UNIVERSITY OF GOTHENBURG

THE SAHLGRENSKA ACADEMY

The efficacy of Rituximab in patients with thyroid associated ophthalmopathy not responding to steroid therapy

Degree project in Medicine

Elin Ekdahl

Programme in Medicine

Gothenburg, Sweden 2019

Main supervisor: Helena Filipsson Nyström, MD and Ass Prof
Department of Endocrinology
Sahlgrenska University Hospital

Co-supervisor: Sofia Manousou, MD and PhD student
Department of Medicine
Frölunda Hospital

Table of contents

List of abbreviations

ATD	Anti thyroid drug
CAS	Clinical activity score
CRF	Case report forms
DON	Dysthyroid optic neuropathy
DRG	Diagnosis-related group (DRG)
EUGOGO	European Group of Graves orbitopathy
fT3	Free triiodothyronine
fT4	Free thyroxine
GC	Glucocorticoid
GD	Graves' disease
iv	Intravenous
MTX	Methotrexate
NR-C	Non-responder patients in the control group that continued with iv GC
NR-RTX	Non-responder patients in the study that switched treatment to Rituximab
R-GC	Responder to treatment with iv GC
RAI	Radioactive iodine

RCT	Randomised clinical trial
RTX	Rituximab
SU/M	Mölndal hospital, a part of the Sahlgrenska University Hospital
t-RTX	The study population of the RTX study, excluding the control group
TAO	Thyroid associated ophthalmopathy
TRAb	TSH receptor antibodies
TSH	Thyroid stimulating hormone
VGR	The region of Västra Götaland
QoL	Quality of life

Abstract

Title: The efficacy of Rituximab in thyroid associated ophthalmopathy not responding to steroid therapy

Degree project in Medicine, Programme in Medicine, Sahlgrenska Academy at the University of Gothenburg, Sweden 2019.

Author: Elin Ekdahl, medical student.

Introduction: Thyroid associated ophthalmopathy (TAO) is an extra-thyroidal manifestation of Graves' disease. Symptoms of TAO are for example diplopia, pain, swelling, erythema, exophthalmos and blindness. Clinical activity score (CAS) measures the degree of inflammation (scale 1-10). In clinically active TAO, CAS is ≥ 4. The primary treatment in clinically active TAO, intravenous glucocorticoids (iv GC), has improved outcome in 75-80% of the patients (responders). A new available treatment for TAO is Rituximab (RTX).

Aim: To compare the outcome (in CAS and quality of life) of RTX and iv GC in patients with active TAO who do not respond to treatment with iv GC.

Method: This is a non-randomized prospective study including 26 patients. After four weeks of treatment with iv GC, patients were separated according to their response to iv GC after 4 weeks in responders – those that improved ≥2 in CAS (R-GC) and continued with iv GC for another 8 weeks, and non-responders – those that improved <2 in CAS and were switched to RTX (NR-RTX). A control group consisting of non-responders that continued with iv GC but were not included in the RTX study was created retrospectively through journal research (NR-C). Data were collected from medical journals and case report forms and thereafter analysed in SPSS.

Results: At 18 weeks, neither did the NR-RTX and NR-C groups differ in CAS, nor did R-GC and NR-RTX. Both groups of non-responders contained more males than R-GC. The responders had improved their quality of life in some aspects at week 18, as opposed to the non-responders.

Conclusion: This study shows no benefit of treatment with RTX compared to iv GC for patients with moderate-severe TAO who do not respond to treatment with iv GC. Further studies of the same groups are necessary in order to assess other aspects of the treatment.

Key words: Rituximab, Thyroid associated ophthalmopathy, Glucocorticoids, Clinical Activity Score, Quality of life

Introduction

Background

Graves' disease (GD) is of autoimmune origin and is the most common cause of hyperthyroidism. It has a reported incidence in Sweden of 2100 patients / year (1). The incidence of GD is higher in women than in men and it peaks in the middle age of the life span (2). Both genetic, epigenetic and environmental factors are important for its development (3). Common symptoms of hyperthyroidism are weight loss, fatigue, heat intolerance, palpitations, tremor, anxiety, hair loss and dyspnoea (3). In GD, these symptoms occur as thyroid stimulating hormone (TSH) receptor antibodies (TRAb) stimulates the thyroid with the result of hyperactivity, which leads to hyperthyroidism and sometimes enlargement of the thyroid gland (4).

In order to diagnose GD, laboratory test can be used. It is important to analyse TSH and free thyroxine (fT4) (4). The latter is a prohormone produced by the thyroid. The combination of TSH and fT4 is the most sensitive measure for thyroid dysfunction and free triiodothyronine (fT3), the active thyroid hormone, better indicates the degree of hyperthyroidism. The reason to hyperthyroidism is, however, detected by the presence of TRAb (3) and GD is hereby confirmed. If TRAb is negative a scintigraphy is necessary, as there are other causes of hyperthyroidism, for example toxic adenoma or subacute thyroiditis (5).

There are different treatments available for GD. The first choice of treatment is anti-thyroid drugs (ATD), which inhibit the thyroid gland and reduce the production of thyroid hormones. ATD is usually used for at least 12 months and thereafter until the level of TRAb has normalized (4). After ATD treatment, there is a risk of relapse of the disease and it is most

likely to occur within 0-12 months after finished treatment (6). Therefore, it is recommended that these patients have continuous and regular follow up visits the year after treatment (5).

A second treatment is radioactive iodine (RAI), which can be used for patients with comorbidities, patients who have had side effects of the ATD or who have had relapse after treatment with ATD (4, 5). Pregnancy and breastfeeding are contraindicated, and for both female and male patients it is recommended to avoid conception until at least 6 months after the treatment (5).

Surgery is also an option as treatment for GD (4). Total thyroidectomy is the method of choice, as the risk of relapse is lower, and the risk of complications is the same as when performing a subtotal thyroidectomy (5). Surgery is indicated for example for women who are planning pregnancies within half a year, in cases of enlargement of the thyroid gland or in moderate or severe thyroid associated ophthalmopathy (TAO) (4).

After treatment with RAI or surgery, many patients need replacement with oral levothyroxine (3) in order to remain euthyroid (5) since the treatments lead to hypothyroidism with elevations of TSH concentration (7).

Thyroid associated ophthalmopathy

GD has several extra-thyroidal manifestations, among which the most common is thyroid associated ophthalmopathy (TAO), also named Graves' orbitopathy. TAO affects up to 60% of the patients' with GD (8), is clinically relevant in 25-40% (9) and sight threatening in 3-5% (10). Symptoms of TAO are for example diplopia, pain, exophthalmos, proptosis, epiphora, conjunctival injection and impaired vision due to dysthyroid optic neuropathy (DON) (3, 10).

Like GD, TAO is more common in women than men, but its severe forms have higher prevalence in men (11). TAO most often occurs between the ages of 30-50 and most often within 18 months after being diagnosed with GD (11). An important risk factor is cigarette smoking (12). TAO is triggered by both genetic as well as environmental factors (11). TAO is triggered by both genetic as well as environmental factors (11). TAO is also of autoimmune origin, but in opposite to GD, it involves both the cell-mediated and humoral immune system. The affected patients obtain inflamed connective tissue, as well as inflammation in the extraocular muscles and increased fat production within the orbits (13). TAO has a major impact on the patients' well-being and it decreases the patients' quality of life (QoL) (14).

In order to measure the inflammatory activity of TAO, the European Group on Graves' Orbitopathy (EUGOGO) has established the Clinical Activity Score (CAS) (15). CAS contains ten different inflammatory variables (Table 1) (11, 15, 16). At the first visit, the patient can at most obtain 7 points (parameter 1-7), so called 7-point scoring system of CAS. At follow up visits, the 10-point scoring system of CAS is used (parameter 1-10) (11), which is used for continuous assessment of the patient, since the latter three scoring points compares the eye status from a previous visit (12, 15). TAO is classified as active when the patient obtains three or more points

Table 1. Clinical Activity Score*. The inflammatory activity of thyroid associated ophthalmopathy is measured according to this scoring system. Each parameter has a score of 0-1 points. The first seven scores are used at the first visit, thereafter the 10-score system is used. If CAS-7 ≥ 3 points or CAS-10 ≥ 4, the disease is seen as clinically active and treatment is indicated.*

1. Bulbar or retrobulbar pain
2. Pain with eye movement
3. Periocular swelling
4. Lid erythema
5. Chemosis
6. Bulbary injection
7. Inflammation of caruncle
8. Increased proptosis
9. Increased vision loss
10. Decrease of eye motility

according to the 7-point score and four or more points according to the 10-point score (11).

Besides inflammation, the patients with TAO also are evaluated regarding severity of the disease. This is done with the severity classification according to the Bartalena *et al* (16). The eleven items of the severity scoring system are presented in table 2. With this system, sight threatening TAO can be identified (16).

Table 2. Severity. *This scoring system is used in order to evaluate how severe the patients thyroid associated ophthalmopathy (TAO) is. With this system, sight threatening TAO can be identified. Each item is graded from 0-3 points. Thus, maximum score is 33 points.*

1. Lid retraction
2. Proptosis
3. Eyelid swelling
4. Eyelid erythema
5. Chemosis
6. Bulbar injection
7. Inflammation caruncle
8. Subjective diplopia
9. Eye motility
10. Corneal impact
11. Opticus impact

Treatment of thyroid associated ophthalmopathy

As mentioned earlier, if CAS-7 ≥ 3 points or CAS-10 ≥ 4, the disease is seen as clinically active and is assessed as moderate-severe TAO, which is an indication for anti-inflammatory treatment. The primary treatment of moderate-severe TAO is intravenous (iv) administered glucocorticoids (GC) (12). The most common scheme that is recommended in international guidelines (2) is 500 mg methylprednisolone once per week for six weeks followed by 250 mg / week for another six weeks.

Furthermore, if the disease progresses into severe TAO and becomes sight threatening with DON, the treatment regimen is changed to a more intensive iv GC scheme (2). If DON, the recommended treatment is 500-1000 mg of methylprednisolone for three days / week for two weeks. If the patient's visual acuity is still impaired, urgent orbital decompression is necessary to restore vision. However, if the patient responds to the treatment with high dose of iv GC, the patient should continue its treatment with GC iv or *per oral* for its TAO (2).

