Outcome of Graves' Disease

Mats Holmberg, MD

Department of Internal Medicine
Institute of Medicine
Sahlgrenska Academy, University of Gothenburg

UNIVERSITY OF GOTHENBURG

Gothenburg 2020

Outcome of Graves' Disease

mats.o.holmberg@gu.se

ISBN978-91-7833-772-9(PRINT)
ISBN 978-91-7833-773-6 (PDF)

http://hdl.handle.net/2077/62211

Printed in Mölndal, Sweden 2020
Printed by BrandFactory

Deep in the forest there is an unexpected glare that only can be found by one who's lost

Tomas Tranströmer

Outcome of Graves' Disease

Mats Holmberg, MD

Department of Internal Medicine, Institute of Medicine
Sahlgrenska Academy, University of Gothenburg
Gothenburg, Sweden

ISBN978-91-7833-772-9(PRINT)
ISBN 978-91-7833-773-6 (PDF)

http://hdl.handle.net/2077/62211

ABSTRACT

Graves' disease (GD) is an autoimmune condition with elevated thyroid hormone levels and symptoms suggesting an affected brain. These symptoms often resolve with treatment but, for some patients, GD results in a long period of reduced well-being. The overall aim of this thesis is to characterize the consequences of GD with a special focus on the brain. Three studies were conducted with data from questionnaires and clinical assessments of patients with GD in both the hyperthyroid and the euthyroid phase.

Paper I was a longitudinal cohort study that assessed long-term treatment results 6–10 years after the onset of GD. Among 1186 included patients, the 774 who were initially treated with antithyroid drugs had a 40% chance of being euthyroid without thyroid medication at follow-up. One in four patients did not feel fully recovered.

Paper II was a longitudinal case-control study designed to characterize affective and cognitive symptoms in 65 premenopausal women with newly diagnosed untreated GD. At onset of GD, the patients were significantly more affected with depression, anxiety, and mental fatigue compared to controls. At follow-up after 15 months, these symptoms remained in a significant proportion of patients. Patients with eye symptoms or a history of psychiatric conditions were more likely to be affected with brain-related symptoms.

Paper III was a longitudinal case-control study of 65 premenopausal women with newly diagnosed untreated GD designed to investigate the effect of GD on hippocampal volumes. At onset of GD, hippocampus and amygdala volumes of the patients were smaller compared to controls. These brain structures became larger with treatment and, after 15 months, only the left amygdala remained smaller than in controls. At inclusion, there was an inverse correlation between thyroid-stimulating hormone receptor antibody (TRAb) and the volumes of both amygdalae and the right hippocampus. There were also inverse correlations between TRAb and free triiodothyronine recovery and the recovery of most of the four assessed brain volumes.

GD is a condition where a minority of patients can hope for a long-lasting restored thyroid function. A large proportion of GD patients do not feel recovered after 8 years. Mental fatigue is an important concept for understanding the brain-derived symptoms in GD. In summary, the studies demonstrate that Graves' hyperthyroidism has unexpected long-term consequences for many patients, provide extensive new data on the serious and chronic nature of GD, and show for the first time that GD is accompanied by reversible brain changes.

Keywords: Graves' disease, Quality of life, Depression, Anxiety, Mental fatigue, Magnetic resonance imaging, Volumetry, Hippocampus, Amygdala

SAMMANFATTNING PÅ SVENSKA

Giftstruma är en autoimmun sjukdom med många ansikten. I många fall är den en relativt enkel sjukdom att hantera medan den hos andra är betydligt mer komplicerad. Flera tidigare studier har visat att det är ganska vanligt med bestående problem efter behandling för giftstruma men det saknas fortfarande mycket kunskap om hur problemen ser ut, hur vanliga de är, hur man ska kunna förutspå dem, förhindra dem och behandla dem, och slutligen vad de beror på.

Denna avhandling är baserad på tre studier där den första följde upp en stor grupp giftstrumapatienter som insjuknade 2003–2005 och kartlade hur deras sjukdom yttrat sig under de år som gått. Den andra och tredje studien inkluderade 65 kvinnor som nyinsjuknade i giftstruma under åren 2011–2018. De jämfördes med 65 matchade, friska kontroller. Alla deltagare fyllde i frågeformulär, genomgick intervjuer, deltog i tester och genomförde slutligen en magnetkameraundersökning av hjärnan. Efter 15 månaders behandling gjorde alla patienterna och många av kontrollerna om samma undersökningar.

De viktigaste resultaten från den första studien var att betydligt färre än man tidigare trott återfår en normal funktion i sin sköldkörtel efter behandling med tabletter. Nästan var fjärde patient som behandlas med tabletter kräver hormonbehandling så småningom. Var fjärde patient kände sig inte helt återställd 8 år efter insjuknandet.

Resultaten i den andra studien pekar mot att de som har ögonsymptom eller tidigare psykiatrisk problematik ofta drabbas hårdare av mentala symtom vid insjuknandet i giftstruma. Studien visar också att hjärntrötthet är vanlig vid GD, inte bara vid insjuknandet utan också efter avslutad behandling.

I den tredje studien undersöktes volymen av hjärnstrukturerna hippocampus och amygdala och båda dessa strukturer var mindre hos patienterna än hos kontrollerna. Med behandling så blev volymerna större igen, och efter 15 månader var alla utom vänster amygdala lika i storlek jämfört med kontrollerna. Vi fann dock ingen direkt koppling mellan de mentala symtomen och hormonnivåerna, antikropparna eller förändringarna i hjärnan.

Sammanfattningsvis visar resultaten att det är viktigt att betrakta giftstruma som en sjukdom med många kroniska karakteristika. Även om vi inte såg något direkt samband mellan de mentala symtomen och förändringarna i hjärnan är det tydligt att giftstruma är en sjukdom med vanligt förekommande, långvariga psykiska konsekvenser och med en betydande effekt på hjärnans struktur.

LIST OF PAPERS

This thesis is based on the following studies, which are referred to in the text by their Roman numerals.

I. Sjölin G*, **Holmberg M***, Törring O, Byström K, Khamisi S, de Laval D, Abraham-Nordling M, Calissendorff J, Lantz M, Hallengren B. Filipsson Nyström H, Wallin, G. The long-term outcome of treatment for Graves' hyperthyroidism. *Thyroid* 2019;29(11):1545-57.

*Gabriel Sjölin and Mats Holmberg contributed equally

II. **Holmberg M**, Malmgren H, Berglund PF, Johansson B, Klasson N, Skau S, Filipsson Nyström H. Predictors for mental symptoms in Graves' disease. Manuscript under preparation.

III. **Holmberg M**, Malmgren H, Berglund PF, Bunketorp-Käll L, Heckemann RA, Johansson B, Klasson N, Olsson E, Skau S, Starck G, Filipsson Nyström, H. A longitudinal study of medial temporal lobe volumes in Graves' disease. Manuscript under preparation.

CONTENTS

ABBREVIATIONS

ATD	Antithyroid drug
CA	Cornu ammonis
CAS	Clinical Activity Score
CPRS-S-A	Comprehensive Psychopathological Rating Scale for Affective Disorders
DSM	Diagnostic and Statistical Manual of Mental Disorders
D2	Deiodinase type 2
D3	Deiodinase type 3
fNIRS	Functional near infrared spectroscopy
fT3	Free triiodothyronine
fT4	Free thyroxine
GD	Graves' disease
GREAT	Graves' Recurrent Events After Therapy
ICV	Intracranial volume
IQR	Interquartile range
MAPER	Multi-atlas propagation with enhanced registration
MFS	Mental Fatigue Scale
MR	Magnetic resonance
MRI	Magnetic resonance imaging
MTL	Medial temporal lobe
PGWB	Psychological General Well-Being
QoL	Quality of life
RAI	Radioactive iodine
rT3	Reverse triiodothyronine
SCID-I	Structured Clinical Interview for DSM-IV Axis I Disorders
SD	Standard deviation
T3	Triiodothyronine
T4	Thyroxine
TAO	Thyroid-associated ophthalmopathy
ThyPRO	Thyroid-specific Patient-Reported Outcome
TPO	Thyroperoxidase
TPOAb	Thyroperoxidase antibody
TR	Thyroid hormone receptor
TRAb	Thyroid-stimulating hormone receptor antibody
TRH	Thyrotropin-releasing hormone
TSH	Thyroid-stimulating hormone (thyrotropin)
TSHR	Thyroid-stimulating hormone receptor

1 INTRODUCTION

This thesis is about the long-term consequences of Graves' disease (GD) with special focus on long-standing mental symptoms.

GD owes its name to the Irish surgeon Robert James Graves who, in 1835, described several females affected by heart palpitations, goiter, and psychiatric symptoms, which were at the time described as "hysteria".[1] One of his major points was that the medical profession should consider that "hysteria" may be a mental reaction to an underlying physical event. Since then, the mental symptoms in GD have been subject to debate.

GD is a common disorder and, in the endocrinology out-patient clinic, GD patients are seen on a daily basis. Restoring thyroid hormone levels is usually a rather simple procedure for the clinician. At the same time, most endocrinologists share the experience of seeing GD patients with disabling mental symptoms despite having achieved euthyroidism for months or even years.

The medical profession rarely has much to offer these patients suffering from long-standing mental symptoms. This is frustrating for both the patients who are left alone trying to find ways to cope with disabling symptoms and also for the clinicians whose contribution to the patients' well-being is limited to compassion.

So, to search for ways to better understand, predict, treat, and possibly prevent long-standing mental symptoms in GD seems an obvious task. This was also the driving force for myself to participate in this scientific project, being a clinician working in daily practice among patients with hormonal disturbances, an endocrinologist interested in psychiatry, and an individual fascinated by human behavior.

