Predictors of outcome and comparison of different drug regimens for the prevention of relapse in patients with Graves’ disease

Bjørn G Nedrebø 1,6, Pål I Holm 1, Sverre Uhving 4, Jan Inge Sørheim 5, Svein Skeie 4, Geir Egil Eide 3, Eystein S Husebye 1, Ernst A Lien 3 and Sylvi Aanderud 1

1 Division of Endocrinology, Department of Internal Medicine, 2 Centre for Clinical Research & Section for Medical Statistics, University of Bergen, 3 Department of Biochemical Endocrinology, Haukeland University Hospital, Bergen, Norway, 4 Department of Internal Medicine, Rogaland Central Hospital, Stavanger, Norway, 5 Department of Internal Medicine, Stord Hospital, Stord, Norway and 6 Department of Internal Medicine, Haugesund Hospital, Haugesund, Norway

(Correspondence should be addressed to Bjørn G Nedrebø, Division of Endocrinology, Department of Internal Medicine, Haukeland University Hospital, N-5021 Bergen, Norway; Email: bjorn.nedrebo@ikb.uib.no)

Abstract

Objective: To investigate the effect of different antithyroid drug (ATD) regimens on relapse rates of Graves’ disease, and to look for predictors of relapse.

Design and Methods: In a prospective two-way factorial randomized clinical trial, 218 patients with Graves’ disease were assigned to ATD combined with L-thyroxine (L-T4) or ATD alone for 12 months. After discontinuation of antithyroid therapy, each group was stratified to either 12 months further treatment with L-T4 or no treatment. Clinical and biochemical assessments were carried out before treatment, after 3 and 6 weeks, and every third month for 12 months. If the patients lacked symptoms of relapse, laboratory tests were performed every third month for the second year, and thereafter annually.

Results: The proportion of all patients with relapse was 47.7% 2 years after withdrawal of ATD. There was no difference in relapse rates between the treatment groups (P = 0.217; log-rank test). Smokers had a higher relapse rate than non-smokers (58.4% vs 38.8%, P = 0.009). Patients who were thyrotropin-receptor antibody (TRAb) positive after 12 months of antithyroid therapy had a higher relapse rate than those who were negative (72.5% vs 36.8%, P = 0.0001). Similarly, relapse was more frequent (55.5%) in patients having large goiter compared with subjects with small goiter (36.3%, P = 0.0007).

Conclusions: Relapse rates of Graves’ disease were independent of ATD regimen whether followed by L-T4 therapy or not. Smoking, large goiter and presence of TRAb at the end of ATD therapy were strong predictors of relapse.

European Journal of Endocrinology 147 583–589

Introduction

Graves’ disease is an autoimmune disorder with causative relationship to elevated thyrotropin-receptor antibodies (TRAb) (1). The prevalence of Graves’ disease is 2–2.5% in females and 0.2–0.6% in males (2, 3). Radioiodine ablation of the thyroid gland, surgery and medical therapy with antithyroid drugs (ATD) are the therapeutic options. The most frequent therapeutic policy in Europe to achieve euthyroidism in Graves’ disease is to use ATD over a period of several months (4). Modalities of ATD treatment are considerably different between countries (4). The optimal regimen has been difficult to establish because the natural course of the disease is variable, and reliable predictors of disease activity have yet to be identified (5). Suppression of thyrotropin (TSH) stimulation of the thyroid gland by adding exogenous L-thyroxine (L-T4) to ATD treatment may be beneficial (6, 7), and direct immunomodulatory effects of high doses of ATD or L-T4 have also been proposed (1, 8). In clinical studies, however, the relapse rates of Graves’ disease have varied from 2 to 60% (8, 9), and the impact of different treatment regimens and predictors of outcome, such as TRAb and smoking, remains a matter of controversy (3, 10–13). It is therefore of clinical interest to identify treatment regimens that lower relapse rates.