Only 75-80% of the patients with moderate-severe TAO respond to treatment with iv GC by reducing CAS at least two points (9), but the response rate is higher than with *per oral* GC treatment (17). Also, the frequency of side effects and adverse events with iv GC treatment is lower than when per oral treatment is used, although there are some more rare severe side-effects in connection to iv GC. Still many patients obtain severe side effects of treatment with iv GC, such as weight gain, insomnia, palpitations (12, 17). These symptoms affect the patients' QoL, beyond the consequences on QoL by TAO itself (17).

A secondary option of treatment for moderate-severe TAO is rituximab (RTX) (12). RTX, a chimeric humanized monoclonal antibody, targets the CD20 antigen, which is a protein located on B lymphocytes (9, 13). Treatment with RTX thus leads to elimination of the B lymphocytes that are expressing the CD20 antigen, but it does not affect the humoral immune system otherwise (18). The effect of RTX lasts for approximately four to six months (18). In previous studies, the dose of RTX when treating moderate-severe TAO is 1000 mg administrated as iv infusion twice with two weeks in between (9, 18, 19). This dose is also used in patients with rheumatic diseases and it is most often combined with Methotrexate in order to avoid production of antibodies that target RTX (20).

Observational studies indicate that RTX may be used successfully as treatment against moderate-severe TAO (18, 19) and a randomised clinical trial (RCT) is coherent with these observations (9). RTX is reported to cause fewer adverse events than iv GC (2, 9). The most frequent adverse event caused by RTX is infusion-related reactions, but it may cause more severe side effects as well, for instance leukoencephalopathy (2). It has also been reported that RTX may improve QoL, when used as treatment against TAO (2). However, another RCT showed no benefit of RTX compared to NaCl (21). In a meta-analysis (22), the conclusion is

benefit from RTX, and in the last international guideline from 2016 (2), RTX is recommended as a second line treatment, in cases of no response to iv GC.

Although recommended, there is no study evaluating the efficacy of RTX in patients with TAO that are non-responsive to steroids. The hypothesis of this study is that patients who have not responded to the first 4 weeks of treatment with iv GC, have better anti-inflammatory effect, measured by CAS, with RTX than by continued treatment with iv GC.

Aim

The primary aim of this study is to evaluate the effect of RTX in TAO patients who are non-responsive to treatment with iv GC (NR-RTX), with patients (controls) that continued with iv GC even though they did not respond (NR-C). A second aim is to compare the NR-RTX with the steroid responsive ones (R-GC) as well as the two groups of non-responders (NR-RTX and NR-C), regarding background variables (gender, age, smoking status, duration of GD, duration of eye symptoms, thyroid treatment and lab) and efficacy variables (CAS, CAS SU/M, severity, proptosis and subjective diplopia). Lastly, a third aim is to evaluate and compare the QoL between the R-GC and the NR-RTX in the study at given time points.

Research questions

Does RTX lower the CAS of the non-steroid responsive TAO-patients, in comparison to NR-C, that continued with iv GC? Do the background variables differ between the groups of NR-RTX and NR-C, or between R-GC and NR-RTX? And lastly, does the QoL differ between the R-GC and the NR-RTX at given measuring points?

Material and Methods

Study design

This master thesis is a part of the RTX-study, which is a study that began in 2011 at Sahlgrenska University Hospital, Gothenburg, Sweden. The RTX-study is a non-randomized prospective study that covers patients with moderate-severe TAO with indication for treatment with iv GC. For this study, a control group (NR-C) was created. The plan of analysis of the RTX study is presented in figure 1. In order to limit this master thesis, data on the follow-up period after week 18 week were not analysed nor analysis of side effects.

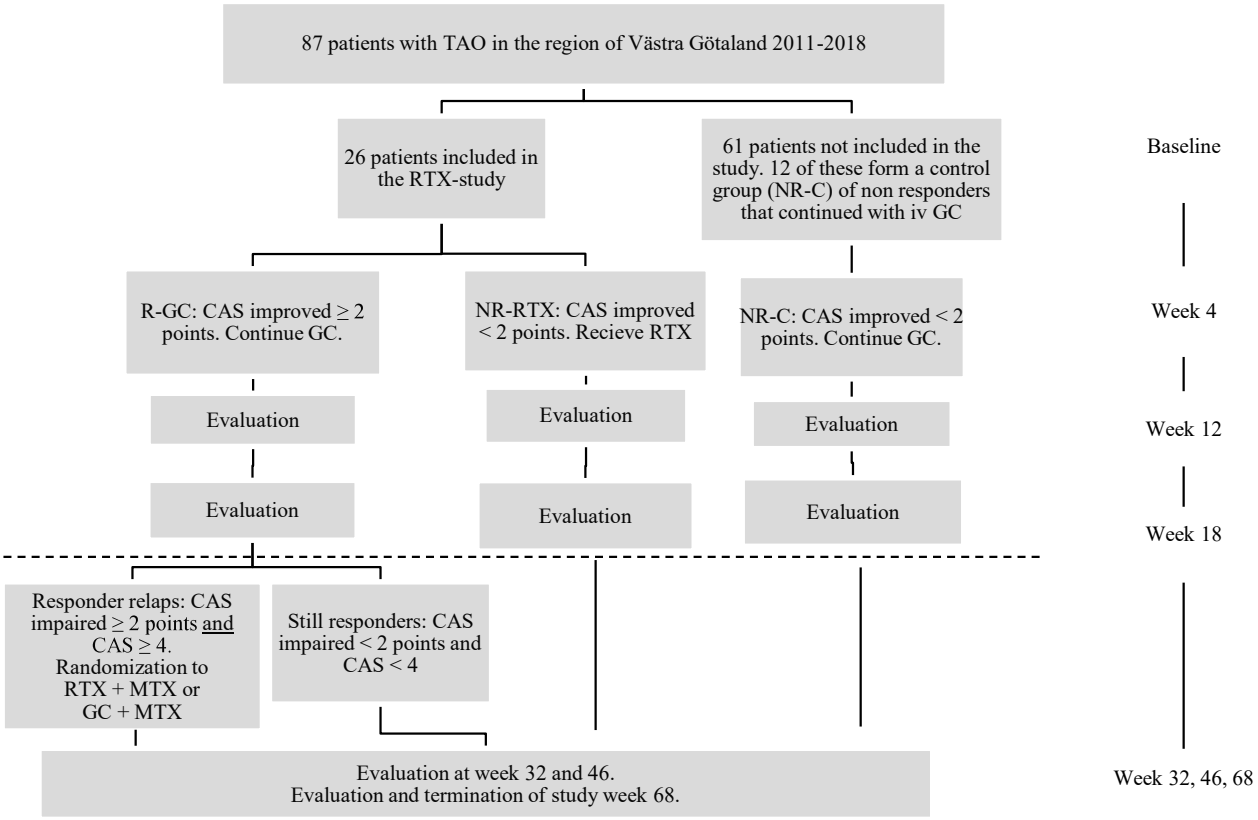

Figure 1. The study design and plan of analysis of the RTX study. Primary endpoint is at week 18, before extracting the relapse at responders from the responder group. The dotted line represents the limitation of this master thesis. Abbreviations: CAS, clinical activity score; GC, glucocorticoids; iv, intravenous; MTX, Methotrexate; NR-C, control group consisting of non-responders that continued with iv GC; NR-RTX, non-responders to treatment with iv GC that obtained RTX; R-GC, responders to treatment with iv GC; RTX, Rituximab; TAO, thyroid associated ophthalmopathy.

Patients with TAO were evaluated for iv GC treatment at Department of Ophthalmology at Sahlgrenska University Hospital/ Mölndal's and were asked for participation if eligible. Before inclusion, the following was tested/analysed: hepatitis B and C, active and latent tuberculosis, and pregnancy. Magnetic resonance imaging (MRI) of the orbit was also performed at inclusion, and again 30 weeks after inclusion (these data are also outside this master thesis).

Variables that were assessed at each visit at the ophthalmologist were CAS, CAS SU/M, severity, proptosis, photographs, ductions vertical and horizontal (in this master thesis CAS, CAS SU/M, severity and proptosis are included). At the endocrinologist, ThyPRO self-assessment questionnaire was filled in. For each visit, the following biochemical tests were collected: fT4, fT3, TSH, TRAb, liver function tests (ALP, ALAT, ASAT, bilirubin) and fasting blood glucose. A metabolic evaluation was also made at every visit, measuring blood pressure, waist, weight and length. In this master thesis, only the thyroid biochemical tests at baseline and week 4 have been analysed.

All patients in the RTX-study started the 12-week scheme of iv GC at baseline (2). After four weeks of treatment, an evaluation of the treatment outcome was made. The primary evaluation form was CAS. If the patients had improved their CAS score by ≥ 2 points, they were classified as responders to iv GC (R-GC) and continued their 12-week iv GC scheme. The remainder of the patients were classified as non-responders to the treatment with iv GC and were switched to 1000 mg iv RTX in week 5 and 7, combined with a weekly oral dose of 12.5 – 15 mg MTX (NR-RTX). The RTX was administered at the department of Rheumatology at Sahlgrenska University Hospital. All included patients had thereafter follow-up visits at week 12, 18, 32, 46 and 68 weeks after baseline.

At week 18 and at all follow-up timepoints thereafter, patients in R-GC group were evaluated for possible relapse. If a patient in R-GC had a CAS ≥ 4 points and CAS had worsened by ≥ 2 points compared to the value measured in the prior evaluation time-point, the patient was considered having a relapse. Patients with relapse were randomized to either treatment with RTX and MTX or with *per oral* GC and MTX. However, this master thesis is limited to the analyses up and until week 18. Week 18 is also the time-point when the primary outcome is set, due to the split of the initial R-GC group thereafter, leading to a weaker statistical power.

If patients' TAO progressed to severe TAO with DON during the study, the course of treatment was changed to high dose iv GC, consisting of 1 g methylprednisolone every other day in two weeks (with a total of 6 g methylprednisolone). Visual acuity was not restored, they were referred for orbital decompression. After surgery, they obtained treatment with *per oral* GC that slowly was reduced. All patients included in the study were assessed according to study protocol, but if DON developed, the affected patients were treated outside the study protocol in order to follow DON treatment regimen. The patients with DON still had the same evaluation time points as the other patients in the study, but they were excluded from the comparisons between groups after the occurrence of DON.

From the group of patients with moderate – severe TAO with indication for treatment with iv GC that did not meet inclusion/exclusion criteria for the RTX-study or declined participation, a control group (NR-C) of non-responders was created. As they were not included in the RTX study, they continued with their 12 weeks scheme of iv GC treatment.