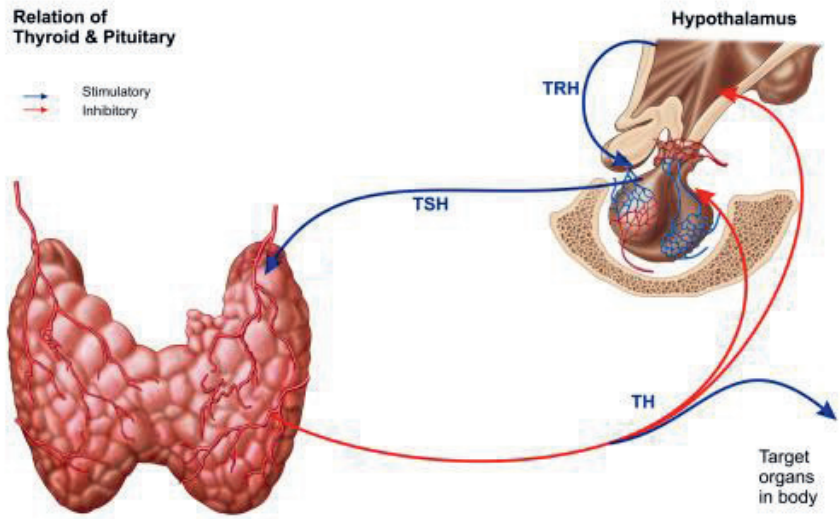
1.1 GRAVES' DISEASE

In Sweden, more than 2000 individuals develop GD every year.[2] The overwhelming majority of these patients are women[3] with a peak incidence between 30 and 50 years of age.[3 4]

Unlike what happens with most autoimmune diseases, the antibodies in GD do not destroy their target organ. Instead, they bind to thyroid-stimulating hormone (TSH) receptors (TSHR) in the thyroid gland and thereby trigger an

overproduction of the thyroid hormones, thyroxine (T4) and triiodothyronine (T3).[3] This results in elevated circulating levels of T4 and T3, and diminished circulating levels of TSH via negative feed-back on the production of TSH. The combination of elevated levels of TSH receptor antibody (TRAb) and free T4 (fT4) together with suppressed levels of TSH constitutes the diagnosis of GD. In the rare TRAb-negative cases, a diffuse uptake on technetium scintigraphy confirms diagnosis. The normal thyroid axis is shown in Figure 1.

Figure 1. The thyroid axis. *Abbreviations:* TH, thyroid hormone; TRH, thyrotropin-releasing hormone; TSH, thyroid-stimulating hormone [with the courtesy of Getty Images].



The high thyroid hormone levels affect most organ systems in the body through genomic pathways, although there are also non-genomic effects of T3 on membrane receptors[5] (e.g., on mitochondria and on the receptors of sympathetic nerves[6]). Thyroid hormones are metabolically active[7] and the hyperthyroid state is associated with somatic symptoms including weight loss, tremor, heat intolerance, palpitations, and mental symptoms such as fatigue, altered mood, insomnia, and anxiety.[3 8] GD can also be complicated by extrathyroidal autoimmune manifestations, such as thyroid-associated ophthalmopathy (TAO) with an incidence of 4.9%,[9] thyroid dermopathy with an incidence of 1%–4%,[10] and acropachy (swelling of the hands and clubbing of the fingers).[3]

1.1.1 TREATMENT OF GRAVES' DISEASE

Treatment of GD is, in most cases, a choice between three dominant modalities: antithyroid drugs (ATDs), surgery, or radioactive iodine (RAI). These represent two primarily different approaches: thyroid-sparing or ablative treatments. Thyroid-sparing treatment, in the form of ATDs, has the major advantage of offering the patient a chance of cure, i.e., a return to a normally functioning thyroid gland. The ablative treatments (surgery and RAI) have the major advantage of a low rate of disease recurrence.

The choice of treatment is based on the patient's clinical state, the availability of treatments, the patient's wishes, and regional treatment practice. The clinician needs to carefully consider short- and long-term efficacy, and the risks, of each modality when discussing therapeutic choices with patients to adequately prepare them for a well-informed choice.[8 11 12] Internationally, there are large variations in which method is preferred: RAI dominates in some countries[11] and ATD in others.[13] Sweden, in line with most of Europe,[14] is somewhere in between.

2.1.1 ANTITHYROID DRUG TREATMENT

The most commonly used treatment method in Sweden is ATD therapy. By tradition, daily oral administration of one of the two thionamides (propylthiouracil or thiamazole) is used, with thiamazole as a first choice. After about 2 weeks, the thionamide is combined with levothyroxine in order to avoid hypothyroidism. This treatment modality is called block-and-replace treatment. After 12–18 months of treatment, the level of TRAb is checked and, if it is low, the treatment is ended. An alternative to a block-and-replace regimen is low-dose monotherapy with an ATD. In the hyperthyroid state, beta-blockers such as propranolol are sometimes used to control tachycardia, tremor, and palpitations. ATDs exert their effect by inhibition of thyroperoxidase (TPO)-catalyzed synthesis of thyroid hormones[15] and the levels of circulating thyroid hormones are usually normalized within weeks of treatment, while antibody levels persist much longer.[16] The mechanism behind the declining TRAb level is not fully understood: a direct immunosuppressive effect of the ATD[15] as well as a consequence of lower thyroid hormone levels[17] have been suggested.

There are two main disadvantages with ATD treatment: side effects and the risk of relapses. Some side effects are common and mild, such as pruritus and rash, and some are rare but serious, such as agranulocytosis and liver failure.[18-20] The risk of disease relapse after a completed ATD course is about 50%.[20] Besides having to live through another period of distressing symptoms,

recurrent disease is also troublesome for the patient in that prolonged exposure to hyperthyroidism may cause adverse health outcomes.[3 16 21]

According to the 2016 American Thyroid Association guidelines,[22] clinical situations that favor medical treatment are mild disease, small goiters, negative or low-titer TRAb, contraindications or high risk with surgery or RAI, patients with moderate to severe TAO, and patients who need rapid biochemical disease control.

3.1.1 RADIOACTIVE IODINE TREATMENT

In Sweden, RAI is the second most common initial treatment for GD. The rationale for this treatment is that orally administered RAI will accumulate in the thyroid. Thus, high doses of radioactivity are delivered to the thyroid gland, while the rest of the body is exposed to low doses. The most common practice in Sweden is to administer enough of the iodine isotope I^{131} as a single dose to deliver sufficient activity to the thyroid to achieve hypothyroidism. The activity of I^{131} is calculated individually in most patients using thyroid volume, percentage 24-hour I^{131} uptake, and I^{131} half-life to achieve an absorbed dose of 120 Gy to the thyroid gland.

The rate of relapse after RAI treatment is low, but a disadvantage is that the treatment causes hypothyroidism, which requires life-long levothyroxine therapy. Other disadvantages with RAI treatment are the increased risks of TAO and cardiovascular morbidity as well as impaired long-term quality of life (QoL).[23-32]

4.1.1 SURGICAL TREATMENT

Surgical treatment is the least common initial treatment for GD in Sweden. The surgical option becomes more attractive with the occurrence of side effects from other treatment choices and with disease relapse. Surgery is generally performed as total thyroidectomy, meaning that the whole (or almost the whole) thyroid gland is removed. Indications for thyroidectomy include a disturbingly enlarged thyroid gland, moderate to severe TAO, pregnancy or breastfeeding, and persistent or relapsing hyperthyroidism after ATD or RAI treatment.[33]

The risk of recurrence with surgical treatment is low but at the expense of life-long thyroid hormone replacement.[34] Surgery also exposes the patient to surgical risks such as damage to the recurrent nerves or surgically acquired hypoparathyroidism.[35]

5.1.1 FOLLOW-UP AFTER TREATMENT

In Sweden, GD is usually diagnosed in primary care. After diagnosis, the patient is referred to an endocrine out-patient clinic at a hospital, where the hyperthyroidism is treated. When euthyroidism has been reached and the patient's antithyroid treatment has been ended, the patient is referred back to primary care for follow-up. If the patient later relapses in GD, a new hospital referral is sent for evaluation and, eventually, a new period of treatment. This method of organizing the treatment of GD gives endocrinologists considerable experience in how to treat hyperthyroidism, while general practitioners generally deal with the long-term consequences of having been treated for GD.

As a contrast, hypothyroidism is usually both diagnosed and treated at a primary care level without involvement of endocrinologists.

6.1.1 MENTAL SYMPTOMS IN GRAVES' DISEASE

Numerous publications have described the associations between GD and affective symptoms.[16 36-49] The most commonly reported symptoms are depression and anxiety, but bipolar disorder has also been suggested as being more common among GD patients.[50-52]

Most of the studies focus on symptoms present in the hyperthyroid state. However, experiences of stress-related life events[53 54] as well as psychiatric symptoms[55] have been reported to be more prevalent even before the onset of GD. Treatment usually improves the symptoms,[36 56] but several long-term follow-up studies have found a higher prevalence of depression and anxiety in treated GD patients compared to the general population.[50 57-59] The mechanisms behind the affective symptoms are to a large extent unknown, but psychological factors including stressful life events,[54 60 61] childhood abuse and neglect,[62] personality traits,[63] and experience of recurrent GD[63] as well as biological factors such as excess thyroid hormones levels[49 64] and elevated levels of TPO antibody (TPOAb)[50 65 66] have all been suggested. Most studies that have investigated the relationship between thyroid hormones and mental symptoms in GD have found no such association and, in population-based investigations in healthy individuals, no relationship between thyroid hormone levels and depression or anxiety has been found.[67 68]

Another prevalent symptom in GD is cognitive dysfunction.[36 38 43 45 69 70] However, results from cognitive testing of GD patients are conflicting.[36 46 69-71]

The typical cognitive complaints by patients with GD resemble what is seen in astheno-emotional disorder,[72 73] which is characterized by a difficulty in

maintaining attention and a tendency to develop fatigue during cognitively demanding tasks accompanied by memory difficulties, irritability, and emotional lability. Astheno-emotional disorder, often called mental fatigue,[74] is present in many pathological brain conditions including traumatic brain injury, brain tumors, stroke,[75] and endocrine diseases such as Cushing's syndrome. Possible alternative labels for astheno-emotional disorder/mental fatigue are mild cognitive disorder[76] and mild neurocognitive disorder.[77] However, definitions of the latter two conditions do not emphasize the characteristic mental fatigability in astheno-emotional disorder/mental fatigue and, importantly, do not include emotional signs and symptoms.