We conducted a randomized factorial clinical trial to compare the effect of different ATD regimens on the overall relapse rate and to search for prognostic factors.
of outcome. Patients were either randomized to ATD combined with l-T4 (‘block-replace’ therapy) or titration of ATD doses (titration therapy) for 12 months. After discontinuation of ATD therapy, both groups were stratified to receive l-T4 or no medication for the subsequent 12 months.

 Patients and methods

Subjects and study design

Between January 1997 and July 1999, 218 consecutive patients with Graves’ disease (214 Caucasian, four Asiatic) between 16 and 75 years of age were enrolled. The diagnosis was based on the clinical signs of hyperthyroidism combined with suppressed serum TSH and positive TRAb or ophthalmopathy. The patients were recruited from four hospitals in Norway.

Patients were randomized in a two-way factorial design to one of two main groups each with two subgroups (1A, 1B, 2A and 2B) (Fig. 1). All patients were treated with ATD for 12 months, either with block-replace therapy (group 1 = 110) or with a titration regimen (group 2 = 108). For the subsequent 12 months, group 1A (n = 54) and group 2A (n = 54) continued with l-T4 alone, whereas group 1B (n = 56) and group 2B (n = 54) stopped all medication (Fig. 1). One patient in group 1 and five patients in group 2 started with propylthiouracil at a daily dose of 200–400 mg. All other patients received carbimazole at an initial mean dose of 29.7 mg per day (15–45 mg) in group 1 and 29.3 mg per day (5–45 mg) in group 2. Patients with mild drug reactions to carbimazole were switched to propylthiouracil treatment if they did not refuse to use ATD. Once biochemical euthyroidism was achieved in group 1, l-T4 was added to maintain normal serum levels of free thyroxine (FT₄). When patients in group 2 became euthyroid, ATD was adjusted to maintain normal serum levels of FT₄. Beta-blockers were given initially according to clinical judgment. Clinical and biochemical assessments were carried out before treatment, after 3 and 6 weeks, and every third month during the first year of treatment. The size of goiter was categorized in two equal groups on the basis of physical examination and measurement of length and width of each thyroid lobe: none or small, and large. If the patients had no symptoms indicating relapse, laboratory tests were taken every third month for the second year, and thereafter annually.

Exclusion criteria were as follows: pregnancy, treatment with ATD in the 12 months prior to enrollment, allergy to ATD, ongoing immunosuppressive treatment, non-compliance because of psychiatric or other serious diseases, patients’ preference for surgery or radioiodine treatment, or unwillingness to participate in the study.

Twenty-seven patients (seven in group 1A, four in group 1B, seven in group 2A and nine in group 2B) dropped out before 12 months of treatment. The reasons were pregnancy (n = 3), change in residence (n = 1), non-compliance (n = 4), treatment with surgery or radioiodine (n = 5), development of blocking TRAb (n = 2), side effects of ATD (n = 11), and aggressive ophthalmopathy (n = 1). After 12 months of ATD treatment, two patients (one in group 1B, one in group 2A) refused to continue in the study. Until August 2001, 189 patients (group 1 = 98, 1A = 47, 1B = 51, and group 2 = 91, 2A = 46, 2B = 45) (Fig. 1) had been followed between 13.4 and 41.7 months.

Figure 1 Study design. Bold lines denote periods of active treatment (T) as indicated. Groups 1A/1B: ATD + l-T4 for the first 12 months (block-replace regimen), followed by 12 months with (group 1A) or without (group 1B) l-T4. Groups 2A/2B: ATD for the first 12 months (titration regimen) followed by 12 months with (group 2A) or without (group 2B) l-T4. Number of patients who dropped out during the ATD treatment period: 1A = 7, 1B = 5, 2A = 8, 2B = 9.
after discontinuation of ATD therapy, except for two patients who became pregnant at 3.5 and 5.2 months follow-up and were then excluded.