Power calculation

No proper power calculation could be done at the start of the RTX study due to lack of variability data from any similar study of RTX in non-steroid responders. Several attempts were made to secure a comparable group of non-responders that had received the full iv GC scheme through international co-operations, but without success. This was one main reason for extracting a control group from the medical records. A post-hoc power calculation was done after 26 included patients, using preliminary data from the control group and based on calculation, the RTX study was closed for inclusion.

Study population

In the years 2011 – 2018, there were 87 patients with clinically active TAO in the region of Västra Götaland (VGR), thereof 26 patients were included to the RTX-study at Sahlgrenska University Hospital, Gothenburg, Sweden. Inclusion criteria for the RTX study were: man or woman between 18-70 years old, TAO with 7-point score system of CAS $\geq$ 4 (less than 3 months old), euthyroid for at least 6 weeks.

Exclusion criteria for the RTX study were: DON, ulcerative corneal involvement, previous treatment with steroids for TAO (except for prophylaxis for TAO in connection with RAI), previous treatment with RTX, positive Hepatitis B or C serology, obtained a living vaccine within 4 weeks prior RTX+MTX to evaluation, history of recurrent significant infection or history of recurrent bacterial infections, assessment by the investigator that the patient may not attend to the protocol, current pregnancy or lactation, significant cardiac disease, including significant or uncontrolled arrhythmia, previous active tuberculosis, bone marrow depression with leukopenia, thrombocytopenia or significant anaemia, allergy to the active

substance or any other substance in the medications or murine proteins, active, severe infections (such as tuberculosis, sepsis or opportunistic infections), patients with severe immunosuppression, severe cardiac failure or severe uncontrolled heart disease.

In VGR 30% of TAO patients with indication for iv GC were eligible for the RTX study, but 61 patients were not included. The reasons for non-inclusion were: attending at other hospital of VGR than Sahlgrenska University Hospital for treatment of TAO (n=5), DON at baseline (n=7) or any other exclusion criteria (n=39), denied participation (n=3), or study inclusion that was impossible due practical reasons (the personnel or the patients were on vacation, the ophthalmologist forgot to ask about study participation etcetera) (n=7). The 12 participants of the NR-C were collected from this group of patients. To be a part of the NR-C, the following inclusion criteria had to be fulfilled: moderate-severe TAO with CAS ≥4 and no DON at baseline, less than 3 months with CAS ≥4, classification as non-responders at week 4 of treatment with iv GC (ΔCAS < 2 from treatment start), and that they had the intention to follow the 12 week schedule with iv GC.

The patients in the NR-C remained a part of the control group even if they did not fulfil their 12-week protocol of iv GC treatment because of adverse events. If they had changed treatment during their 12-week scheme, they were excluded after the treatment change.

The TAO patients in VGR not included in the RTX-study were identified using the diagnosis-related group (DRG) registry according to the 10th revision (ICD-10). In DRG, a combination of different codes was used (TAO [H06.2], Graves' disease [E05.0] and iv infusion [DT016]). In order to select which of the non-included patients that could be a part of the group of comparison, their medical records were evaluated.

Data collection

The data for this master thesis were collected from the patients' medical journals and in case of RTX-study patients also from case report forms (CRF) that were filled in during the study. The background variables collected for this thesis were age, gender, current smoking status, time since debut of GD, time since debut of eye symptoms and current lab values (fT3, fT4, TRAb, TSH). Efficacy variables collected for the RTX-study patients were CAS, CAS SU/M, severity, proptosis, change in subjective diplopia and results of the QoL form ThyPRO questionnaire (23).Of the efficacy variables, only CAS and proptosis were available for the control group, as these were the only variables that were registered consecutively in clinical practice.

Efficacy variables

Both CAS and severity score was used, as described earlier in the background (table 1 and 2). In the study, both patients' eyes were evaluated and the score of and any improvement of the worst eye was used.

CAS SU/M is an extended version of CAS (table 10 in Appendices), created by the Department of Ophthalmology at Sahlgrenska University Hospital/Mölndal (SU/M). It is an extended score, where each item may score from one to maximum three points within each item of CAS (see table 1 for the items of CAS) with a total score of 15 or 18 if the 7-point or 10-point CAS is used, respectively.

The proptosis was measured in both eyes using the Hertel exophtalmometer. The proptosis was defined as improved if the Hertel score had decreased by ≥ 2 points in one of the eyes and not increased by > 2 points in the other eye between two consecutive evaluation time-points.

The ThyPRO questionnaire is a thyroid specific QoL self-assessment questionnaire (23). It is one questionnaire that measures QoL within 13 different dimensions: goitre symptoms, hyperthyroid symptoms (constitutes two dimensions), eye symptoms, tiredness, cognitive function, anxiety, depressivity, emotional susceptibility, impairment of social, daily and sex life and lastly cosmetic complaints. ThyPRO consists of totally 84 questions. The scores for each dimension can be presented as raw or normalised score (raw score/max score x 100) that ranges from 0 to 100. A high normalised score indicates decreasing QoL (i.e. more symptoms or greater impact of disease). Separated from the 13 dimensions, ThyPRO has one last 85th question that regards QoL generally. This question was not used in this study as it was not filled in by the majority of patients.

Statistical methods

In this study, the programme SPSS (Statistical Package for the Social Sciences), version 25 (IBM Corp., Armonk, New York) was used. Normality was tested using Sharpio-Wilks test. If the variables were normally distributed, a *t*-test was used for independent samples and paired *t*-test for related samples. If the variables were not normally distributed, Mann-Whitney *U* test was performed for independent samples and Wilcoxon signed-rank test for related samples. Lastly, in order to compare categorical variables, Chi Square Test was used. If the group of categorical variables was small, Fishers Exact test was used instead. The Bonferoni correction has not been used. All statistical significance was set at alpha level of 0.05.

Ethics

This study, as well as the total RTX study, was conducted according to the Helsinki Declaration. The patients that are included in the RTX study have given their written consent for data collection and the study is ethically approved by the Ethical Review Board in Gothenburg (DNR 276-11, date of approval 2011-04-26). The Ethical Review Board in Gothenburg Sweden has also approved (DNR T392-18, date of approval 201805-02) the review of files for the TAO patients in VGR that was not included in the study. The fact that the control patients not having individually approved the participation may be an ethical dilemma, but the benefit for this research and the low risk for harm for these patients laid ground for the permission from the Ethics Committee. Furthermore, the care unit manager approved that variables from the patients in the RTX study and from control patients could be registered for this research. The RTX study is also approved by the Medical Product Agency.

Results

In this evaluation, 26 patients from RTX-study (t-RTX) and 12 controls (NR-C) with moderate-severe TAO were included (Figure 1). Among the t-RTX, there was one drop out at baseline and two patients that had developed DON at their visit week 4. Except for these patients, all patients followed study protocol until primary end at week 18. At week 18, another one of the NR-RTX had developed DON and changed course of treatment after the evaluation at week 18. At week 18, a total of 38.5% of the R-GC had a relapse in their disease (Figure 2), among which one that also developed DON. The death at week 18 among the R-GC was due to suicide. All patients in t-RTX followed study protocol and obtained treatment according to international guidelines (2).

As seen in figure 2, one patient was excluded from NR-C after 4 weeks because of change of treatment to RTX, since the patient had no effect of iv GC treatment. Two patients in the control group did not complete the 12-week protocol with iv GC because of adverse events (elevated liver function tests), but they remained a part of the control group until week 18. After week 18, one of these obtained RTX as well. One patient among the NR-C developed DON at week 12 and changed treatment regimen thereafter.

A total of five patients (among which four attended to an evaluation at week 18) in the NR-C obtained a scheme of *per oral* glucocorticoids that was tapered down. None of the participants in NR-C had a relapse in their disease at week 18. The patients that obtained *per oral* steroids after the 12-week scheme did not differ in the CAS compared to the other in the NR-C at week 18.

The frequency of DON is 11.5% at baseline, among the patients in VGR with moderate-severe TAO between 2011-2018, that were not a part of either the t-RTX or the NR-C. After 4 weeks, 15.3% of t-RTX had developed DON. At week 18, 20% of the NR-RTX, 7.7% of the R-GC and 9.1% of the NR-C had developed DON.