Although euthyroidism is usually achieved rapidly under treatment, recovery of well-being is delayed for months in many patients and some never regain full pre-morbid health.[37-39 58] Several prospective studies in GD patients have indicated a short-term negative impact on QoL.[37 38 40 78] One year after treatment approximately 20% of GD patients still experience reduced QoL[40] and many are not fully recovered after 3 years.[16] In a randomized, prospective follow-up 17–21 years after treatment, GD patients still had decreased vital and mental QoL compared to the general population.[31]

Sick leave is more prevalent in GD patients 1 year after diagnosis compared to the general population.[16 79] Patients with hyperthyroidism have a doubled risk of being absent from work during the first year after starting treatment and fewer return to work even during subsequent years.[80] In a follow-up 6 years after GD diagnosis, more than 30% of the patients reported complete or partial incapacity for work.[57] The specific symptoms that cause this prolonged inability to return to full-time work are currently unknown.

1.2 THYROID HORMONES AND THE BRAIN

Thyroid hormones exert a profound influence on the fetal brain.[81] Historically, cretins were a sad proof of the deleterious effect of thyroid hormone deficiency on brain growth and function.[82] Since then, several studies have investigated more subtle associations between fetal thyroid hormone levels and brain function later in life.[83-85] In the adult, thyroid hormones also play an important role in the maintenance of brain function.[86]

Thyroid hormones, both T4 and T3, are actively transported into the brain by several different transport proteins.[87] Inside the brain cells, T4 is converted to the active thyroid hormone T3 by deiodinase type 2 (D2),[88] while both T4 and T3 are inactivated to reverse T3 (rT3) by deiodinase type 3 (D3). The activity of the D2 and D3 enzymes regulate intracellular T3 levels and thereby protect

the brain from hypo- and hyperthyroidism. As an example, D2 activity is downregulated in hyperthyroidism.[89]

T3 binds to nuclear thyroid hormone receptors (TRs) distributed throughout most brain regions, with the highest levels found in the olfactory bulbs, the hippocampus*, and the cerebellar cortex.[90-92] Two distinct genes, TRα and TRβ, encode several receptor isoforms with specific functions that define the tissue-specific action of thyroid hormone in the brain.[93-95]

Besides TR activation and non-genomic pathways of thyroid hormones, the suppressed levels of hypothalamic thyrotropin-releasing hormone (TRH) may also be involved in the pathogenesis of GD.[96]

1.3 THYROID ANTIBODIES AND THE BRAIN

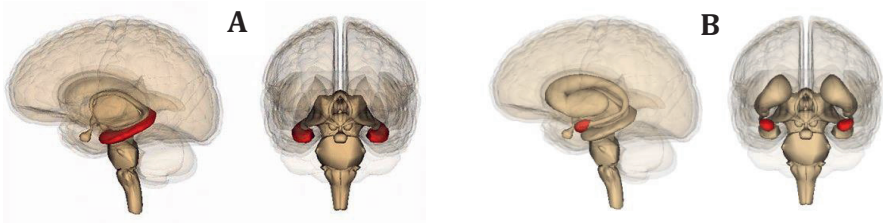
The antibodies directed at the TSH receptor, TRAb, are part of the pathogenesis in GD.[3] TSH receptors are seen in the amygdala, cingulate gyrus, frontal cortex, hippocampus, hypothalamus, and thalamus,[97] and their action is different from that of TRs.[98] Moreover, the production of antibodies in GD is not limited to TRAb. Autoantibodies against thyroglobulin,[99] TPO,[99] D2,[100] insulin-like growth factor 1 receptor,[101] angiotensin II receptor type 1,[102] and β1-adrenergic and M2 muscarinic receptors[103 104] have all been described as elevated in GD, some with proposed pathogenic effects.[102 103] Hence, brain pathology related to GD may be caused by intracerebral hyperthyroidism, by thyroid-related antibodies, and/or, more directly, by the autoimmunity *per se*.[105] The fact that depression and anxiety appear more prevalent in GD than in toxic nodular goiter[37] gives some support to the latter two alternatives.

1.4 THE MEDIAL TEMPORAL LOBE

The medial temporal lobe is a system of anatomically and functionally related regions deep in the brain.[106] Centrally located in this system are the two paired structures: the hippocampus and amygdala (Figure 2).

*Hippocampi and amygdalae are often more adequate terms than hippocampus and amygdala. These forms are, however, rarely used in publications. For simplicity, hippocampus and amygdala will be used from now on except when grammar necessitates the plural forms. The same holds for other bilateral structures.

Figure 2. Location of hippocampus (**A**) and amygdala (**B**) in the medial temporal lobe [with the courtesy of Wikimedia commons/Life science databases].

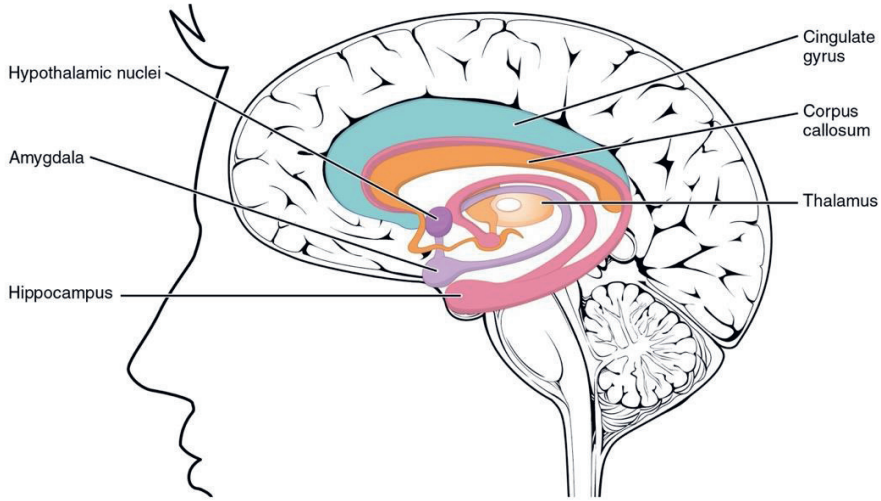


Although these neighboring structures are highly interconnected, the hippocampus is believed to be more involved in declarative memory and orientation, while the amygdala is more involved in emotional processing.[107] The term hippocampus is derived from Greek mythology. The brain structure vaguely resembles the sea horse that drew Poseidon's wagon. When referring to the hippocampus in literature, there is still no consensus among authors regarding exactly what parts are included. The hippocampus can be subdivided into a number of different areas and frequently used subfield terms include the dentate gyrus and the cornu ammonis (CA) regions CA1 to CA4.[108] Amygdala means almond in Greek. Just as the name states, the external shape is less complex than the shape of the hippocampus. Common subdivisions of the amygdala include the basolateral complex, the cortical, medial, and central nuclei, and the intercalated cell clusters.[109] A schematic description of some of the connections of the hippocampus and amygdala to the rest of the brain is presented in Figure 3.

Since 1998, when direct evidence for adult neurogenesis in humans first was presented,[110] hippocampal dynamics have been the focus of numerous studies.[111-116] Although the question whether there is significant adult neurogenesis remains a matter of controversy,[117] this specific region of the brain has been suggested as having an annual turnover of neurons of almost 2%.[118]

Hippocampal gray matter volume alterations have been demonstrated in connection with normal physiological variability like exercise,[119] the menstrual cycle,[120] and aging[121] as well as in pathological conditions like Alzheimer's disease,[122] mild cognitive impairment,[123] post-traumatic stress disorder,[124] major depressive disorder,[125] [126] Parkinson's disease,[127] and Cushing's syndrome,[112-116] all of which are illnesses with mental symptoms. These wide-ranging results are not easy to interpret. For Alzheimer's disease, there is a

Figure 3. Schematic description of some of the hippocampal and amygdalar connections to the rest of the brain [with the courtesy of Wikimedia commons/OpenStax College].



plausible direct coupling to the destructive pathogenetic process, which often involves the hippocampus at an early stage. For the other illnesses, there is no such plausible direct coupling and the hippocampal change is probably part of a more complex causal process that involves other brain regions. In several disorders such as depression and Cushing's syndrome, the hippocampal disturbance may be transient. For depression and Cushing's syndrome, the response to treatment includes both recovery of hippocampal volumes and improvement of mental symptoms.[112 128-130] It is not known whether these changes involve reductions and later restorations of neuron numbers, or whether they are due to other mechanisms.

In some of these illnesses, reduced amygdalar volume was observed.[116 131-133] The same considerations about possible causal mechanisms should be applied to these findings.

Studies on hyperthyroidism and hippocampal volume are rare. One study has reported reduced hippocampal volume in hyperthyroid GD patients,[134] while one large population-based study found no association between hyperthyroidism and hippocampal volume.[135]

2 KNOWLEDGE GAPS

Long-lasting mental symptoms in GD exist. This is evident from both clinical experience and previous research. Questions that remain to be answered are therefore more related to factors influencing these symptoms, their clinical management, and the pathophysiological mechanisms behind them.

The choice of treatment modality is naturally an important component to investigate in the search for factors that influence outcome. Short-term treatment results, side effects, and efficacy of different treatments are rather well studied. However, we lack information on how the choice of treatment or treatments influences the long-term possibility of preserved thyroid function and what influences the patients' feeling of recovery after many years.