Relapse was defined as FT4 > 25 pmol/l (normal range 8–20 pmol/l) combined with TSH < 0.05 mIU/l. When relapse was suspected in patients using L-T4 (groups 1A and 2A), L-T4 was discontinued for 3 months follow-up and were then excluded.

The Regional Ethics Committee, University of Bergen, approved the study protocol, and the study subjects gave their informed consent prior to participation in the study.

**Biochemical methods**

TSH and FT4 in serum were measured using the AutoDELFIA hTSH Ultra kit and AutoDELFIA FT4 kit respectively (Wallac Oy, Turku, Finland). The within-assay variation of TSH was 4.9% in the range between 0.5 and 8.3 mIU/l. The within-assay variation of FT4 was <4% between 11 and 23 pmol/l. TRAb was determined by radioimmunoassay (DLD Diagnostika GmbH, Hamburg, Germany). The within-assay variation was 5–7%. TRAb was defined as positive above 13 U/l.

**Statistical analysis**

Clinical characteristics of the patients before and after 12 months of ATD treatment are given as proportions, medians or means ± S.D. as indicated in Tables 1 and 2. Comparability of the four randomized groups with respect to clinical characteristics was investigated using the chi-square test and analysis of variance, parametric and nonparametric as appropriate. Time from discontinuation of ATD treatment to relapse was analyzed with Kaplan–Meier curves to describe proportion of relapse-free patients according to treatment modalities, smoking, TRAb and goiter size (14). If no relapse occurred, time to end of follow-up was considered as a censored observation time. The differences in the treatment effects were assessed with the log-rank test (15). Multivariate survival analysis was done with Cox’s proportional hazards model producing estimates of hazard ratios with 95% confidence intervals (16). To obtain a statistical power of 80% to detect a reduction from 40% to 20% between group 1 and group 2 after 2 years at a significance level of 5%, the estimated number of patients required was >89 in each group. All statistical analyses were performed with Statistical Package for Social Sciences (SPSS, Chicago, IL, USA), version 11.0.

### Table 1 Characteristics of the patients with Graves’ disease before treatment. Values are means ± S.D.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Group 1 (n = 110)</th>
<th>Group 2 (n = 108)</th>
<th>P-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male sex (%)</td>
<td>14.8</td>
<td>16.1</td>
<td>0.471†</td>
</tr>
<tr>
<td>Age (year)</td>
<td>43.9±10.2</td>
<td>40.2±11.5</td>
<td>0.362</td>
</tr>
<tr>
<td>Body mass index (kg/m²)</td>
<td>23.0±3.0</td>
<td>23±3.3</td>
<td>0.158</td>
</tr>
<tr>
<td>Systolic blood pressure (mmHg)</td>
<td>130±18</td>
<td>126±18</td>
<td>0.760</td>
</tr>
<tr>
<td>Diastolic blood pressure (mmHg)</td>
<td>72±9</td>
<td>71±11</td>
<td>0.450</td>
</tr>
<tr>
<td>Heart rate (beats per minute)</td>
<td>91±16</td>
<td>94±18</td>
<td>0.375</td>
</tr>
<tr>
<td>Loss of weight before treatment (kg)</td>
<td>5.3±4.5</td>
<td>5.4±4.8</td>
<td>0.846</td>
</tr>
<tr>
<td>Smokers (%)</td>
<td>40.7</td>
<td>41.1</td>
<td>0.979†</td>
</tr>
<tr>
<td>Daily number of cigarettes, regular smokers</td>
<td>11.1±7.5</td>
<td>12.0±5.4</td>
<td>0.940</td>
</tr>
<tr>
<td>TRAb-negative (%)</td>
<td>76.6</td>
<td>73.9</td>
<td>0.428†</td>
</tr>
<tr>
<td>FT4 (pmol/l)‡</td>
<td>14.0</td>
<td>6.2</td>
<td>0.009§</td>
</tr>
<tr>
<td>T3 (nmol/l)</td>
<td>1.8±0.3</td>
<td>1.9±0.5</td>
<td>0.638</td>
</tr>
<tr>
<td>T4 (nmol/l)†</td>
<td>26.5±6.2</td>
<td>27.2±5.0</td>
<td>1.000.919¶</td>
</tr>
</tbody>
</table>

* ANOVA unless otherwise indicated; † Chi-square; ‡ P-value = 0.003 is between group 1B and groups 2A/2B; § P-value = 0.009 is between group 1A and group 2A; ¶ P-value = 1.0 between group 1A and group 1B, P-value = 0.919 between group 2A and group 2B.