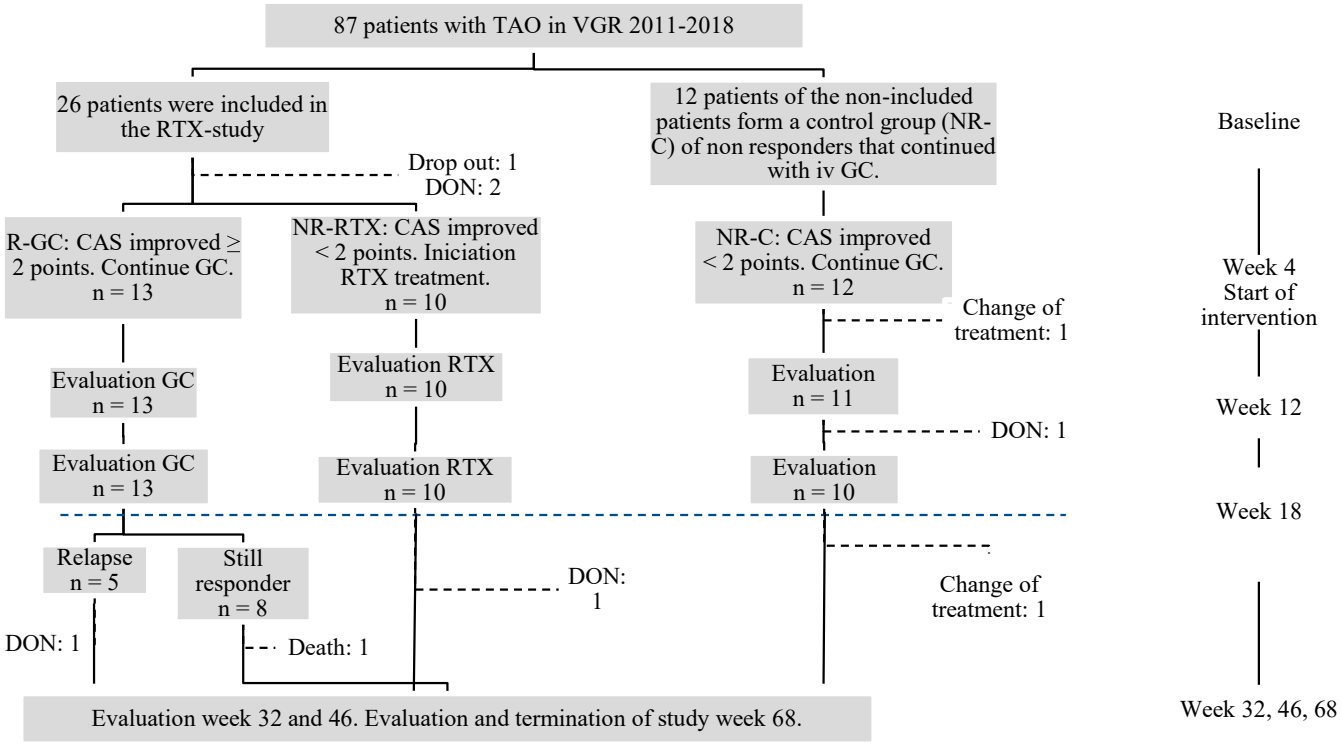

Figure 2. Development of the study population throughout the study. This flow chart displays how many patients that that was a part of the different groups at the different measuring points in the study. The blue dotted line represents the limit of this master thesis.
Abbreviations: CAS, clinical activity score; GC, glucocorticoids; iv, intravenous; MTX, Methotrexate; NR-C, control group consisting of non-responders that continued with iv GC; NR-RTX, non-responders to treatment with iv GC that obtained RTX; R-GC, responders to treatment with iv GC; RTX, Rituximab; TAO, thyroid associated ophthalmopathy.

Study population characteristics

Background characteristics for baseline and week 4, before the start of RTX intervention, are presented in tables 3 and 4, respectively. Women were equally represented between the total RTX-study population (t-RTX) and NR-C, but there were more men among the non-responders. NR-RTX did not differ by sex compared to NR-C. There were no differences regarding age, smoking status, other previous or current treatment of GD, duration of GD or TAO symptoms, nor any differences in laboratory measurements. No differences were observed regarding CAS, CAS-SU/M, severity or diplopia at baseline except for a higher CAS in NR-C compared with NR-RTX.

Table 3. Background characteristics at baseline for the total RTX study population (t-RTX) and the control group (NR-C). The t-RTX is presented both as whole group and separated by the response at 4-weeks: responder to (R-GC) and non-responder (NR-RTX). Variables are presented as n (%), mean ± standard deviation or median (interquartile range). Statistical significance is set at alpha level of 0.05.

	t-RTX (n=26)	**t-RTX separated by response at 4-weeks** n=25[a]		**NR-C** (n=12)	**Test between groups, p-value**		
		R-GC n=13	**NR-RTX** n=12		**t-RTX NR-C**	**NR-RTX R-GC**	**NR-RTX NR-C**
Female sex	19 (73.1)	12 (92.3)	6 (50.0)	7 (58.3)	0.46[d]	0.03*[d]	1.0 [d]
Age – year	54.2 ± 8.2	54.6 ± 6.9	52.8 ± 9.1	57.5 ± 12.0	0.33[e]	0.6[e]	0.3[e]
Smoking status							
Smoker	9 (34.6)	2 (15.4)	7 (58.3)	5 (41.7)			
Previous smoker	13 (50)	9 (69.2)	4 (33.3)	2 (16.7)	0.2[f]	0.08[f]	0.4[f]
Non-smoker	3 (11.5)	2 (15.4)	1 (8.3)	3 (25)			
Unknown	1 (3.8) [b]			2 (16.7)			
Current treatment thyroid disease							
Levothyroxine	25 (96.2)	13 (100)	12 (100)	12 (100)	[c]	[c]	[c]
ATD	10 (38.5)	5 (38.5)	5 (41.7)	9 (75)	0.08[d]	1.0[d]	0.2[d]
Unknown	1 (3.8) [b]			-			
Previous treatment thyroid disease							
Thyroidectomy	7 (26.9)	2 (15.4)	5 (41.7)	2 (16.7)	0.7[d]	0.20[d]	0.4[d]
RAI	7 (26.9)	4 (30.8)	3 (25.0)	6 (50)	0.3[d]	1.0[d]	0.4[d]
ATD	6 (23.1)	2 (15.4)	4 (33.3)	1 (8.3)	0.4[d]	0.4[d]	0.3[d]
Unknown	1 (3.8) [b]			-	-		
Graves' disease – year	1.0 (1.0, 2.0) *n=25*	2.0 (0.4, 4.0)	1.0 (1.0, 2.0)	1.6 (0.5, 2.8)	1.0[g]	1.0[g]	0.8[g]
Duration eye symptoms – months	6.0 (3.5, 13.5) *n=25*	6.0 (4.0, 13.0)	5.5 (2.0, 18.8)	4.50 (3, 17.5)	0.5[g]	0.4[g]	1.0[g]
fT4	17.7 ± 3.8 *n = 24*	17.7 ± 2.9	17.7 ± 4.7 *n=11*	18.2 ± 3.3	0.7[e]	1.0[e]	0.8[e]
fT3	4.4 ± 0.6 *n = 22*	4.5 ± 0.6 *n=12*	4.3 ± 0.4 *n=10*	3.9 ± 1.2 *n = 6*	0.4[e]	0.3[e]	0.5[e]
TSH	0.7 (0.2, 2.0) *n=23*	0.3 (0.2, 1.0) *n=12*	1.6 (0.07, 2.6) *n=11*	0.3 (0.09, 2.0)	0.3[g]	0.3[g]	0.3[g]
TRAb	13.0 (6.4, 31.0) *n=23*	13.3 ± 9.6 *n=12*	28.2 ± 22.4 28.6 (7.8, 41.0) *n=11*	10.9 (4.7, 45.3) *n=11*	1.0[g]	0.06[e]	0.7[g]
CAS	4.0 (4.0, 5.0) *n=25*	4.0 (4.0, 6.0)	4.0 (4.0, 5.0)	5.0 (5.0, 5.0)	0.2[g]	0.6[g]	0.03*[g]
CAS SU/M	10.0 (8.0, 10.5) *n=25*	9.7 ± 3.1	9.8 ± 1.3			0.9[e]	
Severity	11 (8.50, 14.00) *n=25*	11.9 ± 3.5	10.8 ± 3.6			0.5[e]	
Subjective diplopia							
No diplopia	5 (19.2)	3 (23.1)	2 (16.7)				
Intermittent	6 (23.1)	5 (38.5)	1 (8.3)				
Can be corrected with prism	8 (30.8)	2 15.4)	6 (0.5)			0.2[f]	
Constant	4 (15.4)	2 (15.4)	2 (16.7)				
Missing data	3 (13.0)	1 (7.7)	1 (8.3)				

** Significant p-value*

[a] One of the patients in t-RTX dropped out of the study, thus there are only 25 patients when t-RTX are split

[b] Drop out before completed baseline.

[c] P-value is missing due to 100% use of levothyroxine

[d] Fishers Exact test was used with a set alpha level of 0.05.

[e] T-test was used with a set alpha level of 0.05.

[f] Chi Square test was used with a set alpha level of 0.05.

[g] Mann Whitney U test was used with a set alpha level of 0.05.

Abbreviations: CAS, clinical activity score; NR-C, control group consisting of non-responders to treatment with glucocorticoids but continued with the treatment; NR-RTX, non-responders to treatment with glucocorticoids that changed treatment to RTX; GC, responders to treatment with glucocorticoids, RTX, Rituximab; t-RTX, total study group of the RTX study; SU/M, Mölndal hospital.

Table 4. Background and efficacy variables at week 4 for the total RTX study population (t-RTX) and the control group (NR-C). The t-RTX is presented separated by the response at 4-weeks: responder to (R-GC) and non-responder (NR-RTX). Variables are presented as n (%), mean ± standard deviation or median (interquartile range). Statistical significance is set at alpha level of 0.05.

Variables	t-RTX separated by response at 4-weeks		NR-C *n = 12*	Test between groups, p-value	
	R-GC *n = 13*	NR-RTX *n = 10*		NR-RTX R-GC	NR-RTX NR-C
Female sex	12 (92.3)	5 (50.0)	7 (58.3)	0.05[c]	1.0[c]
Age – year	54.6 ± 7.0	52.3 ± 9.5	57.5 ± 12.0	0.5[d]	0.3[d]
Smoking status					
Smoker	2 (15.4)	6 (60.0)	5 (41.7)		
Previous smoker	9 (69.2)	3 (30.0)	2 (16.7)	0.08[e]	0.5[e]
Non-smoker	2 (15.4)	1 (10)	3 (25)		
Unknown	-	-	2 (16.7)		
Current treatment thyroid disease					
Levothyroxine	13 (100)	10 (100)	12 (100)	[a]	[a]
ATD	5 (38.5)	5 (50)	9 (75.0)	0.69[c]	0.4[c]
Previous treatment thyroid disease					
Thyroidectomy	2 (15.4)	3 (30.0)	2 (16.7)	0.6[c]	0.6[c]
RAI	4 (30.8)	3 (30.0)	6 (50.0)	1.0[c]	0.4[c]
ATD	2 (15.4)	3 (30.0)	1 (8.3)	0.6[c]	0.3[c]
Graves' disease – year	2.0 (0.38, 4.0)	1.2 (1.0, 2.0)	1.6 (0.5, 2.8)	1.0[f]	1.0[f]
Duration eye symptoms – months	6.0 (4.0, 13.0)	5.5 (2.0, 22.5)	4.5 (3.0, 17.5)	0.5[f]	0.9[f]
fT4	18.4 ± 2.6	19.3 ± 2.4	18.3 ± 3.1	0.4 [f]	0.4 [f]
fT3	4.5 (4.0, 4.8) *n=10*	4.4 (2.0, 5.0) *n=8*	4.9 (4.7, 5.6) *n=4*	1.0	0.1
TSH	0.5 (0.2, 1.3)	1.1 (0.8, 2.5)	0.5 (0.2, 2.4)	0.07 [f]	0.3 [f]
TRAb	7.7 (2.6, 12.5)	23.5 (3.6, 41.0)	3.8 (1.7, 10.0) *n=5*	0.1 [f]	0.09 [f]
CAS	2.7 ± 1.4	4.4 ± 1.2	4.2 ± 0.9	0.005* [f]	0.6 [f]
ΔCAS from baseline [b]	2.0 (1.5, 3.0)	0.0 (-1.0, 1.0) -0.1 ± 1.0	0.75 ± 0.9	<0.001* [f]	0.045* [d]
CAS SU/M	5.2 ± 2.0	8.3 ± 2.3		0.002* [d]	
ΔCAS SU/M from baseline [b]	4.5 ± 2.8	1.3 ± 1.6		0.004* [d]	
Severity	7.0 (5.0, 8.5)	9.5 (7.8, 12.3)		0.04* [f]	
ΔSeverity from baseline [b]	4.5 ± 1.8	0.2 ± 2.0		<0.001* [d]	
Subjective diplopia					
No diplopia	5 (38.5)	2 (20.0)			
Intermittent	4 (30.8)	3 (30.0)		0.7[e]	
Can be corrected with prism	2 (15.4)	3 (30.0)			
Constant	2 (15.4)	2 (20.0)			