In the clinical management of the mental symptoms, we lack predictors for long-lasting problems. It is largely unclear whether the patients' psychiatric diagnoses or biomarkers such as the levels of thyroid hormones and thyroid antibodies are correlated with outcome.

Regarding mechanistic explanations, one previous cross-sectional study has suggested a reduced hippocampal volume in untreated GD.[134] This finding has to be confirmed and supplemented with longitudinal information on whether treatment of GD is accompanied by reversal of the structural changes in the brain. Information is also lacking on whether the structure of the closely related amygdala is affected.

Finally, it is unknown to what extent the brain's morphological changes over time are related to hormonal and immunological markers of GD and to changes in mental symptoms.

Very relevant, but outside the scope of this thesis, is that we do not know how to ameliorate the patients' long-standing symptoms.

3 GENERAL AIM

The general aim of this thesis is to study what impact GD and its therapies have on patient outcome at disease onset, after 15 months of treatment, and several years later, long after the restoration of hyperthyroidism.

3.1 SPECIFIC AIMS

7.1.1 PAPER I

To describe the long-term consequences of the initial choice of treatment modality for GD. Specifically: a) the efficacy of each therapeutic modality in preventing disease relapse; b) the patterns of initial and subsequent utilization of therapies during the whole period; c) the efficacy of each therapy to normalize thyroid function and the proportion of patients maintaining euthyroidism without levothyroxine replacement; and d) the patients' long-term feeling of recovery.

8.1.1 PAPER II

To systematically characterize the mental symptoms in women with GD both in hyperthyroidism and in euthyroidism after 15 months of treatment. Specifically: a) the prevalence and degree of depression, anxiety, and mental fatigue; b) the prevalence and degree of cognitive dysfunction; c) the prevalence and degree of reduced QoL; and d) the relationship between mental symptoms and biochemical measurements, previous psychiatric conditions, and eye symptoms.

9.1.1 PAPER III

To explore neural correlates to mental symptoms in women with GD. Specifically: to investigate a) the volume of medial temporal lobe (MTL) structures in hyperthyroidism[†]; b) the relationship between MTL volumes, symptoms, and biochemical measurements; c) the MTL volume response to euthyroidism by treatment; and d) the relationship between the degree of recovery in mental symptoms and the degree of recovery in MTL volumes.

[†]The initial formulation of the aim for this study only mentioned the hippocampus and the power calculation was therefore based on hippocampal volume.

4 STUDY DESIGN

4.1 STUDY DESIGN – PAPER I

This was a longitudinal, observational study with a combination of prospective and retrospective design. It was prospective in the sense that a predefined cohort of 2916 patients from a previous incidence study of hyperthyroidism was contacted after 6–10 years for follow-up. It was retrospective in the sense that the questionnaire used partly referred to past events. Data were collected from a questionnaire and the patients' medical files. Invitation to participate was sent by ordinary mail.

CONSIDERATIONS: STUDY DESIGN

The study design for Paper I caused problems typically associated with retrospective studies such as recall bias. Ideally, the study should have been planned *a priori* in 2003–2005 as a prospective, long-term follow-up. That would also have secured the availability of important information such as laboratory values and a clear definition of disease remission. Another consequence of the retrospective design was problems associated with flaws in the earlier study such as incorrect identification numbers for over 50 patients. A common problem with this type of follow-up study is that relatively high numbers of patients are lost to follow-up.

4.2 STUDY DESIGN – PAPERS II AND III

Papers II and III were based on the same prospective, case-controlled, observational study with a combined cross-sectional and longitudinal design. Participants were assessed by questionnaires, endocrine evaluation including eye assessment, a psychiatric interview, cognitive tests, laboratory measurements, brain morphology assessments with magnetic resonance imaging (MRI), and brain function assessment with functional near infrared spectroscopy (fNIRS), all within 2 weeks from the start of treatment for hypothyroidism and in euthyroidism after 15 months of treatment. Patients were invited to participate when referred to the Thyroid Unit at Sahlgrenska University Hospital, Gothenburg, Sweden.

CONSIDERATIONS: STUDY DESIGN

The study design of Papers II and III has many strengths such as the use of an extensive assessment battery and the use of MRI investigators blinded to the participant category. The main limitation of the design was the time that passed

between the start of treatment and inclusion. Ideally, the patients should have been investigated in the untreated phase of GD before the administration of beta-blockers and ATDs. This was unfortunately not possible for practical and ethical reasons. Another consideration of the design was that the investigations were quite time consuming for the participants, which might have led to a selection bias towards persons who can offer their time to medical investigations. It is impossible to predict what effect this might have had on the results but, in order to control for selection bias regarding psychiatric morbidity, an enquiry was sent to the Swedish National Board of Health regarding previous psychiatric diagnoses in patients who accepted participation compared to those who declined. These results are not yet available. The extensive test battery is also likely responsible, together with some participants' negative experiences of MRI, for the rather high drop-out rate that actually made us expand the number of patients included. The inclusion of participants has taken almost 9 years. This could have caused problems with a change of the participating researchers performing the investigations, magnetic resonance (MR) upgrades, etc. Luckily, there were no such changes that had appreciable effects on the study during this period.

5 PARTICIPANTS AND METHODS – PAPER I

5.1 PARTICIPANTS – PAPER I

From the original 2003–2005 cohort, 940 patients (children, emigrated, deceased, and patients with toxic nodular goiter) were excluded. The remaining 1976 patients with GD received an invitation letter and, of these, a total of 1186 agreed to participate, while 245 patients declined participation. Of the invited GD patients, 545 did not reply and were defined as "non-repliers" (Table 1).

CONSIDERATIONS: PARTICIPANTS

The patients included in the original study were all the patients with *de novo* GD at seven endocrine clinics in southern Sweden in 2003–2005. These seven clinics covered approximately 40% of the Swedish population. Sweden is an iodine-sufficient country[136] with small regional differences in GD treatment. Despite this, regional differences in incidence were found in the original study and, therefore, it cannot be excluded that the sample was not representative for the GD population in Sweden. The response rate of 60% was not optimal but is in line with other follow-up studies of this size and duration.[137 138]

5.2 METHODS – PAPER I

The invited patients received a 68-item clinical patient questionnaire developed for this specific study (see Appendix in Paper I). It includes questions regarding autoimmunity, smoking habits, demographic data, menopause, comorbidity, symptoms, presence and treatment of TAO, initial treatment of hyperthyroidism, recurrence of hyperthyroidism, its successive and final treatment, and a question whether the patient felt recovered from the thyroid disease. Two additional questionnaires regarding general (Short Form Health Survey-36 [SF-36]) and thyroid-specific (Thyroid-specific Patient-Reported Outcome [ThyPRO]) QoL were also sent to the patients. The results from these questionnaires are not, however, included in this thesis. The patients' replies to the clinical questionnaire were scanned automatically.

Table 1. Baseline data at inclusion in the original 2003–2005 cohort for the 1186 participants with Graves' disease included in the 2011–2013 follow-up and for the 545 individuals who did not reply to the study invitation (non-repliers). The number of patients included in the analysis of each baseline parameter is noted in square brackets when data were missing.

	Included (n=1186)	Non-repliers (n=545)	p-value
Gender			
Women, *n* (%)	973 (82.0)	428 (78.5)	0.084
Men, *n* (%)	213 (18.0)	117 (21.5)	
Age			
Years, mean (SD)	46.9 (14.4)	42.5 (14.2)	<0.001
Women, years, mean (SD)	46.5 (14.4)	42.4 (14.2)	<0.001
Men, years, mean (SD)	48.6 (14.2)	42.9 (14.2)	<0.001
Country of origin	[n=1184]	[n=411]	
Sweden, *n* (%)	911 (76.9)	270 (65.7)	<0.001
Europe outside Sweden, *n* (%)	148 (12.5)	35 (8.5)	0.029
Outside Europe, *n* (%)	125 (10.6)	106 (25.8)	<0.001
Smoking at diagnosis	[n=1174]	[n=483]	
Non-smoker, *n* (%)	523 (44.5)	210 (43.5)	0.690
Former smoker, *n* (%)	312 (26.6)	111 (23.0)	0.127
Smoking at diagnosis, *n* (%)	339 (28.9)	162 (33.5)	0.060
Eye symptoms	[n=1185]	[n=544]	
Eye symptoms, *n* (%)	250 (21.1)	102 (18.8)	0.260
Initial treatment	[n=1186]	[n=364]	
ATDs, *n* (%)	774 (65.3)	295 (81.0)	<0.001
RAI, *n* (%)	324 (27.3)	63 (17.3)	<0.001
Surgery, *n* (%)	54 (4.6)	3 (0.8)	0.001
Conservative, *n* (%)	34 (2.9)	3 (0.8)	0.026

Abbreviations: ATDs, antithyroid drugs; RAI, radioactive iodine; SD, standard deviation.

In addition to the clinical questionnaire, the included patients' medical files were reviewed with respect to their treatment modality, the order of treatments, and recurrence of disease after each treatment course. If an answer to a question in the questionnaire necessitated verification, this was also sought in the records.

The medical records of the non-repliers were reviewed with respect to their treatment modality, the order of treatments, and recurrence of disease after each treatment course.

CONSIDERATIONS: METHODS

If one wants be able to rely on a questionnaire, its construct validity (does it measure what is says it will) and reliability (will it measure it again) should be tested. This was not performed with the questionnaire especially designed for Paper I. On the other hand, the variables in the questionnaire were mostly of a concrete nature with questions such as "Which treatment did you receive?" In order to minimize automatic reading problems, 5% of the answers were manually validated. To further secure data and minimize selection bias, the records of both the included patients and the non-repliers were reviewed. Since the information in these files was not produced with a scientific follow-up in mind, the file reviews were likewise accompanied by uncertainties.

