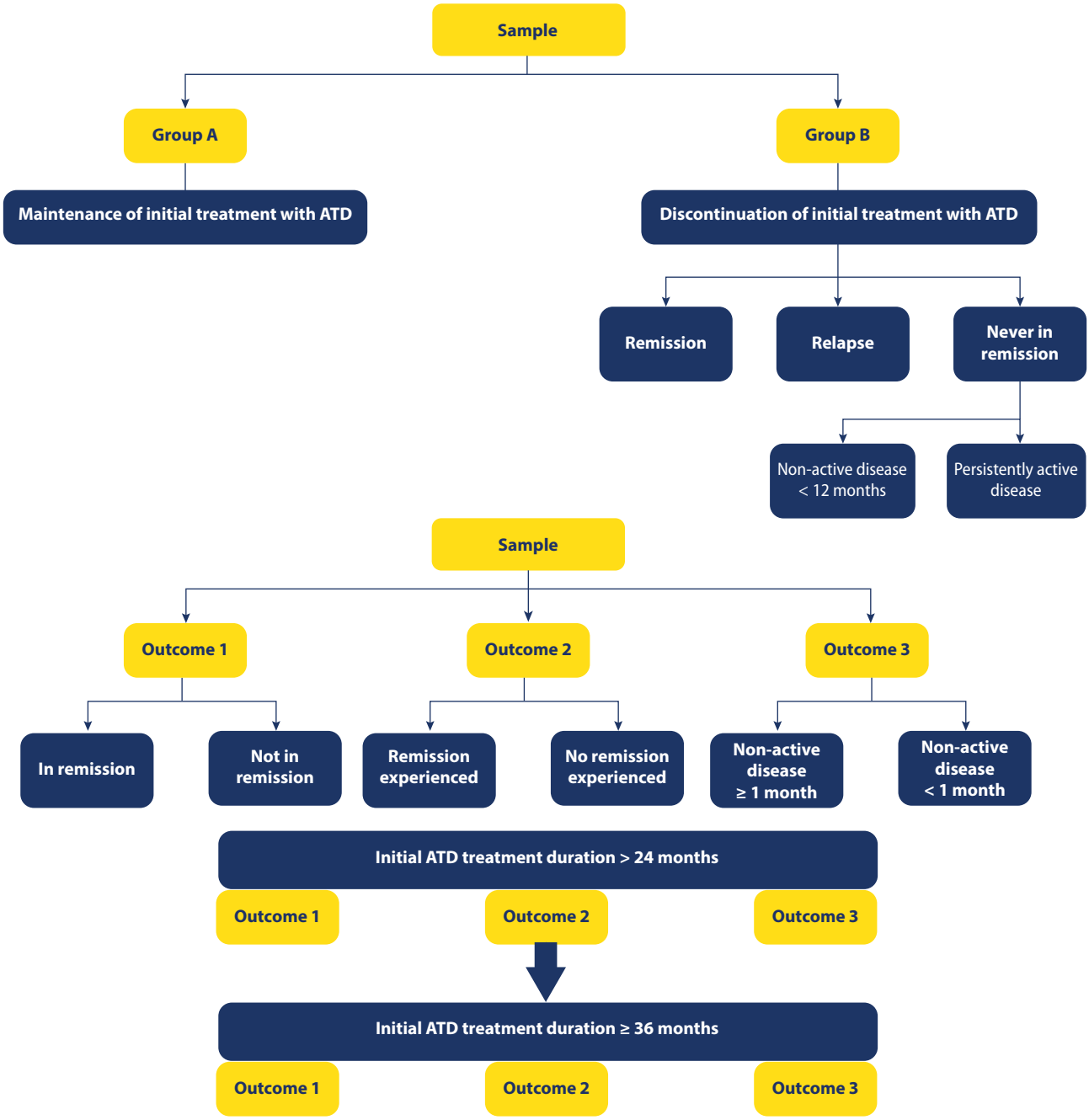
Puberty was defined as Tanner stage ≥ 2 for breast development in girls and Tanner stage ≥ 2 for genital development (testicular volume ≥ 4 mL) in boys (19). Thyroid volume was determined by ultrasound using the ellipsoid formula (width x length x depth x 0.52), standardized by the age-specific ULN, through the ratio of measured volume to ULN (7 years: 4.4; 8 years: 4.5; 9 years: 4.9; 10 years: 5.5; 11 years: 6.4; 12 years: 6.7; 13 years: 7.7; 14 years: 7.8). For children younger than 7 years and older than 14 years, ULN values of 4.4 and 7.8, respectively, were applied (20,21). Not all patients underwent thyroid ultrasound.

Serum FT4 and FT3 levels were measured using competitive chemiluminescent immunoassays, while TSH levels were measured by immunochemiluminescent assay (Atellica IM Analyzer, Siemens Medical Solutions Diagnostics, Erlangen, Germany) and TRAb by radioimmunoassay until 2017 and by chemiluminescent immunoassay from 2017 onwards (IMMULITE 2000 Systems Analyzers, Siemens Medical Solutions Diagnostics). Results were standardized by dividing each by its respective ULN. Reference values for FT4 were 0.9–2.3 ng/mL for ages 0–12 months, 0.8–1.8 ng/mL for ages 1–5 years, 1.0–2.1 ng/mL for ages 6–10 years, and 0.8–1.9 ng/mL for ages 11–18 years. Reference values for TRAb were < 1 U/L or < 9 U/L before 2017 and < 0.55 U/L thereafter (22).