** Significant p-value*
[a] P-value is missing due to 100% use of levothyroxine
[b] Δ of variable = (value at baseline – value at 4 weeks). As high values in these variables represent worse status, positive Δ of variable means decrease in the score through the study and therefore improvement in the patients' TAO or QoL.
[c] Fishers Exact test was used with a set alpha level of 0.05
[d] T test was used with a set alpha level of 0.05
[e] Chi Square test was used with a set alpha level of 0.05
[f] Mann Whitney U test was used with a set alpha level of 0.05
Abbreviations: CAS, clinical activity score; NR-C, control group consisting of non-responders to treatment with glucocorticoids but continued with the treatment; NR-RTX, non-responders to treatment with glucocorticoids that changed

treatment to RTX; GC, responders to treatment with glucocorticoids, RTX, Rituximab; t-RTX, total study group of the RTX study; SU/M, Mölndal hospital.

Efficacy variables

At baseline, there were no differences regarding efficacy variables between NR-RTX and R-GC, nor between t-RTX and NR-C. The NR-C presented with higher CAS compared to the NR-RTX (table 3).

In the R-GC, the CAS improved by iv GC, whereas it remained the same in the NR-RTX. R-GC and NR-RTX differed also at 4 weeks regarding CAS-SU/M and severity, but these differences were not evident when diplopia was evaluated. The CAS levels of the NR-RTX and NR-C were similar at week 4 (Table 4). Though, since NR-C had a higher CAS at baseline compared to NR-RTX (Table 3), the ΔCAS from baseline to week 4 was significantly larger for the NR-C compared to NR-RTX (Table 4).

After 12 weeks, R-GC still had a significantly lower CAS and CAS SU/M than NR-RTX, but they had not improved CAS since week 4 (table 5). NR-RTX and NR-C are presented without difference in CAS. ΔCAS from week 4 shows an increase in CAS score for R-GC, comparing to week 4, whereas it shows a decrease in CAS for both NR-RTX and NR-C. No differences in severity, subjective diplopia or proptosis can be shown.

In table 6, the outcome for the primary endpoint at week 18 is presented cross-sectionally. No differences are seen in CAS score between NR-RTX and R-GC, or between NR-RTX and NR-C. A difference is presented in ΔCAS from week 4, comparing R-GC and NR-RTX, where R-GC increases their CAS, whereas NR-RTX decreases in CAS. No differences between NR-RTX and R-GC were seen in either CAS SU/M, severity, subjective diplopia or proptosis at week 18 (table 6).

Table 5. Efficacy variables at week 12 for the total RTX study population (t-RTX) and the control group (NR-C). The t-RTX is presented separated by the response at 4-weeks: responder to (R-GC) and non-responder (NR-RTX). Variables are presented as n (%), mean ± standard deviation or median (interquartile range). Statistical significance is set at 0.05.

Variable	t-RTX separated by response at 4-weeks		NR-C n = 11	Test between groups, p-value	
	R-GC n=13	NR-RTX n=10		NR-RTX R-GC	NR-RTX NR-C
CAS	2.9 ± 1.3	4.2 ± 1.0	3.6 ± 2.0 *n = 9*	0.02* [c]	0.4[c]
Δ*CAS* from week ***4*** [a]	-1.5 ± 1.9	0.2 ± 1.6	0.7 ± 1.5 *n=9*	0.6 [c]	0.5 [c]
CAS SU/M	4.9 ± 2.1	8.2 ± 2.9		0.005* [c]	
ΔCAS SU/M from week 4 [a]	0.2 ± 2.7	0.1 ± 2.4		0.9 [c]	
Severity	8.3 ± 3.2	10.4 ± 2.6		0.1 [c]	
ΔSeverity from week 4 [a]	0.0 (-2.5, 1.5)	-1.0 (-2.3, 1.50)		0.9 [d]	
Subjective diplopia					
No diplopia	5 (38.5)	4 (40.0)			
Intermittent	3 (23.1)	3 (30.0)			
Can be corrected with prism	2 (15.4)	2 (20.0)		0.9[e]	
Constant	3 (23.1)	1 (10.0)			
Improvement proptosis [b]	0 (0.0)	1 (10) *n=10*	1 (14.3) *n =7*	0.4[f]	1.0[f]

** Significant p-value*

[a] Δ of variable = (value at baseline – value at 4 weeks). As high values in these variables represent worse status, positive Δ of variable means decrease in the score through the study and therefore improvement in the patients TAO or QoL.

[b] Proptosis is seen as improved if one of the eyes has a decrease in Hertel score by ≥ 2 points and if the other eye has not increased in Hertel score by < 1 points.

[c] T test was used with a set alpha level of 0.05

[d] Mann Whitney U test was used with a set alpha level of 0.05

[e] Chi Square test was used with a set alpha level of 0.05

[f] Fishers Exact test was used with a set alpha level of 0.05

Abbreviations: CAS, clinical activity score; NR-C, control group consisting of non-responders to treatment with glucocorticoids but continued with the treatment; NR-RTX, non-responders to treatment with glucocorticoids that changed treatment to RTX; GC, responders to treatment with glucocorticoids, RTX, Rituximab; t-RTX, total study group of the RTX study; SU/M, Mölndal hospital,

Table 6. Efficacy variables at week 18 for the total RTX study population (t-RTX) and the control group (NR-C). The t-RTX is presented separated by the response at 4-weeks: responder to (R-GC) and non-responder (NR-RTX). Variables are presented as n (%), mean ± standard deviation or median (interquartile range). Statistical significance is set at 0.05.

Variable	t-RTX separated by response at 4-weeks		NR-C n = 11	Test between groups, p-value	
	R-GC n=13	NR-RTX n=10		NR-RTX R-GC	NR-RTX NR-C
CAS	3.9 ± 1.4	3.9 ± 1.0	3.3 ± 1.4 *n=7*	0.9[d]	0.3[d]
Δ*CAS* *from week 4*[a]	-1.2 ± 1.2	0.5 ± 1.4	0.6 ± 1.5 *n=7*	0.006* [d]	0.9[d]
Δ*CAS* *from week 12* [a]	-0.83 ± 1.5	0.3 ± 1.4 0.5 (-0.5, 1.3)	0.0 (-1.5, 0.5)	0.08 [d]	0.3[e]
CAS SU/M	6.5 ± 2.6	7.6 ± 2.3		0.3[d]	
ΔCAS SU/M from week 4 [a]	-1.3 ± 2.8	0.7 ± 2.3		0.08[d]	
ΔCAS SU/M from week 12 [a]	-2.0 (-3.5, 1.0)	1.0 (0.0, 2.0)		0.06 [c]	
Severity	8.2 ± 4.2	10.0 ± 2.9		0.3[d]	
ΔSeverity from week 4 [a]	-0.8 ± 2.5	-0.1 ± 2.0		0.5[d]	
Δ*Severity* *from week 12* [a]	0.08 ± 2.3	0.4 ± 2.5		0.6 [d]	
Subjective diplopia					
No diplopia	7 (53.8)	2 (20.0)			
Intermittent	1 (7.7)	3 (30.0)			

Can be corrected with prism	2 (15.4)	3 (30.0)		0.3[f]	
Constant	3 (23.1)	2 (20.0)			
Improvement of proptosis[b]	0 (0.0)	0 (0.0)	1 (20.0) *n=5*	[c]	0.3[g]

** Significant p-value*
[a] Δ of variable = (value at week 4 – value at week 18 or value at week 12 – value at week 18). As high values in these variables represent worse status, positive Δ of variable means decrease in the score through the study and therefore improvement in the patients'TAO.
[b] Proptosis is seen as improved if one of the eyes has a decrease in Hertel score by ≥ 2 points and if the other eye has not increased in Hertel score by < 1 points.
[c] P-value is missing due to 0% of the groups compared had improvement of proptosis according to given criteria..
[d] T test was used with a set alpha level of 0.05
[e] Mann Whitney U test was used with a set alpha level of 0.05
[f] Chi Square test was used with a set alpha level of 0.05
[g] Fishers Exact test was used with a set alpha level of 0.05
Abbreviations: CAS, clinical activity score; NR-C, control group consisting of non-responders to treatment with glucocorticoids but continued with the treatment; NR-RTX, non-responders to treatment with glucocorticoids that changed treatment to RTX; GC, responders to treatment with glucocorticoids, RTX, Rituximab; t-RTX, total study group of the RTX study; SU/M, Mölndal hospital.

Table 7 displays the results of longitudinal analysis of CAS within the three different groups. The R-GC group is the only group that significantly increases in CAS from week 12 to 18 and from week 4 to week 18. The R-GC also has a significant decrease in CAS from baseline to week 4, which is presented for the NR-C as well. After the decrease in CAS of NR-C from baseline to week 4, no other differences in CAS is displayed for this group. No differences for NR-RTX in CAS is presented.

Table 7. Longitudinal comparisons of the clinical activity score (CAS). This table presents how the CAS of each group changes through the course of the study. The total RTX study population (t-RTX) presented separated by the response at 4-weeks: responder to (R-GC) and non-responder (NR-RTX). Variables are presented as mean ± standard deviation or median (interquartile range). Statistical significance is set at alpha level of 0.05.