6 PARTICIPANTS AND METHODS – PAPERS II AND III

6.1 PARTICIPANTS – PAPERS II AND III

The patients included in Papers II and III were 65 consecutive adult, premenopausal women with newly diagnosed GD with fT4 ≥50 pmol/L (reference range: 12–22 pmol/L) and/or T3 ≥6.0 nmol/L (reference range: 1.3–3.1 nmol/L) plus positive TRAb. Inclusion and exclusion criteria are presented in Paper II, page 7.

Age-matched female control subjects were selected from a random sample of the general population in Gothenburg provided by the Swedish Tax Registry. They were invited by mail and were eligible for the trial if they were also matched for educational level and smoking habits to an individual patient.

After being referred to the Thyroid Unit at Sahlgrenska University Hospital, eligible patients were approached either before or at the first clinical visit when ATD treatment was prescribed. After starting ATD, an inclusion visit was scheduled within 2 weeks.

CONSIDERATIONS: PARTICIPANTS

The inclusion of only women in Papers II and III is considered as both a strength and a weakness. Men and women have a different prevalence of psychiatric conditions and only including women makes the results more valid and relevant for the group that is most afflicted both by GD, and by depression and anxiety. The clinical presentation and course of the disease may differ between men and women. It is therefore not possible to say whether these results are generalizable to male GD patients. The included patients also represent a group with severe hyperthyroidism, which has the advantage of minimizing the risk of misdiagnosis and the possibility of investigating them in a state of hyperthyroidism, but the disadvantage that we cannot say if these results are generalizable to GD patients with less severe hyperthyroidism.

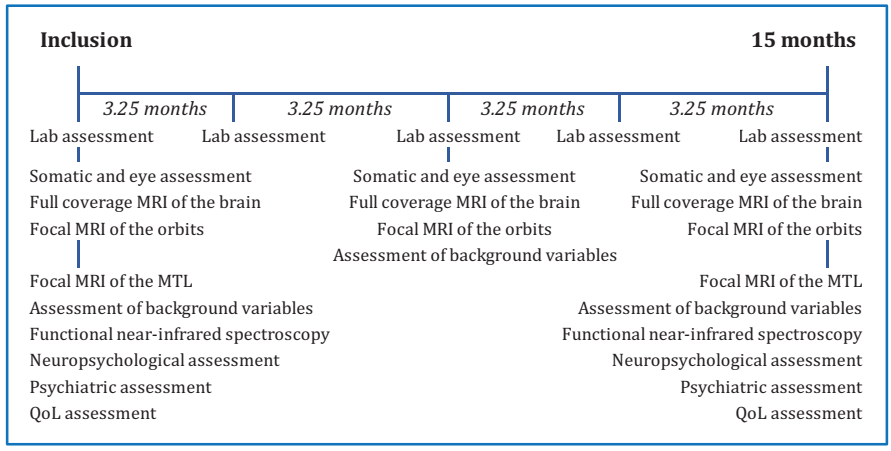
The exclusion of patients with TAO needing glucocorticoid treatment (see the exclusion criteria) and the near total absence of patients with RAI treatment also affect generalizability. These two groups have, however, previously been described as having lower QoL compared to patients without TAO and patients treated with ATD or surgery[31 32 139-142]. Thus, leaving them out should lower the risk of over-estimating a negative effect of GD on QoL.

There is a possibility of selection bias in both directions depending on psychiatric burden. In order to investigate this, a request has been sent to the Swedish National Board of Health to compare the included patients with those that declined inclusion.

6.2 METHODS – PAPERS II AND III

Participants underwent a multimodal assessment battery including demographics, smoking status, previous head trauma, height and weight, electrocardiogram, history of eye symptoms, TAO activity and severity, thyroid hormone and antibody assessment, a psychiatric assessment of past and current illness, a mental fatigue questionnaire, a neuropsychological cognitive assessment, and two QoL questionnaires. The whole battery was used at inclusion and after 15 months (Figure 4). On the same occasions, the participants also underwent MRI for structural MTL volumetry and orbital muscle evaluation. The latter was also done at 7.5 months. A subset of participants underwent fNIRS at inclusion and after 15 months.

Figure 4. Study time line and full assessment battery for Papers II and III. *Abbreviations:* Lab, laboratory; MRI, magnetic resonance imaging; MTL, medial temporal lobe; QoL, quality of life.



10.1.1 STRUCTURED CLINICAL INTERVIEW FOR DSM-IV AXIS I DISORDERS

The Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I) is a semi-structured interview based on open questions. It should be handled by an

interviewer with experience of both SCID-I and psychiatric diagnosis, and is usually done by a psychiatrist. First published in 1996, the SCID-I interview has become an extensively used method to capture previous and present psychiatric disorders.[143] Regardless of whether the interviewee has ever been in contact with the healthcare system or has previously been diagnosed with a psychiatric disease, the interview gathers the participant's symptoms that may or may not fulfil the diagnostic criteria of a DSM-IV axis disorder. The interview usually takes about 1 hour. The Swedish translation from 1998 was used.[144]

CONSIDERATIONS: SCID-I

SCID-I was included with the aim of investigating the impact of previous and present psychiatric conditions on GD symptomatology.[55] Previous psychiatric conditions refer to disorders that started before the onset of GD, including psychiatric conditions that are no longer present and those that are chronic. SCID-I is a well-used and validated method. In the present study, the diagnostic criteria were not always followed literally, since other signs and symptoms sometimes made it clear to the psychiatrist that an exception had to be made. The vast majority of our interviews were made by the same psychiatrist but, in all, there were three interviewers. In order to optimize assessment consistency, all interview protocols were carefully inspected by the main psychiatrist. It is nevertheless not possible to exclude a rater bias. The use of SCID-I may also be criticized since it is too time consuming to be a useful tool in clinical practice. Finally, the interviewer was not blinded to participant category.

An alternative would have been to use registry data, but the latter is less sensitive since patients do not always seek health care for psychiatric problems.

11.1.1 COMPREHENSIVE PSYCHOPATHOLOGICAL RATING SCALE FOR AFFECTIVE SYNDROMES

The Comprehensive Psychopathological Rating Scale for Affective Syndromes (CPRS-S-A) is a self-rating scale developed in Sweden, but widely used internationally.[145] The scale was constructed in 1994 from the original interview-based CPRS as a tool to evaluate symptomatic change in psychiatric disorders. It consists of 19 questions covering depression and anxiety and, on average, takes 15–30 minutes for the participant to complete. A higher score refers to more severe symptoms.

12.1.1 MENTAL FATIGUE SCALE

The Mental Fatigue Scale (MFS) questionnaire was originally developed to capture the mental fatigue that is prevalent following acquired brain injury regardless of the cause.[75] This mental fatigue had previously been described as astheno-emotional disorder.[72 146] The scale consists of 15 questions covering symptoms typical of astheno-emotional disorder like rapid drain of mental energy upon mental activity, impaired concentration, long recovery time, and diurnal variation. Furthermore, associated symptoms like tearfulness, irritability, stress intolerance, trouble with memory, sleep problems, and sensitivity to or intolerance of light and loud noise are covered. The MFS has been validated for traumatic brain injury and stroke patients.[147] A cut-off at 10.5 was used, with a total score ≥10.5 indicating deviation from normality.[147]

13.1.1 PSYCHOLOGICAL GENERAL WELL-BEING

The Psychological General Well-Bing (PGWB) questionnaire is a validated 22-question QoL measure widely used across different conditions. Six dimensions of well-being are assessed within six subscales: anxiety, depressed mood, positive well-being, self-control, general health, and vitality. A higher score is indicative of a higher level of psychological well-being. Results can be presented as a raw or normalized score (raw score/max score × 100). The validated Swedish version was used.[148] No previous publication on GD has used PGWB.

14.1.1 THYROID-SPECIFIC PATIENT-REPORTED OUTCOME

Thyroid-specific Patient-Reported Outcome (ThyPRO) is a self-assessment questionnaire developed in Denmark and designed to capture QoL specifically related to thyroid disease.[149] The questionnaire measures QoL with 13 scales covering physical symptoms, mental symptoms, and the impact of thyroid disease on participation (i.e., social and daily life) plus one question on the effect of thyroid disease on QoL. ThyPRO consists of 84 items, and each scale is rated using scores ranging from 0 to 100, with increasing scores indicating decreasing QoL (i.e., more symptoms or greater impact of disease). The Swedish translation developed in 2011 was used.[150]

CONSIDERATIONS: QUESTIONNAIRES

The choice of questionnaires was based on the ambition to cover psychiatric symptoms (depression and anxiety), mental fatigue, and thyroid-specific and general QoL. All the questionnaires used in the study are validated instruments. They were also handled by experienced persons who were used to presenting

such instruments in a standardized way. However, only ThyPRO has been previously validated in GD patients. This can possibly make the results less valid. As an example, CPRS-S-A was developed to evaluate treatment outcome and not to diagnose depression and anxiety. Therefore, this questionnaire has no cut-offs and only the longitudinal comparisons were in line with its original purpose.

The choice to include MFS was based on long clinical experience with patients who complain of symptoms that resemble mental fatigue. There are many previous publications that have evaluated "tiredness" that remains after treatment for GD. Tiredness is, however, a much more general term than mental fatigue.

PGWB was not included in the study until after patient 13. None of the persons handling the questionnaires were blinded to participant category.

15.1.1 NEUROPSYCHOLOGICAL ASSESSMENT

The neuropsychological examination included assessment of processing speed, attention, working memory, verbal fluency, and reading speed. The tests were administered in a standardized sequence as described in Paper II, page 12.

These tests were added to evaluate whether the cognitive complaints that are so common in GD patients would be captured in a more objective test situation.

CONSIDERATIONS: NEUROPSYCHOLOGICAL TESTS

The tests used in Papers II and III are all very well-known, validated tests. The testing was performed by an experienced tester in the vast majority of cases but there were three testers overall. The tester was not blinded to participant category. Neither of these facts would be expected to influence the results since all tests have highly standardized administration procedures. It is more difficult to evaluate to what extent training influenced the results when the tests were repeated after 15 months.