Data were obtained from medical records. Collected variables included sex, the presence of other autoimmune diseases (AIDs), age at diagnosis, pubertal stage, BMI, thyroid volume, FT4 and TRAb levels, type and dose of initial ATD therapy, and outcome after the initial ATD treatment. The sample was subsequently divided into groups according to the classification schemes shown in **Figure 1**. Group A comprised patients who were still receiving their initial ATD regimen at the time of data collection, while Group B comprised patients who had already discontinued their initial ATD regimen before data collection. Within Group B, patients were further categorized into the following groups: "Remission" (euthyroidism maintained for ≥ 12 months after ATD discontinuation), "Relapse" (suppressed serum TSH levels and elevated FT3 and/or FT4 levels, along with positive TRAb, after achieving remission), and "Never in Remission" (disease inactive for < 12 months or persistently active). For analytical purposes, three outcomes were defined: Outcome 1 – "In Remission", which included patients in remission at the time of data collection, and "Not in Remission", which encompassed all remaining patients, including those who had never discontinued initial ATD therapy (Group A); Outcome 2 – "Remission Experienced", grouping patients who achieved remission at any time, and "No Remission Experienced", grouping those who never achieved remission or relapsed without attaining a subsequent remission; Outcome 3 – "Disease Not Active ≥ 1 Month", which included all patients from the "Remission Experienced" group plus those with inactive disease lasting ≥ 1 month but < 12 months after discontinuation of the initial ATD regimen, and those with "Disease Not Active < 1 month". The selected outcomes were defined to reflect clinically meaningful

Copyright© AE&M all rights reserved.

stages of PGD management. While sustained remission represents the primary therapeutic goal and was therefore defined as the primary outcome ("In Remission"), shorter periods of disease inactivity and remission experienced were included as secondary outcomes, as they are also clinically relevant and frequently inform decisions regarding ATD withdrawal, treatment continuation, and timing of definitive therapy.



ATD: antithyroid drugs.

Figure 1. Characterization of the sample according to maintenance or discontinuation of the initial treatment, by Outcomes 1, 2, and 3, and according to the duration of the initial treatment. **Group A:** initial antithyroid drug treatment ongoing at data collection. **Group B:** discontinuation of initial treatment with antithyroid drugs. "Remission": patients in remission at the time of data collection. "Relapse": patients who experienced relapse after remission (without achieving a second remission). "Never in Remission": patients with inactive disease for less than 12 months or with persistently active disease. **Outcome 1:** "In Remission" included patients who were in remission at the time of data collection; "Not in Remission" included all remaining patients, including those who had never discontinued initial antithyroid drug treatment **(Group A). Outcome 2:** "Remission Experienced" grouped patients who had either achieved remission or relapsed after remission (without reaching a new remission). **Outcome 3:** "Disease Not Active ≥ 1 Month" included the sum of patients from the "Remission Experienced" group and those who had inactive disease lasting at least 1 month and less than 12 months.

Copyright© AE&M all rights reserved.

Statistical analysis

Statistical analyses were performed using IBM SPSS Statistics, version 29.0 (IBM Corp., Armonk, NY, USA). Statistical significance was defined as $p < 0.05$, and 95% confidence intervals (CIs) were used to support the interpretation of the results.

The minimum sample size considered representative of the Portuguese pediatric population was $n \geq 16$, calculated using an online sample size calculator (23), assuming a 95% CI and a 5% margin of error.

Normality of quantitative variables was assessed using the Shapiro–Wilk test, histogram inspection, and evaluation of skewness and kurtosis (24). Depending on the normality assumption, parametric or nonparametric tests were applied.

Normally distributed quantitative variables were summarized as mean and SD, while non-normally distributed variables were reported as median and interquartile range. Qualitative variables were expressed as absolute frequencies (n) and/or relative frequencies (%) (24).

Comparisons between groups were performed using statistical tests selected according to group dependence (paired or independent), variable type (quantitative or qualitative), and distribution pattern of quantitative variables, with parametric or nonparametric methods applied as appropriate (25-27).

Binary logistic regression (LR) models — univariate and multivariate — were used to identify potential predictors of outcomes (28). Variable selection considered statistical significance in group comparisons, sample size, and the number of variables analyzed. To reduce overfitting, the "1:10 rule" (one predictor per ten positive events) (29) was applied, although not strictly, for exploratory purposes, given conflicting evidence in the literature (30,31).

To minimize multicollinearity and improve robustness, only variables with $p < 0.20$ in comparative analyses were included in the univariate LR. Those with $p < 0.20$ in univariate LR were subsequently entered into multivariate LR (10). Results were reported as odds ratios (ORs) with 95% CIs. For significant predictors ($p < 0.05$), Nagelkerke R^2 was calculated to estimate the variance explained by the model (28).

Receiver operating characteristic (ROC) curve analysis was performed for continuous variables of potential clinical interest to assess discriminative ability. For variables with an area under the curve (AUC) > 0.5 — indicating discrimination above chance and improving as AUC approached 1 — the optimal cutoff was determined using the Youden Index (J = Sensitivity + Specificity – 1). A p value < 0.05 indicated statistically significant discriminative capacity (26,28,32).

RESULTS

A total of 45,000 children and adolescents were followed at the Pediatrics Department of Braga Hospital from January 2001 to December 2021, of whom 36 were diagnosed with PGD, corresponding to a prevalence of 0.08%. Two patients treated with PTU were subsequently excluded due to loss to follow-up, resulting in a final sample of 34 patients included in the analysis (**Figure 2**). Clinical and biochemical characteristics at diagnosis, along with details of initial ATD treatment and outcomes, are summarized in **Table 1**. Not all patients had complete laboratory or thyroid ultrasound data available due to missing information in the clinical records.