### Table 2 Characteristics of the patients with Graves’ disease after 12 months of antithyroid drug treatment. Values are means ± S.D.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Group 1A (n = 47)</th>
<th>Group 1B (n = 51)</th>
<th>Group 2A (n = 46)</th>
<th>Group 2B (n = 45)</th>
<th>P-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRAb-negative (%)</td>
<td>76.6</td>
<td>64.7</td>
<td>73.9</td>
<td>62.2</td>
<td>0.428†</td>
</tr>
<tr>
<td>TSH (mIU/l)</td>
<td>2.5±3.0</td>
<td>1.9±2.4</td>
<td>1.9±2.5</td>
<td>1.7±1.5</td>
<td>0.370</td>
</tr>
<tr>
<td>TRAb (U/l)</td>
<td>15.0±19.8</td>
<td>21.3±30.0</td>
<td>26.0±64.6</td>
<td>18.8±18</td>
<td>0.638</td>
</tr>
<tr>
<td>FT4 (pmol/l)</td>
<td>14.0±2.5</td>
<td>15.4±4.2</td>
<td>12.8±3.5</td>
<td>12.8±4.1</td>
<td>0.003§</td>
</tr>
<tr>
<td>Dose of carbimazole (mg per day)</td>
<td>26.5±6.2</td>
<td>27.2±5.0</td>
<td>11.1±5.7</td>
<td>9.6±3.0</td>
<td>1.000.919¶</td>
</tr>
<tr>
<td>Dose of L-T4 (mg per day)</td>
<td>0.093±0.016</td>
<td>0.093±0.021</td>
<td>0.093±0.021</td>
<td>0.093±0.021</td>
<td>1.0</td>
</tr>
</tbody>
</table>

* ANOVA unless otherwise indicated; † Chi-square; ‡ P-value = 0.003 is between group 1B and groups 2A/2B; § P-value = 0.009 is between group 1A and group 2A; ¶ P-value = 1.0 between group 1A and group 1B, P-value = 0.919 between group 2A and group 2B.

www.eje.org
Results

Characteristics of patients

The characteristics of the treatment groups at baseline and after 12 months of ATD treatment are shown in Tables 1 and 2. There were no differences in sex, age, body mass index, blood pressure, heart rate, smoking habits, loss of weight before treatment, serum FT4, total tri-iodothyronine (T3) and TRAb at baseline. As expected, there were more women (n = 188, 86%) than men (n = 30) in the study groups. Age of the women (mean 41.5, range 17.0–72.5 years) was lower than that of men (mean 48.5, range 27.0–67.0 years). 42% of the women and 43% of the men smoked tobacco daily.

Relapse rate of hyperthyroidism

Twenty-four months after withdrawal of the specific antithyroid therapy, the overall estimated relapse rate of hyperthyroidism was 47.7%. There was no difference in relapse rate between the four groups (group 1A = 44.3%, group 1B = 55.3%, group 2A = 52.6%, group 2B = 37.8%, P = 0.217 by log–rank test) after an estimated median relapse-free time of 35.9 months. No difference was found between groups 1 and 2 (49% vs 46.2%, P = 0.889), or between groups A and B (48.5% vs 46.9%, P = 0.763). The proportions of relapse-free patients are shown in Fig. 2. Cox’s multivariable proportional hazards model showed a positive effect (less relapse) of treatment with L-T4 after blocking therapy (group 1), but a negative effect (more relapse) after titration therapy (group 2). Group 1B had a hazard ratio of 1.41 for relapse compared with group 1A, while group 2A had a hazard ratio of 1.71(= (1.41 × 0.415)^(-1)) compared with group 2B. The likelihood ratio test gave P = 0.039 for the interaction term (Table 3).