	R-GC *n = 13*	**NR-RTX** *n=10*	**NR-C**
CAS at baseline	4.8 ± 1.3	4.0 (4.0, 5.0)	5.0 (5.0, 5.0) *n=12*
CAS at week 4	2.7 ± 1.4	4.4 ±1.2	4.0 (3.3, 5.0) *n=12*
CAS at week 12	2.9 ± 1.3	4.20 ± 1.0	3.6 ± 2.0 *n=11*
CAS at week 18	3.9 ± 1.3	3.9 ± 1.0	3.3 ± 1.4 *n=7*
Test between baseline and week 4, p-value	<0.001* [a]	0.7 [b]	0.02* [b]
Test between week 4 and week 12, p-value	0.8 [a]	0.7 [a]	0.2[a]
Test between week 12 and week 18, p-value	0.04* [a]	0.5 [a]	0.6[a]
Test between week 4 and week 18, p-value	0.005* [a]	0.3 [a]	0.4[a]

** Significant p-value*
[a] Paired T test was used with a set alpha level of 0.05.
[b] Wilcoxon signed-rank test was used with a set alpha level of 0.05.
Abbreviations: CAS, clinical activity score; NR-C, control group consisting of non-responders to treatment with glucocorticoids but continued with the treatment; NR-RTX, non-responders to treatment with glucocorticoids that changed

treatment to RTX; R-GC, responders to treatment with glucocorticoids, RTX, Rituximab; t-RTX, total study group of the RTX study.

Quality of life

Table 8 presents the results of the ThyPRO questionnaire. Only one difference can be found when comparing NR-RTX and R-GC, which is that the NR-RTX had larger influence on their sex life at week 12, compared to the R-GC.

In table 9, longitudinal analysis of the ThyPRO results is displayed, R-GC and NR-RTX are showed respectively. Between baseline and week 4, the R-GC decreased their hyperthyroid symptoms 2 (dimension 3), eye symptoms and their cosmetic complaints. Between baseline and week 12, an increase of hyperthyroid symptoms 1 (dimension 2) was observed among the R-GC, as well as decrease of eye symptoms and anxiety. The R-GC group presented with a decrease in eye symptoms and less impairment of the sex life, comparing baseline with week 18. Lastly for the R-GC, an increase of hyperthyroid symptoms 2 (dimension 3) is displayed when comparing the results of week 4 with week 18.

Regarding the NR-RTX, a significant decrease in goiter symptoms was displayed between baseline and week 4. Otherwise, no other differences were displayed for the NR-RTX regarding ThyPRO.

Table 8. The results of the ThyPRO self-assessment quality of life (QoL) questionnaire. The questionnaire consists of 13 different dimensions of QoL affliction. Here, the normalised score is presented (raw score/max score * 100), giving each item a score of 0-100 points. A higher score indicates more symptoms or greater impact on QoL of the disease. The total RTX study population (t-RTX) presented separated by the response at 4-weeks: responder to (R-GC) and non-responder (NR-RTX). Variables are presented as median (interquartile range). Statistical significance is set at alpha level of 0.05.

ThyPRO scale	R-GC	NR-RTX	Test between groups, p-value
Goiter symptoms			
Baseline	9.0 (2.0, 19.0)	10.0 (6.0, 36.8)	0.6[a]
Week 4	7.0 (5.0, 21.5)	9.0 (0.5, 10.5)	0.6[a]
Week 12	14.0 (8.0, 34.0)	16.0 (3.3, 36.0)	0.9[a]
Week 18	9.0 (2.0, 23.0)	13.5 (5.5, 31.5)	0.3[a]
Hyperthyroid symptoms 1			
Baseline	25.0 (16.0, 34.0)	19.0 (6.0, 40.8)	0.5[a]

Week 4	28.0 (9.0, 36.0)	26.5 (19.8, 42.5)	0.7[a]
Week 12	31.0 (22.0, 45.5)	19.0 (6.0, 55.0)	0.4[a]
Week 18	22.0 (9.0 (39.0)	19.0 (6.8, 55.0)	0.9[a]
Hyperthyroid symptoms 2			
Baseline	31.0 (19.0, 41.0)	16.0 (7.8, 54.5)	0.3[a]
Week 4	19.0 (6.0, 28.0)	13.0 (6.0, 25.0)	0.7[a]
Week 12	25.0 (13.0, 47.0)	22.0 (14.5, 42.5)	0.9[a]
Week 18	25.0 (13.0, 56.0)	16.0 (7.8, 28.0)	0.3[a]
Eye symptoms			
Baseline	66.0 (50.0, 86.0)	59.5 (32.8, 80.3)	0.9[a]
Week 4	34.0 (22.0, 53.5)	51.5 (25.0, 65.3)	0.3[a]
Week 12	38.0 (28.0, 53.5)	62.5 (29.8, 81.8)	0.1[a]
Week 18	31.0 (20.5, 59.5)	50.0 (21.3, 76.5)	0.3[a]
Tiredness			
Baseline	46.0 (39.5, 64.5)	48.0 (15.8, 81.0)	0.9[a]
Week 4	36.0 (19.5, 62.5)	55.5 (25.5, 76.3)	0.3[a]
Week 12	46.0 (28.5, 62.5)	53.5 (23.0, 90.3)	0.7[a]
Week 18	57.0 (21.5, 64.0)	46.5 (20.8, 74.0)	0.9[a]
Cognition			
Baseline	17.0 (6.0, 37.5)	15.0 (6.3, 54.0)	0.8[a]
Week 4	17.0 (6.0, 33.5)	23.0 (9.3, 60.8)	0.3[a]
Week 12	21.0 (4.0, 39.5)	25.0 (6.3, 55.0)	0.5[a]
Week 18	13.0 (4.0, 27.0)	19.0 (9.3, 66.8)	0.3[a]
Anxiety			
Baseline	25.0 (12.5, 48.0)	27.0 (8.0, 58.5)	0.8[a]
Week 4	25.0 (12.5, 33.5)	44.0 (7.3, 57.0)	0.3[a]
Week 12	17.0 (0.0, 25.0)	33.5 (2.0, 67.8)	0.3[a]
Week 18	17.0 (4.0, 31.5)	29.5 (3.3, 59.8)	0.6[a]
Depressivity			
Baseline	29.0 (23.0, 44.0)	30.5 (22.0, 61.3)	0.9[a]
Week 4	32.0 (19.5, 39.0)	46.0 (17.8, 68.0)	0.2[a]
Week 12	36.0 (21.0, 44.5)	39.0 (14.0, 79.0)	0.6[a]
Week 18	29.0 (16.0, 48.0)	34.0 (14.0, 85.5)	0.7[a]
Emotional susceptibility			
Baseline	31.0 (18.0, 28.0)	35.0 (7.3, 84.8)	0.9[a]
Week 4	22.0 (15.0, 42.0)	50.0 (10.8, 88.3)	0.2[a]
Week 12	33.0 (16.5, 54.0)	57.0 (13.0, 81.0)	0.3[a]
Week 18	31.0 (12.5, 53.0)	42.0 (8.8, 70.3)	0.6[a]
Impaired social life			
Baseline	13.0 (0.0, 28.0)	15.5 (0.0, 59.5)	0.7[a]
Week 4	6.0 (3.0, 25.0)	25.5 (1.5, 54.5)	0.3[a]
Week 12	13.0 (9.5, 28.0)	22.0 (6.0, 79.5)	0.6[a]
Week 18	19.0 (9.5, 22.0)	31.5 (6.0, 67.5)	0.6[a]
Impaired daily life			
Baseline	42.0 (17.0, 52.0)	35.5 (7.3, 65.8)	0.7[a]
Week 4	25.0 (13.0, 56.0)	31.5 (10.3, 76.8)	0.7[a]
Week 12	42.0 (8.5, 66.5)	38.0 (15.0, 83.5)	0.7[a]
Week 18	46.0 (4.0, 54.0)	39.5 (17.0, 77.0)	0.4[a]
Impaired sex life			
Baseline	50.0 (13.0, 56.5)	31.5 (25.0, 78.5)	0.9[a]
Week 4	25.0 (6.5, 50.0)	25.0 (25.0, 93.8)	0.4[a]
Week 12	25.0 (0.0, 38.0)	69.0 (12.5, 100.0)	0.03*[a]
Week 18	13.0 (0.0, 37.5)	25.0 (3.3, 75.0)	0.2[a]
Cosmetic complaints			
Baseline	38.0 (25.0, 46.0)	42.0 (8.0, 59.8)	0.9[a]
Week 4	21.0 (19.0, 44.0)	45.5 (6.3, 66.0)	0.7[a]
Week 12	33.0 (12.5, 68.5)	46.0 (7.3, 73.0)	1.0[a]
Week 18	38.0 (21.0, 54.0)	50.5 (9.3, 78.0)	0.7[a]

** Significant p-value*

[a] Mann Whitney U test was used with a set alpha level of 0.05.

Abbreviations: NR-RTX, non-responders to treatment with glucocorticoids that changed treatment to RTX; R-GC, responders to treatment with glucocorticoids, RTX, Rituximab; t-RTX, total study group of the RTX study; QoL, quality of life

Table 9. Results of longitudinal analysis of ThyPRO results. This table presents whether there are significant differences within the dimensions of ThyPRO questionnaire between the different weeks. The data that was used is presented in table 8. The total RTX study population (t-RTX) presented separated by the response at 4-weeks: responder to (R-GC) and non-responder (NR-RTX). Variables are presented as median (interquartile range). Statistical significance was set at 0.05.