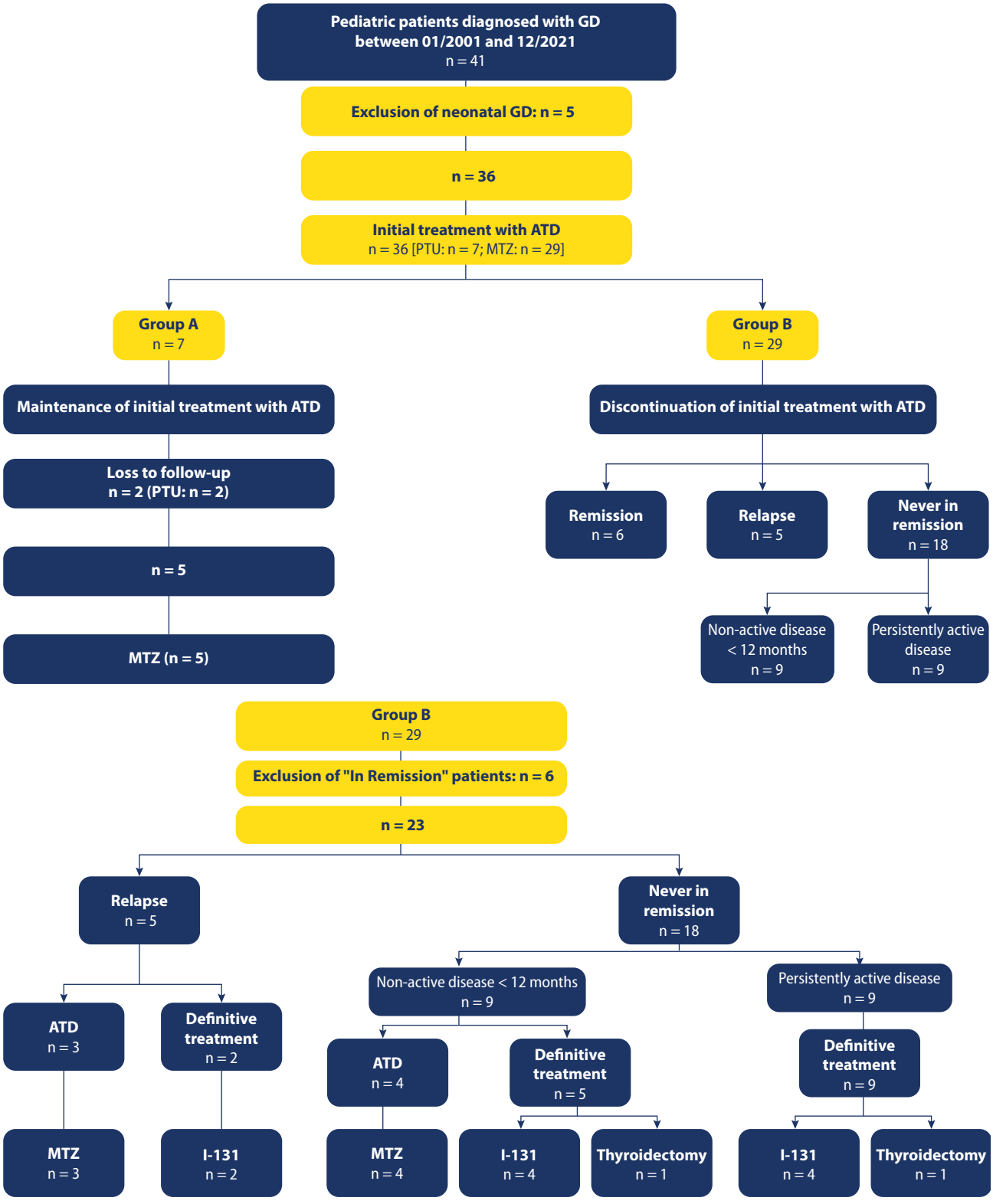
None of the patients was initially treated with definitive therapy. One patient, who developed PTU-induced toxic hepatitis, discontinued ATD after 4 months and underwent subsequent ^{131}I ablation.

In comparative analyses, longer ITD was significantly associated with remission outcomes. Among patients treated for more than 24 months, those "In Remission" had a mean ITD of 58 months compared to 41 months among those "Not in Remission" ($p = 0.049$). A similar trend was observed for "Remission Experienced" among patients treated for at least 36 months, where it was associated with longer ITD ($p = 0.002$) (**Table 2**).

Lower thyroid volume and TRAb levels at diagnosis correlated with "Disease Not Active ≥1 Month" ($p = 0.024$ and $p = 0.037$, respectively). For initial therapy longer than 24 months, standardized TRAb levels remained significantly lower in this group ($p = 0.035$) (**Table 2**).

Comparing the group of patients that maintained the initial ATD treatment with the group that discontinued it, only standardized TRAb levels were significantly higher: 31.64 (84.10) in Group A versus 6.57 (10.24) in Group B ($p = 0.008$).

Copyright© ACE&M all rights reserved.



ATD: antithyroid drug; GD: Graves' disease; ^{131}I: radioiodine; MTZ: methimazole; PTU: propylthiouracil.

Figure 2. Selection of patients and disease progression according to initial treatment with antithyroid drugs; disease progression in patients who discontinued initial antithyroid drug therapy (Group B) and therapeutic choices made in cases of relapse or lack of remission (antithyroid drugs or definitive therapy with radioiodine or thyroidectomy). **Group A:** maintenance of initial treatment with antithyroid drugs. **Group B:** discontinuation of initial treatment with antithyroid drugs. "Remission": patients in remission at the time of data collection. "Relapse": patients who experienced relapse after remission (without achieving a second remission). "Never in Remission": patients with inactive disease for less than 12 months or with persistently active disease.

Copyright© AE&M all rights reserved.

Table 1. Descriptive analysis of clinical and biochemical characteristics at diagnosis, details of initial antithyroid treatment, and outcomes

Sex, n (%)	
Female	29 (80.56)
Male	7 (19.44)
Other autoimmune diseases*, n (%)	
Yes	7 (19.44)
Pubertal stage, n (%)	
Pubertal	28 (77.78)
Body mass index, n (%)†	
Underweight	1 (2.86)
Normal weight	27 (77.14)
Overweight	7 (20.00)
Obese	0 (0.00)
Initial treatment with antithyroid drugs, n (%)	
Propylthiouracil	7 (19.44)
Methimazole	29 (80.56)
Age (years), median (IQR)	13.67 (4.71)
Standardized thyroid gland volume (xN), median (IQR) (n = 31)‡	3.22 (1.94)
Standardized FT4 (xN), median (IQR) (n = 32)§	2.15 (1.24)
Standardized TRAb (xN), median (IQR) (n = 29)¶	7.11 (14.90)
Initial ATD dose (mg/kg/day), mean ± SD	
PTU (n=7)	4.50 ± 1.42
Methimazole (n=29)	0.36 ± 0.02
Duration of initial ATD treatment (months), median (IQR) (n = 35)	31.00 (23.00)
PTU, mean ± SD (n = 6)**	45.50 ± 14.45
Methimazole, mean ± SD (n = 29)	33.24 ± 2.48
Duration of initial ATD treatment > 24 months (n = 24)	
Mean ± SD	44 ± 3
Median (IQR)	40 (19)
Duration of initial ATD treatment ≥ 36 months (n = 16)	
Mean ± SD	51 ± 4
Median (IQR)	46 (12)
Duration of initial ATD treatment until definitive treatment†† (n = 16)	
Mean ± SD	34 ± 3
Median (IQR)	40 (22)
Duration of initial ATD treatment until I-131 ablation†† (n = 13)	
Mean ± SD	33 ± 4
Median (IQR)	39 (25)
Duration of initial ATD treatment until thyroidectomy†† (n = 3)	
Mean ± SD	35 ± 7
Median (IQR)	40 (.)
Outcomes, n (%)	
In Remission	6 (17.65)
Remission Experienced	11 (32.35)
Disease Not Active < 12 Months	9 (26.47)
Disease Not Active ≥ 1 Month	20 (58.82)
Relapse	5 (45.45)

Results are expressed as mean ± standard deviation, median (interquartile range), or number (%). The n value may vary due to missing data. Standardization: values divided by their respective upper limit of normal.