Factors influencing the relapse rate

Using a backward stepwise Cox’s proportional hazards analysis, regular cigarette smoking, large goiter size and positive TRAb at 12 months of treatment were associated with an increased relapse risk (Table 4, Fig. 3). TRAb fell during the 12 months of ATD treatment, from 70.0±87.5U/l to 18.6±25.6U/l in group 1 and from 74.8±135.3U/l to 21.1±45.7U/l in group 2. There was no difference in mean serum concentrations of TRAb or in the proportions of patients in whom TRAb were present between the groups before or after 3, 6 and 12 months of ATD treatment.

Age, gender, body mass index, weight loss before treatment, weight gain during ATD treatment, ATD doses at 12 months of treatment in group 2, and thyroid disease in first-degree relatives did not influence the relapse risk.
Changes of thyroid hormone levels

After 6 weeks of ATD therapy, 77% of the patients had achieved FT4 within the normal range. The mean time ± s.d. required to obtain serum FT4 within the normal range was 6.0±5.8 weeks. Gender and smoking habits did not influence time to normalization. After 12 months of ATD treatment, FT4 in group 1B was higher than in groups 2A and 2B, and T3 in group 2A was higher than in group 1A (Table 2).

Adverse effects

Twenty-one patients developed a rash due to carbimazole, and six of them refused further ATD treatment. The remaining patients were switched to propylthiouracil treatment, of whom five dropped out because of increasing concentrations of alanine aminotransferase (n = 3) or the development of rash (n = 2). None of the patients developed agranulocytosis. Mild drug reactions such as hair loss, gastrointestinal complaints, itching or joint discomforts were common, but change of drug treatment was not necessary.

Discussion

Long-term results of ATD treatment in Graves’ hyperthyroidism are generally unsatisfactory. The overall relapse rate of 47.7% 2 years after discontinuation of ATD in the present, comparatively large study is in accordance with earlier European studies showing relapse rates between 30 and 60% 1–5 years after treatment (3, 10–13, 17).

Differences in genetic background and iodine intake may, at least in part, explain the lower relapse rates observed in other ethnic populations (18).

Both ATD and l-T4 are thought to influence intrathyroidal immune processes by inhibiting the plasma cells production of TRAb, decreasing the expression of autoantigens on the surface of thyrocytes or by other mechanisms (1, 8, 19). In addition, it has been suggested that suppression of TSH by exogenous l-T4 may put the thyroid gland in a resting state, thus reducing risk of relapse due to stimulation of thyrocytes by TSH (8).

Outcome of treatment may be related to ATD dose, duration of treatment and the use of l-T4 during or after withdrawal of ATD treatment. In the present
study, we have not found any differences between the study groups in TSH or TRAb levels at the end of ATD treatment (Table 2) nor in relapse rates 24 months after the cessation of ATD treatment (Fig. 2). Results of earlier studies have been conflicting. Some studies (6, 8), but not all (17, 20), have demonstrated lower TRAb levels after block-replace regimen compared with titration ATD treatment. Paschke et al. (21) were not able to find evidence of a dose-dependent immunosuppressive action of methimazole on the intensity of the intrathyroidal autoimmune processes. In a Norwegian study (7) with a short period (6 months) of ATD treatment, a beneficial effect on relapse rate of a high ATD dose (60 mg methimazole daily) combined with 1-T4 versus titrating methimazole doses could be demonstrated after 1 year, but not after 2 years of follow-up. After a mean follow-up time of 42 months Romaldini et al. (6) achieved a remission rate of 75% using block-replace regimen with 60 mg methimazole daily, which is equivalent to 100 mg carbimazole. High ATD doses may, however, increase the risk of serious adverse events (6, 7). Thus carbimazole doses higher than those used in our trial may increase the likelihood of long-term remission, but increase the risk of serious adverse effects (20).