16.1.1 BIOCHEMICAL MEASUREMENTS

Blood samples were analyzed at the Department of Clinical Chemistry at Sahlgrenska University Hospital, Gothenburg, Sweden for serum T4, fT4, T3, free T3 (fT3), and TSH. TPOAb was analyzed by electrochemiluminescence immunoassay (Roche Elecsys®ECL, Roche Diagnostics International AG, Rotkreuz, Switzerland) and total TRAb by radioreceptor analysis using a Brahms Kryptor automated assay system (Thermo Fisher Scientific, Waltham, MA).

CONSIDERATIONS: BIOCHEMICAL MEASUREMENTS

The laboratory tests in Papers II and III were all performed in the same laboratory using validated methods. The main consideration is about the choice of TRAb as a marker for autoimmunity that causes hyperthyroidism. The TRAb level does not reflect the degree of stimulation of the thyroid gland since only a fraction of TRAb stimulates, while other fractions block and some are neutral. There are several methods available that measure the stimulating fraction of TRAb. Thyroid-stimulating immunoglobulins were included in the study but these levels were measured in batches of frozen serum and were unfortunately not complete at the time of this thesis.

17.1.1 MAGNETIC RESONANCE IMAGING

The resolution of the MR images depends on the strength of the magnetic field, measured in Tesla. In Paper III, a 3-Tesla MR scanner (Philips Gyroscan Achieva 3T, Philips, Best, The Netherlands) was used at the Department of Radiology, Sahlgrenska University Hospital, Gothenburg, Sweden. Structural images of 0.7 × 0.7 × 1 mm voxel size were acquired axially using a 3D T1-weighted Fast Field Echo sequence. Sagittal T2w sections (3–5 mm slices) were additionally acquired.

Scan settings in the MR acquisition determine how tissues appear in the resulting images. In Paper III, a T1-weighted scan sequence was used as this sequence is thought to better separate gray and white matter.

CONSIDERATIONS: MRI

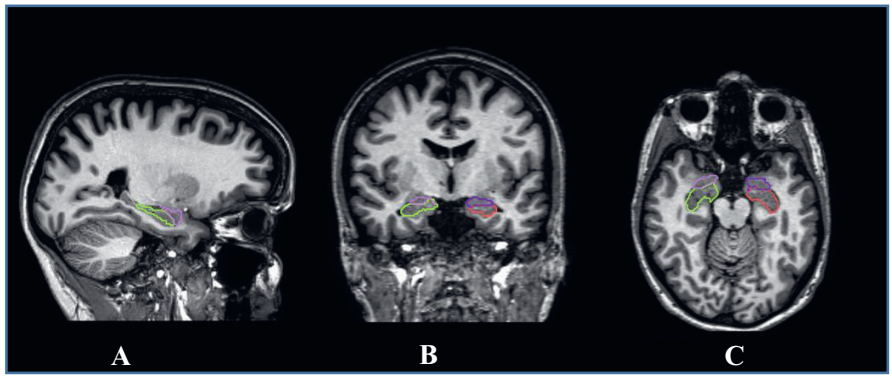
Mental symptoms are undoubtedly a consequence of altered function of the brain and there are many imaging techniques available to describe brain function. The choice of structural MRI, nevertheless, has the advantage of being clinically available. The initial choice to focus on MTL can be questioned. Cognition is complex and, as such, involves widely distributed functional systems and several brain regions. A functional tool, fNIRS, which measures changes in blood oxygenation in the neocortex, was therefore added to the assessment battery after patient 37 in Papers II and III.

18.1.1 SEGMENTATION

Segmentation (drawing and separation) of brain regions can be performed manually or automatically, where the former is considered the golden standard while the latter is less time-consuming, less rater dependent, and essentially stable if repeated on the same data. Many automatic segmentation methods are available and some of these use atlas databases as reference.[151] In Paper III, the

MR images were segmented by multi-atlas propagation with enhanced registration (MAPER) using the Hammersmith atlas database (Figure 5).[152-154]

Figure 5. Magnetic resonance brain images segmented with MAPER. The outer limits of hippocampus (green and red) and amygdala (purple) are marked in the sagittal plane (**A**), coronal plane (**B**), and horizontal plane (**C**).



CONSIDERATIONS: SEGMENTATION

Manual volumetry is still the golden standard when it comes to hippocampal volumetry. The main advantages of an automatic method like MAPER are that several other brain areas can be measured at the same time and that the high test-retest reliability makes it easy to combine results from analyses performed at different times. A subset of images were also measured with manual volumetry and the preliminary results compared to the MAPER results. To avoid rater drift over time the researcher should conduct the manual volumetry as one batch. Therefore, this work will be completed after the inclusion of the last participant.

19.1.1 INTRACRANIAL VOLUME NORMALIZATION

The amygdala and hippocampus volumes as well as head size display large variability between individuals. In order to be able to compare segmentation results between patients and controls while eliminating part of the influence from head size, the volumes of the individual MTL structures were normalized by intracranial volume.[155]

7 STATISTICAL ANALYSES

7.1 PAPER I

Demographic data was analyzed using Pearson's chi-square for categorical variables and unpaired Student's t-test for continuous variables. Odds ratios were analyzed using linear and multivariate regression. Statistical analyses were performed using IBM SPSS Statistics 22.0.0.1 64-bit edition (SPSS Institute, Chicago, IL, USA). Statistical significance was set at $p<0.05$.

7.2 PAPERS II AND III

According to a pre-study sample size calculation, 40 patients and 40 controls should be enough to detect a hippocampal volume group mean difference of 10% with 80% power and a significance level of 0.05. After also considering the risk of drop-outs, 65 patients were included in Papers II and III.

For the statistical analysis, fT3, fT4, and TRAb values exceeding the detection range in either direction were set to the detection limit. When fewer than 64 patients were included in a sub-analysis, the number included was noted. Spearman's rank correlations were used when associations between variables were investigated. Patients and controls were compared using the Mann-Whitney U-test with the exception of MTL volumes, which were normally distributed and were therefore compared using Student's t-test. Intraindividual longitudinal comparisons were performed using the Wilcoxon matched pairs test with the exception of longitudinal MTL comparisons, which were performed using the paired Student's t-test. In the analyses of mental symptomatology and MTL volumes, we performed analyses both in the group with normal TSH at 15 months and in the group with TSH outside the reference range to rule out the possible influence of inadequate thyroid therapy. All statistical calculations were made using Tibco Statistica 13.2 (Tibco Software Inc., Palo Alto, CA).

8 ETHICAL APPROVAL

8.1 PAPER I

Ethical approval was granted by the Regional Ethics Review Board in Uppsala, Sweden (reference no. 2012/035 approved 2012 April 4, 2012 with addition 2012/035/2 approved February 24, 2015). The study was performed according to the Declaration of Helsinki. The evaluation of the medical files in the non-replier group was approved in an amendment.

8.2 PAPERS II AND III

Ethical approval was granted by the Regional Ethics Review Board in Gothenburg, Sweden (reference no. 190-10 approved May 21, 2010 with additions T912-11 approved November 11, 2011, T854-12 approved November 11, 2012, T955-14 approved February 26, 2015, T249-16 approved March 15, 2016, and T1247-18 approved December 28, 2018). The studies were conducted in accordance with the Declaration of Helsinki.

9 RESULTS – PAPER I

9.1 BASELINE CHARACTERISTICS AT DIAGNOSIS

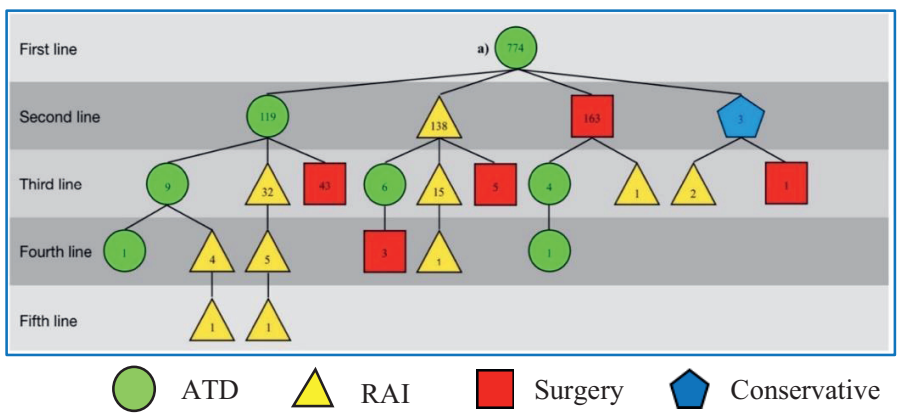
Mean (±SD) follow-up time was 8.0 ± 0.9 years and mean age at diagnosis was 46.9 years. The female-to-male ratio was 4.6:1. Information on TAO was available for 1185 of the 1186 of the patients and 21.1% were considered as having TAO at diagnosis according to an experienced endocrinologist and/or ophthalmologist.[2 156] For additional baseline data, see Paper I, Table 1.

9.2 LONG-TERM TREATMENT OUTCOME

20.1.1 ATDs AS FIRST-LINE TREATMENT

ATDs were used as the first treatment choice in 774 patients (Figure 6). Of these, 596 patients completed a 12–18 month ATD course and, of the latter, 351 (58.9%) were in remission at follow-up. Relapse after the first treatment period was noted in 245 (41.1%) patients and 119 of these were re-treated with ATD, which resulted in 35/119 (29.4%) still in remission at follow-up. The second ATD course resulted in significantly fewer long-term remissions than the first course (p<0.001).

Figure 6. Flowchart illustrating the patients who initially received a) ATD treatment and their subsequent treatment choices of ATD, RAI, surgery, or conservative treatment. Second- to fifth-line treatments include new ATD treatment and changed treatment due to recurrent disease, side effects, failure of previous treatment, or the patient's wishes. *Abbreviations:* ATD, antithyroid drug; RAI, radioactive iodine [reproduced with permission from *Thyroid,* Mary Ann Liebert, Inc., New Rochelle, NY].