*Includes type 1 diabetes mellitus (n = 3), celiac disease (n = 3), juvenile idiopathic arthritis (n = 1); †n = 35 due to 1 missing result in the clinical records; ‡n = 31, due to 5 missing results in the clinical records; §n = 32, due to 4 missing results in the clinical records; ¶n = 29, due to 7 missing results in the clinical records; **n = 6, due to 1 patient having a follow-up duration < 6 months; ††Duration of initial antithyroid drug treatment, in months, for patients who underwent definitive therapy (radioiodine ablation or thyroidectomy). "In Remission" – patients in remission at the time of data collection; "Remission Experienced" – group of patients who achieved remission at any point (patients in remission at data collection plus patients who relapsed and did not achieve a subsequent remission); "Disease Not Active < 12 Months" – group of patients who had a period of inactive disease lasting < 12 months, without achieving remission; "Disease Not Active ≥ 1 Month" – total of patients who experienced remission at any point and those without signs of active disease for at least 1 month but less than 12 months; "Relapse" – patients who experienced remission at any point (n = 11) and subsequently relapsed (n = 5).

FT4: free thyroxine; IQR: interquartile range; SD: standard deviation; TRAb: thyrotropin receptor antibodies.

Copyright© AE&M all rights reserved.

Table 2. Variables with $p < 0.20$ in the comparative analyses for Outcome 1 ("In Remission" versus "Not in Remission"), Outcome 2 ("Remission Experienced" versus "No Remission Experienced"), and Outcome 3 ("Disease Not Active ≥ 1 Month" versus "Disease Not Active < 1 Month"), as well as for the comparisons on initial treatment durations of more than 24 months and at least 36 months

			Test statistic	*p*
Outcome 1 (n = 34)	**In remission (n = 6)**	**Not in remission (n = 28)**		
Standardized thyroid gland volume Mean ± SD (n = 29)	2.54 ± 0.78 (n = 5)	3.57 ± 1.50 (n = 24)	*t*(27) = -1.475	0.152
Duration of initial ATD treatment Median (IQR) (n = 34)	48.50 (43.00) (n = 6)	31.00 (19.00) (n = 28)	*U* = 51.500	0.145
Outcome 2 (n = 34)	**Remission experienced (n = 11)**	**No Remission experienced (n = 23)**		
Age Median (IQR) (n = 34)	14.92 (4.92) (n = 11)	13.00 (4.16) (n = 23)	*U* = 89.500	0.176
Standardized thyroid gland volume Mean ± SD (n = 29)	2.80 ± 1.20 (n = 8)	3.61 ± 1.49 (n = 21)	*t*(27) = -1.367	0.183
Outcome 3 (n = 34)	**Disease not active ≥1 month (n = 20)**	**Disease not active <1 month (n = 14)**		
Standardized thyroid gland volume Mean ± SD (n = 29)	2.89 ± 1.14 (n = 17)	4.10 ± 1.58 (n = 12)	*t*(27) = 2.400	**0.024**
Standardized TRAb Median (IQR) (n = 27)	6.46 (8.20) (n = 15)	13.84 (51.28) (n = 12)	*U*=133.00	**0.037**
Initial ATD treatment duration > 24 months				
Outcome 1 (n = 24)	**In remission (n = 4)**	**Not in remission (n = 20)**		
Standardized thyroid gland volume Mean ± SD (n = 19)	2.62 ± 1.09 (n = 3)	4.06 ± 1.48 (n = 16)	*t*(17) = -1.595	0.129
Duration of initial ATD treatment Mean ± SD (n = 24)	58.25 ± 12.61 (n = 4)	40.80 ± 15.70 (n = 20)	*t*(22) = 2.080	**0.049**
Outcome 2 (n = 24)	**Remission experienced (n=7)**	**No remission experienced (n = 17)**		
BMI SD, n (%) (n = 23) Underweight (n = 1) Normal (n = 18) Overweight (n = 4)	 (n = 0) 33.3 (n = 6) 0.0 (n = 0)	 100 (n = 1) 66.7 (n = 12) 100 (n = 4)	LR(2) =3.448	0.175
Duration of initial ATD treatment Median (IQR) (n=24)	58.43 ± 22.86 (n = 7)	37.65 ± 7.49 (n=17)	*t*(7) = 3.354	0.053
Outcome 3 (n=24)	**Disease not active ≥ 1 month (n=13)**	**Disease not active <1 month (n = 11)**		
BMI SD, n (%) (n=23) Underweight (n = 1) Normal (n = 18) Overweight (n = 4)	 (n = 0) 61.1 (n = 11) 25.0 (n = 1)	 100 (n = 1) 38.9 (n = 7) 75.0 (n = 3)	LR(2) = 3.286	0.193
Standardized TRAb Median (IQR) (n = 19)	6.84 (9.60) (n = 10)	15.56 (79.32) (n = 9)	*U* = 71.000	**0.035**
Duration of initial ATD treatment Mean ± SD (n = 24)	48.62 ± 20.42 (n = 13)	37.91 ± 6.99 (n = 11)	*t*(15) = 1.772	0.096
Initial ATD treatment duration ≥ 36 months				
Outcome 1 (n = 16)	**In remission (n = 4)**	**Not in remission (n = 12)**		
Standardized thyroid gland volume Mean ± SD (n = 12)	2.62 ± 1.09 (n = 3)	3.62 ± 0.92 (n = 9)	*t*(10) = -1.568	0.148
Duration of initial ATD treatment Median (IQR) (n = 16)	57.00 (24.00) (n = 4)	42.50 (10.0) (n = 12)	*U* = 8.500	0.058
Outcome 2 (n = 16)	**Remission experienced (n = 6)**	**No remission experienced (n = 10)**		
Standardized thyroid gland volume Mean ± SD (n = 12)	2.62 ± 1.09 (n = 3)	3.62 ± 0.92 (n = 9)	*t*(10) = -1.568	0.148
Duration of initial ATD treatment Median (IQR) (n = 16)	57.00 (30.00) (n = 6)	41.00 (8.00) (n = 10)	*U* = 3.000	**0.002**
Outcome 3 (n = 16)	**Disease not active ≥ 1 month (n = 10)**	**Disease not active <1 month (n = 6)**		
Standardized TRAb Median (IQR) (n = 11)	6.67 (8.20) (n = 7)	13.84 (76.94) (n = 4)	*U* = 24.000	0.073
Duration of initial ATD treatment Median (IQR) (n = 16)	50.50 (24.0) (n = 10)	41.50 (8.00) (n = 6)	*U* = 16.500	0.147