ThyPRO scale	R-GC Tests between weeks, p-value				NR-RTX Tests between weeks, p-value			
	Baseline – Week 4	*Baseline – Week 12*	*Baseline – Week 18*	*Week 4 – Week 18*	*Baseline – Week 4*	*Baseline – Week 12*	*Baseline – Week 18*	*Week 4 – Week 18*
Goiter symptoms	0.9[a]	0.3[a]	1.0[a]	0.9[a]	0.03*[a]	0.9[a]	0.7[a]	0.1[a]
Hyperthyroid symptoms 1	1.0[a]	0.009*[a]	0.7[a]	0.9[a]	0.6[a]	0.7[a]	0.8[a]	0.7[a]
Hyperthyroid symptoms 2	0.04*[a]	0.7[a]	0.6[a]	0.006*[a]	0.5[a]	0.8[a]	0.5[a]	0.4[a]
Eye symptoms	0.002*[a]	0.002*[a]	0.003*[a]	0.4[a]	0.3[a]	0.8[a]	0.5[a]	0.8[a]
Tiredness	0.07[a]	0.8[a]	0.6[a]	0.2[a]	0.4[a]	0.1[a]	0.7[a]	0.8[a]
Cognition	0.5[a]	0.8[a]	0.5[a]	0.8[a]	0.8[a]	1.0[a]	0.9[a]	0.9[a]
Anxiety	0.3[a]	0.02*[a]	0.1[a]	0.4[a]	0.1[a]	0.5[a]	0.6[a]	0.3[a]
Depressivity	0.8[a]	0.8[a]	0.6[a]	0.8[a]	0.4[a]	0.6[a]	0.6[a]	0.7[a]
Emotional susceptibility	0.3[a]	0.7[a]	0.7[a]	0.2[a]	0.8[a]	0.7[a]	0.7[a]	0.4[a]
Impaired social life	0.4[a]	0.4[a]	0.3[a]	0.2[a]	0.8[a]	0.5[a]	0.5[a]	0.2[a]
Impaired daily life	0.5[a]	0.9[a]	0.9[a]	0.9[a]	0.8[a]	0.7[a]	0.5[a]	0.2[a]
Impaired sex life	0.2[a]	0.08[a]	0.03*[a]	0.2[a]	0.6[a]	0.1[a]	0.6[a]	0.6[a]
Cosmetic complaints	0.02*[a]	0.6[a]	0.9[a]	0.08[a]	0.5[a]	0.4[a]	0.2[a]	0.1[a]

** Significant p-value*

[a] Wilcoxon signed-rank test was used with a set alpha level of 0.05.

Abbreviations: NR-RTX, non-responders to treatment with glucocorticoids that changed treatment to RTX; R-GC, responders to treatment with glucocorticoids, RTX, Rituximab; t-RTX, total study group of the RTX study; QoL, quality of life.

Discussion

This is a unique study of the RTX treatment in patients with TAO that do not respond to the primary treatment that is consisting of steroids. From this study, it cannot be proven that RTX result in a more beneficial outcome than iv GC using the conventional 12-week scheme. On the contrary, the group of patients that were characterized as responders (R-GC), improved their CAS significantly during the first 4 weeks and kept their CAS unchanged for the remaining 8 weeks.

At week 18, 38.5% of the R-GC had a reactivation in their disease, displayed by the five responders that had a relapse. This made the groups of R-GC, NR-RTX and NR-C ending up at similar CAS levels at primary end point. This frequency of relapse for patients receiving iv GC for TAO has been reported earlier, for instance in Salvi *et al* (9), where they in 2015 reported a disease reactivation of 30% in the patients. In this study, the frequency of relapses may be affected by that comparative control group had received a *per oral* glucocorticoid scheme after the 12 weeks of iv GC, which were given to five out of a total of twelve patients in the NR-C patients. It is important to bear in mind that the NR-C was non-respondent to GC treatment. This is agreement with that the patients that obtained *per oral* steroids did not present with better CAS compared to the other patients in NR-C. The clinical experience of relapses post iv GC may be a reason to why the Departments of Ophthalmology in VGR do not follow international guidelines (2), where a tapering down scheme of *per oral* steroids is not recommended.

This study clearly demonstrates the lack of efficient treatments in almost half of the patients in need of anti-inflammatory treatments with TAO. The non-response rate is higher than that

reported by Kahaly *et al* (17) but equal to the figures derived from a recent Lancet publication comparing iv GC alone with a combination with mycophenolate (24). It may be questioned that 4 weeks may be too short to judge if a patient is a responder or not, but the results of the comparison group that fulfilled the 12-week scheme was similar after 12 weeks treatment. This is an argument that steroid non-responsive patients can be classified already at 4 weeks. This evaluation point was chosen, as it is still early in the course, and it was a balance between the lack of efficacy with the risk of side effects giving the patients the option of a new promising treatment (18, 19) as this study was planned before the results of the RCTs (21, 22). Also, there was a report with evaluation at 4 weeks using the same iv GC scheme (25) It seems that the response rate to iv GC is rather 50-70%, not 75-80% (9).

Is RTX a good second line treatment in those that are non-respondent to iv GC in moderate-severe TAO? The argument for adding it as a second line treatment is in the latest international guidelines (2) is that it seems superior to iv GC with less side effect (9) also using the collapsed information from the publication by Stan *et al* (26) in a meta-analysis. The results are contradictory with similar response of RTX as to NaCl in the other RCT (21). The lack of efficacy in that study has been questioned, as the patients may have come too far in the course of the disease. In this study, it is also demonstrated a lack of effect, even though the patients are early in their course, and yet RTX seems not to be better than iv GC against moderate-severe TAO. At least not the first 18 weeks. The full effect of RTX may still be awaited (9) and it will be interesting to look at the 32-68 week data and the profile of adverse events, but it could equally be that neither RTX nor iv GC has any effect in the non-responsive patients, and the changes that are seen illustrate the natural course. Though, based on the data presented in this master thesis, it is hard to believe that RTX is a good second line

treatment in patients with TAO that do not respond to the first line treatment that consists of iv glucocorticoids.

However, it is of great importance to define the group of non-responsive patients early in the course, as there are other promising therapies entering the stage approaching a different mechanism (27). As seen in table 3, the lack of differences between the t-RTX and the NR-C regarding background variables at baseline may imply that the patients included in the RTX-study are representative for the population of patients with clinically active TAO. Though, both the included patients in the RTX-study and the NR-C have been selected through multiple criteria (since it is not a RCT), which might explain their similarities. Smoking is a known risk factor for TAO, but as table 3 and 4 displays, it seems to not have an impact on whether the patient will respond to iv GC or not.

Something that may have an impact on whether the patient will respond to iv GC is gender. Men are more likely to belong to the non-responsive group (11). In fact, only 8% of the responders were men in this study, whereas men represented 50% in the NR-RTX group. The reason to this is unknown but may illustrate that men have a more severe disease, which agrees with what Marocci and Bartalena et al stated in 1989 (28).

Efficacy is determined not only by anti-inflammatory effect and orbital effects but is also measured as patients' satisfaction and QoL. The results of the ThyPRO questionnaire presented no improvement of QoL in either of the 13 dimensions for the NR-RTX at week 18. The effect of lesser goitre symptoms of the NR-RTX was obtained during their initial treatment with iv GC, thus that is nothing RTX can obtain credentials for. The R-GC, however, both have increasing QoL in some aspects and decreasing QoL in other during the

course of the study. The decrease of eye symptoms was significant in all comparisons with baseline for R-GC, but not when comparing week 4 with week 18, most likely because the eye symptoms had reduced by week, much like their CAS.

Strengths and limitations of the study

A strength of this study is that is has a longitudinal design, which makes it possible to analyse long-term effects both of treatment with iv GC and with RTX. Another strength of this study comparison group of non-respondent patients has been collected. This study reflects reality, and with RTX to non-responders according to the current guidelines (2), but it would have been a strength if this step had included a randomisation between RTX and iv GC, but this was limited from the long list of contraindications for TAO treatment and the limited number of available patients, as TAO is a rare disease (29).

A limitation of this study is the limited sample size, but it was unethical to include more patients when the post-hoc analysis showed no effect beyond iv GC at 12-weeks. The sample size of the study is the same size as other studies in this area (9, 19, 21). Though, because of the reasonably small group of patients analysed, the effect of different drop outs and exclusions get even larger. Therefore, it is good that primary endpoint of the study occurs before extracting the relapse at responders at week 18, even though this may affect the results of the responders at week 18 since a total of 5 patients had a relapse. After this, the different groups will get even smaller and it will be harder to obtain a reliable result.

Another limitation is that the control group was created retrospectively. It would have been better if the control group was included in the RTX-study and had its assessment according to the study protocol with the same interval between visits as the t-RTX. This would also have

made it possible to compare other efficacy variables, other than CAS and proptosis. If this had been the case, patients in the NR-C had not received *per oral* steroids after the given treatment of iv GC and then it would have been possible to evaluate the incidence of relapse among the non-responders who received monotherapy with iv GC as well.

Ideas for further research

This study has not shown that it is better nor worse to change the TAO treatment from GC to RTX if you are non-responsive to iv GC. But before a conclusion can be made, adverse events should be analysed, and the full longitudinal data be explored. If the patients that obtained RTX obtain fewer adverse events than the ones that obtains 12 weeks treatment with IV GC, could that maybe be of RTX favour?

Another idea for a study would be, as mentioned earlier, if NR-RTX in a future study would be randomized to either continue with iv GV or to obtain RTX. In that way, a control group would not have to be created retrospectively. Due to the results of this study, a future study maybe would be done in another way. Why not iv GC+RTX in one arm vs iv GC only and make it multicentre?

Conclusions

This study shows no benefit of treatment with RTX compared to iv GC for patients with moderate-severe TAO that are non-responding to treatment with iv GC, and from these data RTX would be no good second line choice. Further evaluations of this data are necessary in order to assess other aspects of the treatment, such as adverse events, in order to evaluate the treatment more properly.

Populärvetenskaplig sammanfattning på svenska

Ny behandling testad för svår ögonsjukdom

Varje år i Sverige insjuknar flera människor i en autoimmun sjukdom, Graves sjukdom (också kallad giftstruma), som påverkar sköldkörteln. Graves sjukdom gör att sköldkörteln blir överaktiv och de drabbade kan till exempel drabbas av viktnedgång, hjärtklappning, hårförlust och trötthet.

Upp till två tredjedelar av de som drabbas av Graves sjukdom har också symptom från ögonen, så kallad Graves oftalmopati eller giftstrumeögon. Denna komplikation drabbar ögonen och den kan bli ytterst allvarlig. Symptom på Graves oftalmopati är till exempel att man ser dubbelt, får rinnande, skavande och utåtstående ögon och i värsta fall påverkan på synnerven som kan leda till blindhet. Sjukdomen delas in i olika stadier och det finns olika behandlingar tillgängliga.

Den vanligaste behandlingen mot Graves oftalmopati är höga doser av läkemedlet kortison, som ges direkt in i blodet. Kortison fungerar tyvärr inte på cirka en fjärdedel av patienterna. De höga doserna av kortison orsakar också många svåra biverkningar, så som sömnpåverkan och hjärtklappning. De senaste åren har studier gjorts på en alternativ behandling mot Graves oftalmopati, Rituximab, och det är idag det läkemedel man ger i andra hand, om kortison inte fungerar.