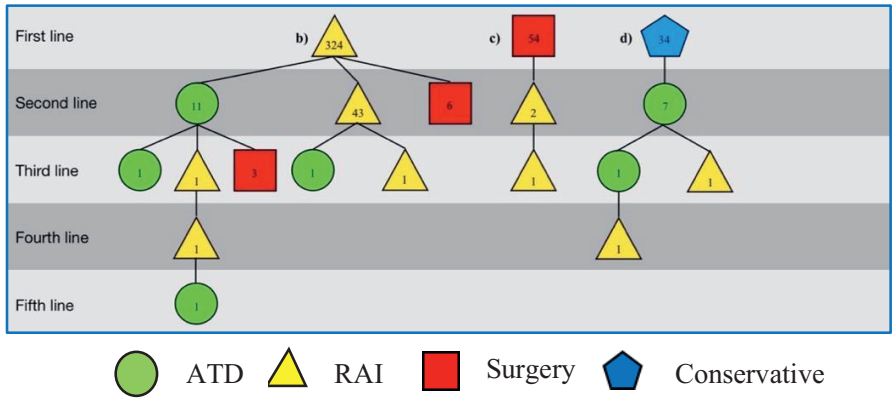


Of the patients who started ATD, 178 patients changed to ablative treatment during the first treatment period. The reasons were: adverse effects, new development of TAO or worsening of an existing TAO, persistent high TRAb levels, or that the patient desired changing therapy. Of the 119 patients who were selected for a second course of ATD after relapse, 63% ultimately received ablative treatment (Figure 6). Taken together, a patient who was selected for ATD as first treatment and subsequently had a possible additional ATD treatment had a 49.7% risk of undergoing ablative treatment during the 6–10 years follow-up.

21.1.1 RAI, SURGERY, OR CONSERVATIVE TREATMENT AS FIRST-LINE TREATMENT

RAI was used as initial treatment in 324 patients, of whom 264 were still in remission at follow-up after one treatment course (Figure 7).

Figure 7. Flowchart illustrating the patients who initially received b) RAI, c) surgery, or d) conservative treatment and their subsequent treatment choices of ATD, RAI, or surgery. Second- to fifth-line treatments include retreatment or change of treatment due to recurrent disease or change of treatment due to side effects, failure of previous treatment, or the patient's wishes. *Abbreviations:* ATD, antithyroid drug; RAI, radioactive iodine [reproduced with permission from *Thyroid,* Mary Ann Liebert, Inc. New Rochelle, NY].



In total, 42.6% of patients (505/1186) were treated with RAI at some time during the whole study period (Figures 6 and 7).

Surgery was chosen as first-line treatment in 54 patients (Figure 7c). Among these, 52 patients were still in remission at the end of the study period.

Considering all GD patients included in the study, 23.4% (278/1186) had undergone thyroidectomy at follow-up (Figures 6 and 7).

22.1.1 PRESERVED THYROID FUNCTION

Of all 774 patients who started ATD treatment, 325 (40.3%) had preserved thyroid function without needing levothyroxine substitution 6–10 years after the diagnosis of GD. Of the 396 patients who remained in remission after one or more ATD courses, 77% did not need substitution with levothyroxine after 6–10 years. RAI treatment at any time resulted in preserved thyroid function in 16.8% and surgery in 4.7% of patients. The latter group underwent subtotal thyroidectomy.

23.1.1 PATIENTS' EVALUATION OF THEIR TREATMENT

At follow-up, 25.3% of the patients reported that they were not yet fully recovered. The most common specific causes for this were tiredness (25.3%), eye-related problems (13.5%), and levothyroxine substitution problems (12.5%). Additional unspecified treatment problems were reported by 10.4% and concerns about living with a life-long disease with the risk of recurrence by another 10.4%. The proportion of patients who felt fully recovered did not differ between the treatment groups, but was higher in the patients without levothyroxine treatment compared to those receiving levothyroxine ($p<0.001$).

24.1.1 COMPARISON TO THE NON-REPLIER GROUP

Detailed baseline data are shown in Paper I, Table 1. Compared to the non-replier group, the included patients were more often born in Sweden or Europe and older by a mean of 4.4 years. They were less frequently treated with ATDs and more often treated with RAI or surgery. The proportion of patients who were in complete remission from first-line treatment was higher in the included group (58.5%, 694/1186) than among the non-repliers (41.8%, 228/545) [$p<0.001$].

10 RESULTS – PAPERS II AND III

Sixty-five patients and 61 controls were included in Papers II and III. At 15 months, 57 of the 65 patients remained but only 48 completed the MR investigations. Immediately after inclusion, one patient was excluded due to menopause. Two patients could not complete the baseline MR investigation due to claustrophobia. The main reasons for the remaining drop-outs were lack of time and not wanting to repeat the MRI investigation.

10.1 STATUS AT DIAGNOSIS

At referral, before inclusion, 92% of the GD patients presented with fT4 ≥50 pmol/L, while 58% had a total T3 ≥6.0 nmol/L. Elevated TRAb was observed in 97% and elevated TPOAb in 67% of the patients. In the two patients with negative TRAb, the diagnosis of GD was confirmed by a diffuse uptake on technetium scintigraphy.

10.2 STATUS AT INCLUSION, INITIAL TREATMENT

At the first clinical visit, thiamazole treatment was initiated in all patients. Before the inclusion visit, two patients had changed to propylthiouracil. At inclusion, 77.4% of the patients were treated with beta-blockers. At inclusion there were no differences between the patients and controls except that the mean body weight was 5.3 kg lower (p<0.05) in patients. At baseline, 46.9% of the patients had CAS 0 and 53.1% CAS 1, 2, or 3. Disease characteristics of the included patients are described in Table 2.

10.3 STATUS AT 15 MONTHS

At 15 months, 32% of the patients had undergone thyroidectomy and four patients had finished antithyroid treatment. There was a mean bodyweight gain of 3.2 kg (p<0.01) during treatment. At 15 months 73.5% had CAS 0 and 26.5% CAS 1, 2 or 3. During the 15 months of follow-up, CAS increased in 4 patients and decreased in 19 patients.

Table 2. Disease characteristics of the 64 patients with newly diagnosed Graves' disease with hyperthyroidism at inclusion.

	Median (IQR)	*n* (%)
Time from blood test before diagnosis to inclusion (days)	13 (11.0–20.0)	
Time from start of ATD to blood test at inclusion (days)	8 (6–12)	
Duration of symptoms before diagnosis (months)	4 (3–6.5)	
Patients treated with beta-blockers		48 (77.4)
Patients treated with ATD		64 (100)
Propylthiouracil		2 (3.1)
Thiamazole		62 (96.9)
TAO Clinical Activity Score		
≥1		34 (53.1)
>3		0

Abbreviations: ATD, antithyroid drugs; IQR, interquartile range.

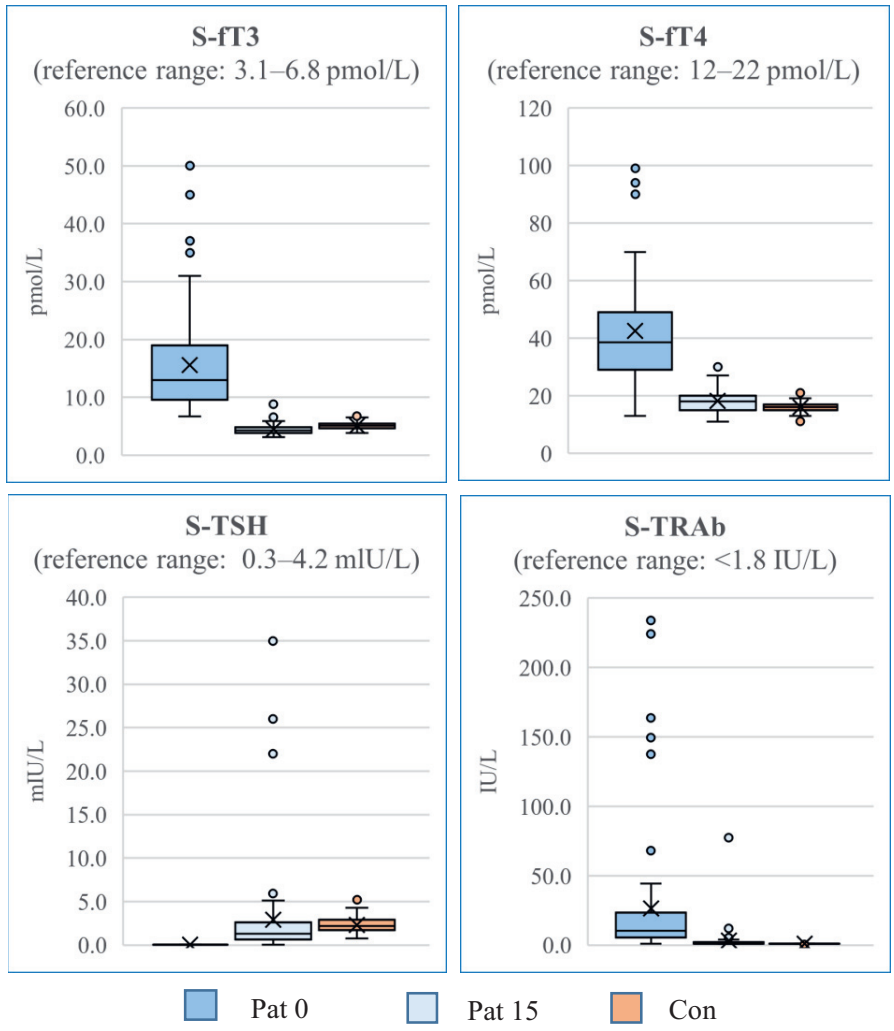
10.4 BIOCHEMICAL EVALUATION

The levels of fT3, fT4, TSH, and TRAb in patients and controls at inclusion and in patients at the 15-month follow-up are presented in Figure 8. At inclusion, the levels of fT3, fT4, and TRAb were higher ($p<0.001$) and TSH lower ($p<0.001$) in the patients compared to controls. At 15 months, thyroid hormones had returned to normal levels for most patients, but fT4 was still higher, fT3 lower, and TSH lower in patients compared to controls ($p<0.001$).