Results are expressed as mean ± standard deviation, median (interquartile range), or number (%); the n value may vary due to missing data. Standardization: values divided by their respective upper limit of normal. "In Remission": patients in remission at the time of data collection. "Remission Experienced": patients who experienced remission at some point (those in remission at the time of data collection, as well as those who relapsed without subsequently achieving a new remission). "Disease Not Active ≥ 1 Month": patients who experienced remission at some point, as well as those with inactive disease for a period of at least 1 month and less than 12 months.

ATD: antithyroid drug; BMI: body mass index; IQR: interquartile range; SD: standard deviation; TRAb: thyrotropin receptor antibodies.

Copyright© AE&M all rights reserved.

In univariate LR, ITD was a significant predictor of "Remission Experienced" in the group with initial therapy longer than 24 months (p = 0.041; OR = 1.14; 95% CI: 1.005–1.297), explaining 47% of outcome variability. For "Disease Not Active ≥ 1 Month", thyroid volume was initially significant (p = 0.042; OR = 0.49) but lost significance after adjustment for TRAb levels (**Table 3**). No further univariate or multivariate LR results reached significance.

An analysis of the ROC curve confirmed the discriminative ability of ITD for remission outcomes. For therapy longer than 24 months, ITD showed excellent performance (AUC = 0.894 for "In Remission" and 0.840 for "Remission Experienced"), with an optimal cutoff of 51 months. For treatment ≥ 36 months, the discriminative ability remained high (AUC = 0.823 and 0.950, respectively) (**Figure 3**).

Standardized TRAb levels demonstrated acceptable discrimination for "Disease Not Active ≥ 1 Month" (AUC = 0.739, p = 0.036), improving to excellent for treatment ≥ 36 months (AUC = 0.857, p = 0.059). The optimal TRAb cutoff point was 7.055-fold the ULN (**Figure 3**).

For standardized thyroid volume, discrimination was acceptable for "Disease Not Active ≥ 1 Month" (AUC = 0.752, p = 0.023), with an optimal cutoff point of 2.50 (**Figure 3**).

DISCUSSION

The optimal treatment for GD in children remains controversial due to the limited efficacy of available therapeutic options and the difficulty in predicting remission. Although ATDs are the recommended first-line therapy (3,7,9), long-term remission after 2 years

Table 3. Odds ratio of variables selected from the comparative analyses between groups (Table 2) for the occurrence of the events "In Remission", "Remission Experienced", and "Disease Not Active ≥ 1 Month"

	Variable	Unadjusted OR			Adjusted OR		
		OR	95% CI	*p*	OR	95% CI	*p*
Outcome 1	Standardized thyroid gland volume	0.462	0.156–1.362	0.162*	0.345	0.080–1.490	0.154
	Duration of initial ATD treatment	1.033	0.987–1.082	0.162*	1.078	0.977–1.191	0.136
Outcome 1 (>24 months)	Standardized thyroid gland volume	0.310	0.064–1.504	0.146*	0.063	0.000–8.311	0.267
	Duration of initial ATD treatment	1.058	0.990–1.132	0.099*	1.800	0.787–4.117	0.164
Outcome 1 (≥36 months)	Standardized thyroid gland volume	0.311	0.059–1.630	0.167*	-	-	-
	Duration of initial ATD treatment	1.038	0.967–1.115	0.297	-	-	-
Outcome 2	Age	1.090	0.869–1.367	0.456	-	-	-
	Standardized thyroid gland volume	0.602	0.283–1.280	0.188*	-	-	-
Outcome 2 (> 24 months)	BMI SD	0.376	0.035–4.037	0.420	-	-	-
	Duration of initial ATD treatment	1.142	1.005–1.297	**0.041***	-	-	-
Outcome 2 (≥36 months)	Standardized thyroid gland volume	0.311	0.059–1.630	0.167*	0.064	0.000–8.395	0.269
	Duration of initial ATD treatment	1.483	0.958–2.296	0.077*	1.793	0.778–4.129	0.170
Outcome 3	Standardized thyroid gland volume	0.492	0.249–0.975	**0.042***	0.519	0.217–1.244	0.142
	Standardized TRAb	0.943	0.871–1.020	0.143*	0.937	0.846–1.038	0.211
Outcome 3 (>24 months)	BMI SD	0.607	0.094–3.943	0.601	-	-	-
	Standardized TRAb	0.923	0.818–1.041	0.190*	0.923	0.811–1.053	0.226
	Duration of initial ATD treatment	1.060	0.981–1.146	0.141*	1.053	0.933–1.188	0.403
Outcome 3 (≥36 months)	Standardized TRAb	0.902	0.718–1.132	0.373	-	-	-
	Duration of initial ATD treatment	1.143	0.936–1.395	0.190*	-	-	-

Variables with $p < 0.20$ in the comparative analysis between groups were included in the univariate logistic regression (unadjusted odds ratio [OR]). Those with $p < 0.20$ in the univariate logistic regression (*) were included in the multivariate logistic regression (adjusted OR). Adjusted ORs with a $p < 0.05$ were considered statistically significant. Standardization: values divided by their respective upper limit of normal. "In Remission": patients in remission at the time of data collection. "Remission Experienced": patients who experienced remission at some point (patients in remission at the time of data collection, as well as those who relapsed without subsequently achieving a new remission). "Disease Not Active ≥ 1 Month": patients who experienced remission at some point, as well as those with inactive disease for a period of at least 1 month and less than 12 months.
ATD: antithyroid drug; BMI: body mass index; OR: odds ratio; SD: standard deviation; TRAb: thyrotropin receptor antibodies.