Continuing 1-T4 monotherapy after completion of ATD treatment has been shown to improve relapse rates in one study (8). Using Cox’s multivariable proportional hazards model, the present study suggests that this may be true for patients who received block-replacement therapy (Table 3). The clinical importance of this finding is uncertain as the result was only marginally significant.

The need for reliable predictors of relapse in Graves’ disease is obvious, since these could help to identify patients needing prolonged ATD or ablative treatment. It is generally agreed that TRAb cause the hyperthyroidism of Graves’ disease (22). A meta-analysis based on studies including almost 1000 patients documented a significantly reduced risk of relapse in patients with undetectable TRAb levels at the end of ATD treatment (23). The predictive effect of smoking is controversial (24), and both increased and decreased thyroid function has been described in smokers (24, 25). In susceptible individuals, smoking may increase sympathetic activity (25) or may alter the structure of the TSH receptor and promote the development of autoimmune thyroid disease (26). Variations in iodine intake may modulate individual responsiveness (27). Positive TRAb at the end of treatment and current smoking indicated a 100% risk of relapse in a study including only 11 smokers (13), whereas Raber et al. (3) concluded that ‘smoking did not allow prediction of remission, irrespective of treatment group’. In our study, 58% of current smokers experienced a relapse compared with 40% of non-smokers, and the relapse rates were significantly higher in patients with persistently elevated TRab. Twenty-four months after withdrawal of ATD treatment as many as 80% of the current smokers who were TRAb-positive had relapsed. Current smoking and TRAb are both correlated to goiter size (28). Goiter size > 40 ml is associated with failure of medical treatment (29), and the effect on relapse rate by TRAb and smoking could therefore be explained by large goiter. However, using Cox’s proportional hazards regression analysis, we found that both TRAb and smoking acted to increase the risk of a relapse after adjusting for goiter size (Table 4).

No serious adverse events such as agranulocytosis or thrombocytopenia occurred, and a daily carbimazole dose of 30–40 mg appeared to be safe. Increase of alanine aminotransferase concentrations occurred in 3 of 22 patients taking propylthiouracil compared with none of the patients taking carbimazole. This suggests that propylthiouracil is more hepatotoxic than carbi- mazole, as reported earlier (30).

In conclusion, we found no difference in relapse rates in the patients with Graves’ disease treated with ATD alone or higher doses of ATD in combination with 1-T4 for 12 months. After medical treatment, smoking, large goiter size and TRAb at the end of ATD treatment are strong predictors of relapse in these patients.

Acknowledgements

We are indebted to Inger Bleskestad and John Cooper for the care of patients at their respective institutions, and John Cooper for critically reading the manuscript. This work was supported by grants from the Norwegian Research Council and Helse Vest.

References

2 Bjoro T, Holmen J, Krüger Ø, Midtbjell K, Hunstad K, Schreiner T et al. Prevalence of thyroid disease, thyroid dysfunction and

Table 4 Cox’s proportional hazards regression of prognostic covariates in the model for relapse-free follow-up time.

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>Standard error</th>
<th>Hazard ratio</th>
<th>95% CI</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRAb-positive at 12 months</td>
<td>1.14</td>
<td>0.22</td>
<td>3.14</td>
<td>2.04–4.84</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Current smoking</td>
<td>0.53</td>
<td>0.22</td>
<td>1.70</td>
<td>1.11–2.60</td>
<td>0.015</td>
</tr>
<tr>
<td>Goiter size</td>
<td>0.77</td>
<td>0.22</td>
<td>2.16</td>
<td>1.40–3.34</td>
<td>0.001</td>
</tr>
</tbody>
</table>


Received 30 April 2002
Accepted 8 July 2002