I en ny studie gjord vid Sahlgrenska Universitetssjukhuset har man studerat Rituximabs effekt hos patienter med Graves oftalmopati som inte blivit bättre i sin sjukdom av kortisonbehandlingen. Detta gjordes genom att effekten av kortison utvärderades fyra veckor

in i behandlingen. Bedömdes det att kortison inte haft önskad effekt hos patienterna i studien bytte patienten behandling till Rituximab.

I studien hade kortison, i de fall behandlingen fungerade, en snabb effekt mot sjukdomen, men när den 12 veckor långa behandlingen avslutats så försämrades två femtedelar av patienterna inom sex veckor i sin sjukdom igen. Rituximab gav varken en snabb eller tydlig effekt på sjukdomen. Däremot såg man att patienterna som behandlats med Rituximab blev mer stabila i sin sjukdom och de drabbades inte av den typ av försämring i sin sjukdom, som de som fick kortison gjorde. Dock jämfördes resultatet hos de som fått Rituximab med patienter som fortsatt med kortisonbehandling trots utebliven effekt. I de jämförelserna såg man att sjukdomen stabiliserats även i de fall då patienterna fortsatt med kortison.

Andra analyser gjordes i studien, som visade att Graves oftalmopati är vanligare bland kvinnor, men att det är vanligare att män inte svarar på behandlingen med kortison. Studien visade också att livskvaliteten varken förbättras eller försämras av Rituximab, medan den, i olika avseenden, både förbättras och försämras under kortisonbehandlingen.

En fortsättning av studien är att vänta, där även biverkningarna av de två olika behandlingarna kommer att analyseras. Patienterna följdes även i drygt ett år, så resultat på en mer långsiktig effekt av både kortison och Rituximab är att vänta. Men en slutsats man redan nu kan dra är att Rituximab inte är en bättre behandling än kortison mot Graves oftalmopati, även om fler studier på området krävs.

Acknowledgement

I would like to express my gratitude to my supervisor Helena Filipsson Nyström and thank her for believing in me enough to be a part of this study and for her help, guidance and support during this work. Also, I would like to say a sincerely thank you to my co-supervisor Sofia Manousou, especially for her patience with me during the statistical part of this work.

I would also like to express my appreciation to Oscar Ragnarsson and to my opponent Margareta Bånnsgård for listening to my oral presentation and giving me valuable feedback.

Appendices

Table 10. Clinical Activity Score SU/M (CAS SU/M). *This is an extension of the international used Clinical Activity Score (Table 1). It was created at the Department of Ophthalmology at Mölndal hospital, a part of the Sahlgrenska University hospital. Instead of each item having a maximum score of one point, each item can score 0-3 points. With CAS SU/M it is possible to even more accurate score the inflammatory activity in TAO.*

Item	Point and criteria
Pain	1. Pain occasionally during the last months 2. Pain occurring daily. 3. Constant pain.
Pain on eye movement	1. Pain or discomfort on eye movement during the last months. 2. Pain on eye movement occurring daily. 3. Pain with every eye movement.
Eyelid swelling	1. New swelling of upper or lower eyelid. 2. New swelling of upper **and** lower eyelid. 3. Bags by upper and/or lower eyelid
Eyelid erythema	1. New redness pretarsally (not blepharitis) 2. New redness preseptally
Chemosis	1. Significant conjunctival swelling, but < 1/3 of rima in pp 2. Conjunctival swelling of ≥ 1/3 of rima in pp
Conjunctival injection	1. Significant conjunctival redness covering < 50% of visible bulb 2. Conjunctival redness ≥ 50% of visible bulb
Inflammation caruncle	1. Injection over muscle attachments 2. Inflammatory redness of plica and/or caruncle
Increased proptosis	1. ≥ 2 mm proptosis from previous visit either eye
Worsened visual impairment	Visual impairment (TAO-related) ≥ 0.1 logMAR since previous visit either eye
Worsened motility restriction	Decreased duction any eye muscle ≥ 8° since previous visit

Abbreviations: CAS, clinical activity score; SU/M, Mölndal hospital.

References

1. 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(6):899-905.
2. Bartalena L, Baldeschi L, Boboridis K, Eckstein A, Kahaly GJ, Marcocci C, et al. The 2016 European Thyroid Association/European Group on Graves' Orbitopathy Guidelines for the Management of Graves' Orbitopathy. Eur Thyroid J. 2016;5(1):9-26.
3. Smith TJ, Hegedus L. Graves' Disease. N Engl J Med. 2016;375(16):1552-65.
4. Subekti I, Pramono LA. Current Diagnosis and Management of Graves' Disease. Acta Med Indones. 2018;50(2):177-82.
5. Kahaly GJ, Bartalena L, Hegedus L, Leenhardt L, Poppe K, Pearce SH. 2018 European Thyroid Association Guideline for the Management of Graves' Hyperthyroidism. Eur Thyroid J. 2018;7(4):167-86.
6. Mohlin E, Filipsson Nystrom H, Eliasson M. Long-term prognosis after medical treatment of Graves' disease in a northern Swedish population 2000-2010. Eur J Endocrinol. 2014;170(3):419-27.
7. Samuels MH, Kolobova I, Niederhausen M, Janowsky JS, Schuff KG. Effects of Altering Levothyroxine (L-T4) Doses on Quality of Life, Mood, and Cognition in L-T4 Treated Subjects. J Clin Endocrinol Metab. 2018;103(5):1997-2008.
8. Gillespie EF, Smith TJ, Douglas RS. Thyroid eye disease: towards an evidence base for treatment in the 21st century. Curr Neurol Neurosci Rep. 2012;12(3):318-24.
9. Salvi M, Vannucchi G, Curro N, Campi I, Covelli D, Dazzi D, et al. Efficacy of B-cell targeted therapy with rituximab in patients with active moderate to severe Graves' orbitopathy: a randomized controlled study. J Clin Endocrinol Metab. 2015;100(2):422-31.
10. Bartalena L, Tanda ML. Clinical practice. Graves' ophthalmopathy. N Engl J Med. 2009;360(10):994-1001.
11. Sahli E, Gunduz K. Thyroid-associated Ophthalmopathy. Turk J Ophthalmol. 2017;47(2):94-105.
12. Genere N, Stan MN. Current and Emerging Treatment Strategies for Graves' Orbitopathy. Drugs. 2019;79(2):109-24.
13. Minakaran N, Ezra DG. Rituximab for thyroid-associated ophthalmopathy. Cochrane Database Syst Rev. 2013(5):Cd009226.
14. Weiler DL. Thyroid eye disease: a review. Clin Exp Optom. 2017;100(1):20-5.
15. Mourits MP, Prummel MF, Wiersinga WM, Koornneef L. Clinical activity score as a guide in the management of patients with Graves' ophthalmopathy. Clin Endocrinol (Oxf). 1997;47(1):9-14.
16. Bartalena L, Baldeschi L, Dickinson A, Eckstein A, Kendall-Taylor P, Marcocci C, et al. Consensus statement of the European Group on Graves' orbitopathy (EUGOGO) on management of GO. Eur J Endocrinol. 2008;158(3):273-85.
17. Kahaly GJ, Pitz S, Hommel G, Dittmar M. Randomized, single blind trial of intravenous versus oral steroid monotherapy in Graves' orbitopathy. J Clin Endocrinol Metab. 2005;90(9):5234-40.
18. Salvi M, Vannucchi G, Campi I, Beck-Peccoz P. Rituximab in the treatment of thyroid eye disease: science fiction? Orbit. 2009;28(4):251-5.
19. Khanna D, Chong KK, Afifiyan NF, Hwang CJ, Lee DK, Garneau HC, et al. Rituximab treatment of patients with severe, corticosteroid-resistant thyroid-associated ophthalmopathy. Ophthalmology. 2010;117(1):133-9.e2.
20. Einarsson JT, Evert M, Geborek P, Saxne T, Lundgren M, Kapetanovic MC. Rituximab in clinical practice: dosage, drug adherence, Ig levels, infections, and drug antibodies. Clin Rheumatol. 2017;36(12):2743-50.

21. Stan MN, Garrity JA, Carranza Leon BG, Prabin T, Bradley EA, Bahn RS. Randomized controlled trial of rituximab in patients with Graves' orbitopathy. J Clin Endocrinol Metab. 2015;100(2):432-41.
22. Shen WC, Lee CH, Loh EW, Hsieh AT, Chen L, Tam KW. Efficacy and Safety of Rituximab for the Treatment of Graves' Orbitopathy: A Meta-analysis of Randomized Controlled Trials. Pharmacotherapy. 2018;38(5):503-10.
23. Watt T, Hegedus L, Rasmussen AK, Groenvold M, Bonnema SJ, Bjorner JB, et al. Which domains of thyroid-related quality of life are most relevant? Patients and clinicians provide complementary perspectives. Thyroid. 2007;17(7):647-54.
24. Kahaly GJ, Riedl M, Konig J, Pitz S, Ponto K, Diana T, et al. Mycophenolate plus methylprednisolone versus methylprednisolone alone in active, moderate-to-severe Graves' orbitopathy (MINGO): a randomised, observer-masked, multicentre trial. Lancet Diabetes Endocrinol. 2018;6(4):287-98.
25. Zhu W, Ye L, Shen L, Jiao Q, Huang F, Han R, et al. A prospective, randomized trial of intravenous glucocorticoids therapy with different protocols for patients with graves' ophthalmopathy. J Clin Endocrinol Metab. 2014;99(6):1999-2007.
26. Stan MN, Salvi M. MANAGEMENT OF ENDOCRINE DISEASE: Rituximab therapy for Graves' orbitopathy - lessons from randomized control trials. Eur J Endocrinol. 2017;176(2):R101-r9.
27. Smith TJ, Kahaly GJ, Ezra DG, Fleming JC, Dailey RA, Tang RA, et al. Teprotumumab for Thyroid-Associated Ophthalmopathy. N Engl J Med. 2017;376(18):1748-61.
28. Marcocci C, Bartalena L, Bogazzi F, Panicucci M, Pinchera A. Studies on the occurrence of ophthalmopathy in Graves' disease. Acta Endocrinol (Copenh). 1989;120(4):473-8.
29. Perros P, Hegedus L, Bartalena L, Marcocci C, Kahaly GJ, Baldeschi L, et al. Graves' orbitopathy as a rare disease in Europe: a European Group on Graves' Orbitopathy (EUGOGO) position statement. Orphanet J Rare Dis. 2017;12(1):72.