Copyright© AE&M all rights reserved.

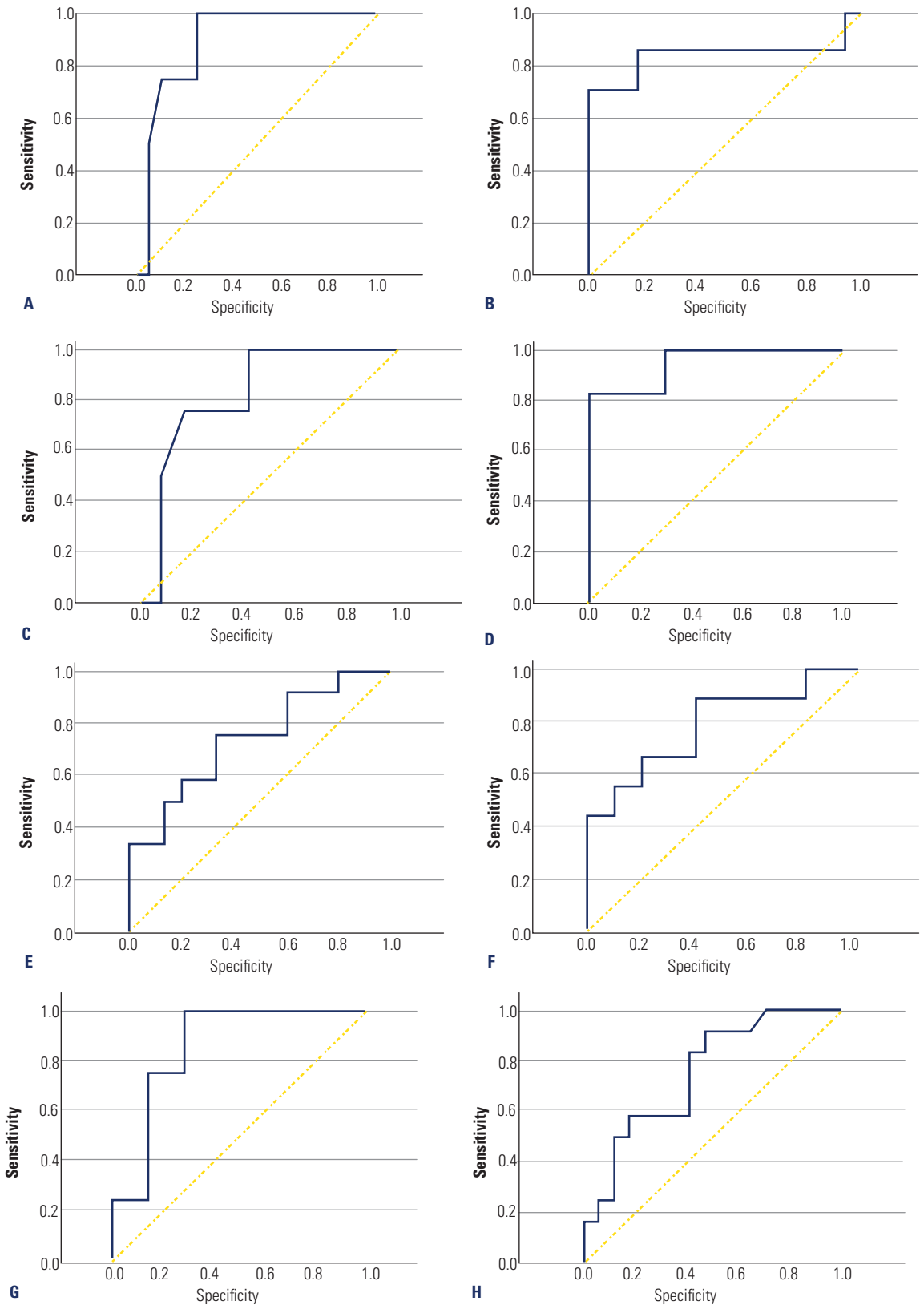


Figure 3. Receiver operating characteristic (ROC) curves. (**A**) ROC curve assessing the discriminative ability of the variable "Initial Treatment Duration" for the event "In Remission" (**Outcome 1**, treatment > 24 months). (**B**) ROC curve assessing the discriminative ability of the variable "Initial Treatment Duration" for the event "Remission Experienced" (**Outcome 2**, treatment > 24 months). (**C**) ROC curve assessing the discriminative ability of the variable "Initial Treatment Duration" for the event "In Remission" (**Outcome 1**, treatment ≥ 36 months). (**D**) ROC curve assessing the discriminative ability of the variable "Initial Treatment Duration" for the event "Remission Experienced" (**Outcome 2**, treatment ≥ 36 months). (**E**) ROC curve assessing the discriminative ability of the variable "Standardized TRAb" for the event "Disease Not Active ≥ 1 Month" (**Outcome 3**).

Copyright© AE&M all rights reserved.

is achieved in only 20%–30% of patients (1,3,13). The ETA currently advises at least 3 years of ATD treatment, extendable to 5 or more years depending on clinical and biochemical evolution (3). Identifying reliable prognostic markers at diagnosis could help individualize therapy and avoid unnecessary exposure to prolonged medication or premature definitive treatment.

Baseline characteristics in our sample — predominance of females, pubertal stage at diagnosis, and normal BMI — were consistent with previous studies (4–6,10,13,14). Notably, PTU was prescribed only before 2010, reflecting historical practice. Following INFARMED's (National Authority of Medicines and Health Products, I.P. in Portugal) 2010 recommendation highlighting the hepatotoxicity risk of PTU, its use in pediatric patients was gradually discontinued, and methimazole became the preferred first-line ATD in accordance with national and international guidelines.

Longer ITD was associated with remission outcomes, suggesting a potential benefit of extended therapy. Analysis of the ROC curve identified an optimal ITD of approximately 51 months, consistent with ETA recommendations and other reports showing increased remission rates with prolonged therapy (1,3,4). Studies by Léger and cols. and Ohye and cols. also observed higher remission rates after 4–6 years of ATD treatment (11,12). These findings support the rationale for maintaining ATD therapy for at least 3 years.

Smaller thyroid volume at diagnosis was associated with disease inactivity, in line with previous studies reporting poorer outcomes in patients with larger goiters (9). Analysis of the ROC curve in our sample suggested that a thyroid volume ≥ 2.5-fold the ULN reduced the likelihood of disease inactivity.

Lower TRAb levels at diagnosis were also associated with better outcomes. Similar trends have been reported by Gastaldi and cols. (8) and Rho and cols. (14), who showed that persistently high TRAb titers predicted relapse or failure to achieve remission. Our cutoff of 7.0-fold the ULN aligns with these results and supports the role of TRAb monitoring in guiding clinical decisions.

No statistically significant associations were found between remission outcomes and sex, pubertal stage, BMI, FT4 levels, or the type of initial ATD used (PTU versus methimazole), consistent with prior reports (8,14,33). Although the coexistence of other AIDs was more frequent among patients who experienced remission or relapse, this trend did not reach statistical significance and warrants further investigation. A similar observation was reported by Gatta and cols., who suggested that concurrent autoimmune conditions might modulate immune activity and favor GD remission. According to the GRAPHE study, the most reasonable hypothesis is that the immunological cascade involving multiple autoimmune diseases may contribute to shifts in TRAb profiles, particularly from stimulating to blocking antibodies, together with a lower overall immune response potentially related to differences in B-cell subsets (34).

Although our findings support the potential benefit of prolonged ATD therapy, the clinical implications of extended treatment must be interpreted with caution. This study did not systematically evaluate treatment-related adverse effects or adherence to long-term oral therapy, both of which are clinically relevant considerations, particularly in pediatric and adolescent populations. While serious adverse events associated with ATDs are uncommon, prolonged treatment may increase the cumulative burden of minor side effects and pose challenges to sustained adherence, potentially influencing real-world effectiveness. Therefore, decisions regarding extended ATD therapy should be individualized, balancing the potential benefits of remission against treatment tolerability, patient adherence, and family preferences.

Importantly, although several variables demonstrated significant associations with remission-related outcomes in univariate analyses, no independent predictors remained statistically significant after multivariate adjustment. This likely reflects limited statistical power due to the small sample size and the low number of outcome events. Therefore, the identified associations should be interpreted as exploratory and hypothesis-generating, rather than establishing independent predictive or causal relationships. Larger, adequately powered multicenter studies are required to confirm these findings and to determine whether these factors retain independent prognostic value in PGD.

Copyright© AE&M all rights reserved.

Limitations

This study has several limitations. The small sample size and retrospective single-center design limited statistical power, particularly for subgroup analyses, and did not allow standardized long-term follow-up. Missing clinical, laboratory, and imaging data further constrained the analyses. Inter-laboratory variability and suboptimal standardization methods over the long study period may have influenced biochemical measurements. In addition, the small number of outcome events reduces the reliability of multivariate models and ROC curve analyses; therefore, any derived cutoff points or predictive estimates should be considered exploratory rather than confirmatory. Larger prospective multicenter studies with standardized follow-up are needed to validate these findings and refine prognostic markers in PGD.

In conclusion, in this 20-year retrospective cohort study, prolonged ATD therapy was associated with higher remission rates, supporting treatment durations of at least 3 years and identifying an optimal cutoff of around 51 months. In addition, smaller thyroid volume and lower TRAb levels at diagnosis were associated with disease inactivity, suggesting a supportive role of these markers in clinical risk stratification. Given the small sample size and the retrospective single-center design, further multicenter studies are needed to validate these associations and refine prognostic tools for clinical practice.

Statement of use of generative AI and AI-Assisted technologies in the writing process: during the preparation of this article, the authors used ChatGPT (OpenAI) for language editing and text refinement. The authors subsequently reviewed and edited the content as needed and take full responsibility for the final manuscript.

Funding: this research received no external funding.

Disclosure: no potential conflict of interest relevant to this article was reported.

Data availability: datasets related to this article will be available upon request to the corresponding author.

REFERENCES

1. Léger J, Kaguelidou F, Alberti C, Carel JC. Graves' disease in children. Ann Endocrinol (Paris). 2018;79(6):647-55. doi: 10.1016/j.ando.2018.08.001.
2. Sociedade Portuguesa de Endocrinologia e Diabetologia Pediátrica. Abordagem da Doença de Graves na Criança e Adolescente – Protocolo de Diagnóstico e Tratamento [under revision since early 2024 – not yet published].
3. Mooij CF, Cheetham TD, Verburg FA, Eckstein A, Pearce SH, Léger J, et al. 2022 European Thyroid Association Guideline for the management of pediatric Graves' disease. Eur Thyroid J. 2022;11(1):e210073. doi: 10.1530/ETJ-21-0073.
4. Song A, Tan L, Ong YH, Chai JYH, Ang LW, Lim R, et al. Long-term antithyroid drug treatment of Graves' disease in children and adolescents: A 20-year single-center experience. Front Endocrinol (Lausanne). 2021;12:687834. doi: 10.3389/fendo.2021.687834.
5. Kaguelidou F, Alberti C, Castanet M, Guitteny MA, Czernichow P, Léger J, et al. Predictors of autoimmune hyperthyroidism relapse in children after discontinuation of antithyroid drug treatment. J Clin Endocrinol Metab. 2008;93(10):3817-26. doi: 10.1210/jc.2008-0842.
6. Marques O, Antunes A, Oliveira MJ. Treatment of Graves' disease in children: The Portuguese experience. Rev Endocrinol Diabetes Nutr. 2018;65(3):143-9. doi: 10.1016/j.endinu.2017.11.014.
7. 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(10):1343-421. doi: 10.1089/thy.2016.0229.
8. Gastaldi R, Poggi E, Mussa A, Weber G, Vigone MC, Salerno M, et al. Graves' disease in children: Thyroid-stimulating hormone receptor antibodies as remission markers. J Pediatr. 2014;164(5):1189-94.e1. doi: 10.1016/j.jpeds.2013.12.047.
9. Glaser NS, Styne DM. Predictors of early remission of hyperthyroidism in children. J Clin Endocrinol Metab. 1997;82(6):1719-26. doi: 10.1210/jcem.82.6.3986.
10. Wong T, Ngan E, Lee ACW, Fan CY, Ng SM. Remission in pediatric Graves' disease treated with antithyroid drug and the risk factors associated with relapse. Ann Pediatr Endocrinol Metab. 2022;27(4):308-14. doi: 10.6065/apem.2244038.019.
11. Léger J, Gelwane G, Kaguelidou F, Benmerad M, Alberti C; French Childhood Graves' Disease Study Group. Positive impact of long-term antithyroid drug treatment on the outcome of children with Graves' disease: national long-term cohort study. J Clin Endocrinol Metab. 2012 Jan;97(1):110-9. doi: 10.1210/jc.2011-1944.
12. Ohye H, Minagawa A, Noh JY, Mukasa K, Kunii Y, Watanabe N, et al. Antithyroid drug treatment for Graves' disease in children: A long-term retrospective study at a single institution. Thyroid. 2014;24(2):200-7. doi: 10.1089/thy.2012.0612.
13. Jevalikar G, Rathi M, Saran S, Jain A, Bhansali A, Dutta P, et al. Long-term outcomes of pediatric Graves' disease. J Pediatr Endocrinol Metab. 2014;27(11-12):1131-6. doi: 10.1515/jpem-2013-0342.
14. Rho J, Jeong HR, Kim S, Lee HS, Hwang JS. Long-term outcomes of Graves' disease in children and adolescents receiving antithyroid drugs. Ann Pediatr Endocrinol Metab. 2021;26(4):266-71. doi: 10.6065/apem.2040286.143.
15. Rabon S, Burton AM, White PC. Graves' disease in children: Long-term outcomes of medical therapy. Clin Endocrinol (Oxf). 2016;85(4):632-5. doi: 10.1111/cen.13099.
16. Gu Y, He H, Xie H, Chen C, Chen Y, Xu X, et al. Clinical features and predictors of remission in children under the age of 7 years with Graves' disease. Pediatr Investig. 2020;4(3):198-203. doi: 10.1002/ped4.12219.
17. World Health Organization. Child growth standards. Geneva: WHO; 2024. Available from: https://www.who.int/toolkits/child-growth-standards/standards/body-mass-index-for-age-bmi-for-age. Accessed August 11, 2025.
18. World Health Organization. BMI-for-age (5–19 years). Geneva: WHO; 2024. Available from: https://www.who.int/tools/growth-reference-data-for-5to19-years/indicators/bmi-for-age. Accessed August 11, 2025.
19. Tanner JM. Growth and maturation during adolescence. Nutr Rev. 1981;39(2):43-55. doi: 10.1111/j.1753-4887.1981.tb06734.x.

Copyright© AE&M all rights reserved.

20. Konca C, Elhan AH. Unveiling the accuracy of ultrasonographic assessment of thyroid volume: A comparative analysis of ultrasonographic measurements and specimen volumes. J Clin Med. 2023;12(20):6619. doi: 10.3390/jcm12206619.
21. LaFranchi S. Approach to acquired goiter in children and adolescents. UpToDate. 2023. Available from: https://www.uptodate.com/contents/approach-to-acquired-goiter-in-children-and-adolescents. Accessed August 4, 2025.
22. Yebra Yebra J, Oliva Pérez P, Gómez Lorente JL, Guerrero-Fernández J. Disfunción tiroidea. Interpretación de las pruebas tiroideas. In: Manual de Diagnóstico y Terapéutica en Endocrinología Pediátrica v1.1. Cap. 36.
23. Sample Size Calculator. *Calculator.net*; 2024. Available from: https://www.calculator.net/sample-size-calculator.html?type=1&cl=95. Accessed July 10, 2024.
24. Chan YH. Biostatistics 101: Data presentation. Singapore Med J. 2003;44(6):280-5. Available from: http://www.smj.org.sg/sites/default/files/4406/4406bs1.pdf. Accessed August 11, 2025.
25. Chan YH. Biostatistics 102: Quantitative data – Parametric & non-parametric tests. Singapore Med J. 2003;44(8):391-6. Available from: http://www.smj.org.sg/sites/default/files/4408/4408bs1.pdf. Accessed August 11, 2025.
26. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd ed. Hillsdale, NJ: L Erlbaum Associates; 1988. Available from: https://utstat.utoronto.ca/brunner/oldclass/378f16/readings/CohenPower.pdf. Accessed August 11, 2025.
27. Chan YH. Biostatistics 103: Qualitative data – Tests of independence. Singapore Med J. 2003;44(10):498-503. Available from: https://paulogentil.com/pdf/Biostatistics%20103.%20Qualitative%20data%20-%20tests%20of%20independence.pdf. Accessed August 11, 2025.
28. Chan YH. Biostatistics 202: Logistic regression analysis. Singapore Med J. 2004;45(4):149-53. Available from: https://lo.unisa.edu.au/pluginfile.php/1020307/mod_book/chapter/147886/Logistic%20Regression%20Analysis.pdf. Accessed August 11, 2025.
29. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol. 1996;49(12):1373-9. doi: 10.1016/s0895-4356(96)00236-3.
30. van Smeden M, de Groot JA, Moons KG, Collins GS, Altman DG, Eijkemans MJ, et al. No rationale for 1 variable per 10 events criterion for binary logistic regression analysis. BMC Med Res Methodol. 2016;16:163. doi: 10.1186/s12874-016-0267-3.
31. Vittinghoff E, McCulloch CE. Relaxing the rule of ten events per variable in logistic and Cox regression. Am J Epidemiol. 2007;165(6):710-8. doi: 10.1093/aje/kwk052.
32. Zou KH, O'Malley AJ, Mauri L. Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models. Circulation. 2007;115(5):654-7. doi: 10.1161/CIRCULATIONAHA.105.594929.
33. Sato H, Minagawa M, Sasaki N, Sugihara S, Kazukawa I, Minamitani K, et al. Comparison of methimazole and propylthiouracil in the management of children and adolescents with Graves' disease: Efficacy and adverse reactions during initial treatment and long-term outcome. J Pediatr Endocrinol Metab. 2011;24(5-6):257-63. doi: 10.1515/jpem.2011.194.
34. Gatta E, Pirola I, Gotti A, Fredi M, Bellini P, Dondi F, et al. Clinical phenotype of Graves' disease in autoimmune polyglandular syndrome or as isolated disease: The GRAPHE study. J Clin Endocrinol Metab. 2025;dgaf144. doi: 10.1210/clinem/dgaf144.

Copyright© AE&M all rights reserved.