At follow-up, TSH levels were below normal in 9/57 patients, above normal in 5/57 patients, and within the normal range in 43/57 patients.

TRAb levels improved with treatment ($p<0.001$), but were still higher at follow-up compared to the controls and had not returned to the normal range in 40% of the patients.

Figure 8. Box and whisker plots of serum free triiodothyronine (S-fT3), serum free thyroxine (S-fT4), serum thyroid-stimulating hormone (S-TSH), and serum TSH receptor antibody (S-TRAb) in patients at inclusion (dark blue, Pat 0) and after 15 months' treatment (light blue, Pat 15), and matched controls (orange, Con). Numbers of participants, median values, interquartile ranges, and levels of significance for group comparisons are specified in Paper II, Figure 1.



10.5 MENTAL SYMPTOM SCORING

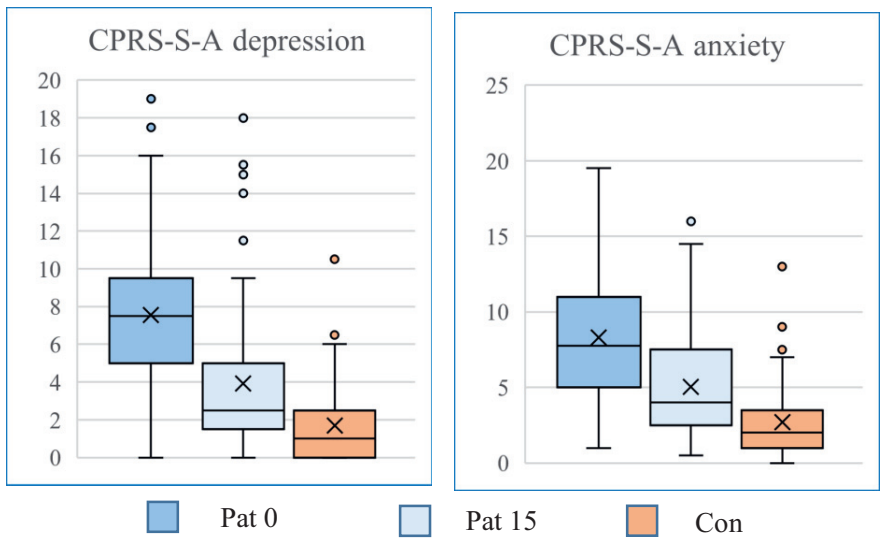
25.1.1 SCID

At inclusion, 64 patients and 61 controls were interviewed according to the SCID protocol. Of these participants, 52% of the patients and 54% of the controls were diagnosed with previous depression, anxiety, or both.

26.1.1 CPRS-S-A

CPRS-S-A was completed by 64 patients and 55 controls at inclusion and by 53 patients at follow-up. At inclusion, patients reported more depression and anxiety than controls (p<0.001) [Figure 9]. At follow-up, patients had improved (p<0.001), but still reported more depression and anxiety compared to controls (p<0.001) [Figure 9].

Figure 9. Box and whisker plots for comprehensive Psychopathological Rating Scale (CPRS-S-A) depression and anxiety scores in patients at inclusion (Pat 0, dark blue) and after 15 months' treatment (Pat 15, light blue), and matched controls (Con, orange). Median values, interquartile ranges and levels of significance for group comparisons are specified in Paper II, Figure 2. Higher scores represent more symptoms.



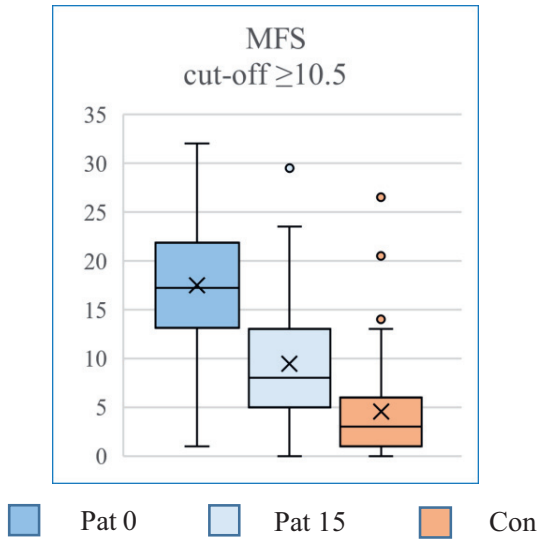
At inclusion, patients with a previous psychiatric condition reported more depression ($p<0.01$) and anxiety ($p<0.05$) than patients without such history. However, the patients who neither had a previous psychiatric condition nor any eye symptoms also had higher CPRS-S-A depression ($p<0.001$) and CPRS-S-A anxiety scores ($p<0.001$) than controls. At follow-up, anxiety scores were still higher in patients with a previous psychiatric condition ($p<0.05$) than in those without, whereas depression scores were not.

No correlations were found between the levels of fT3, fT4, TSH, or TRAb and CPRS-S-A scores either at inclusion or at follow-up.

27.1.1 MENTAL FATIGUE ASSESSMENT

The MFS was completed by 64 patients and 56 controls at inclusion and by 53 patients at follow-up. Result are reported in Figure 10. Patients experienced more mental fatigue than controls ($p<0.001$) at inclusion, and 91% of the patients scored above cut-off for healthy individuals (≥10.5) compared to 11% of the controls (chi^2-test, $p<0.001$). Patients improved with treatment ($p<0.001$), but at follow-up, patients still reported more mental fatigue than controls ($p<0.001$) and 38% still exceeded the cut-off.

Figure 10. Box and whisker plot for Mental Fatigue Scale (MFS) score at inclusion (Pat 0, dark blue) and after 15 months' treatment (Pat 15, light blue), and matched controls (Con, orange). Median values, interquartile ranges, and levels of significance for group comparisons are specified in Paper II, Figure 2.



Patients with previous psychiatric conditions had higher MFS scores than patients without such history ($p<0.001$). However, at follow-up, a history of psychiatric conditions did not influence the MFS score. For patients without psychiatric burden and without eye symptoms, the MFS score at inclusion was higher compared to controls ($p<0.001$). This difference remained at follow-up ($p<0.05$).

There were no correlations between MFS and fT3, fT4, or TRAb at inclusion. However, there was an inverse correlation between MFS and fT4 at 15 months ($\rho=-0.30$, $p<0.05$).

28.1.1 NEUROPSYCHOLOGICAL ASSESSMENT

Cognitive tests were performed by 64 patients and 56 controls at inclusion, and by 50 patients at follow-up. There were no significant differences in any of the tests between the patients and the controls at inclusion. At follow-up, the only difference was that patients were faster in Trail Making Test A ($p<0.01$) compared to controls. With treatment, patients improved in all tests except Digit span and reading speed (Table 3). There was no difference in cognitive test scoring between patients with previous psychiatric conditions and patients without such a history.

Table 3. Results of cognitive tests in patients at inclusion and 15 months. Comparisons are performed with Wilcoxon matched pairs test.

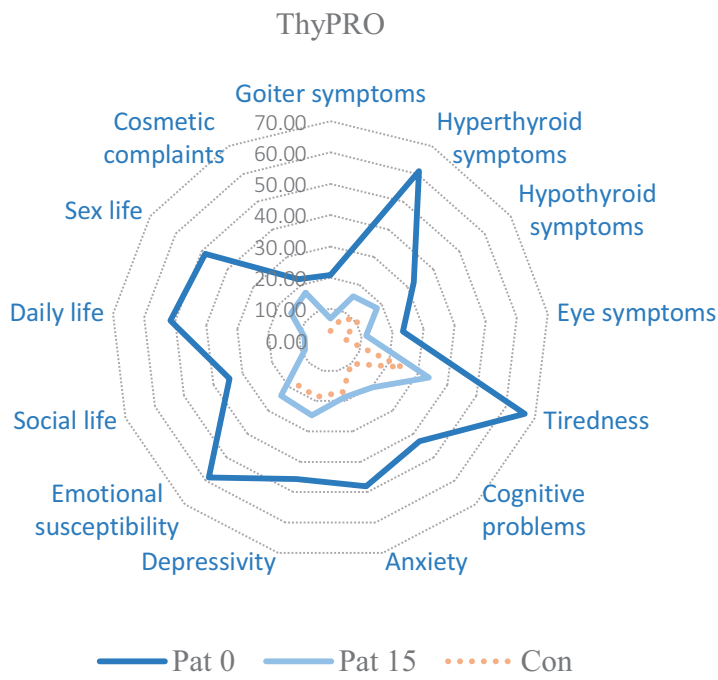
	Patients at inclusion		Patients at 15 months		
	N	Median (IQR)	*N*	Median (IQR)	*p*-value
Trail Making Test A	48	26.8 (20.9–34.0)	48	21.2 (18.0–27.2)	<0.001
Trail Making Test B	51	64.4 (52.5–85.0)	51	55.0 (45.2–69.0)	<0.01
Trail Making Test C	49	64.0 (47.6–81.8)	49	57.0 (40.0–67.0)	<0.05
Trail Making Test D	50	113.1 (87.0–139.5)	50	93.0 (75.0–112.6)	<0.05
Digit Symbol Coding	50	77.0 (67.0–89.0)	50	81.0 (72.0–90.0)	<0.01
Digit Span	42	15.0 (13.0–17.0)	42	15.0 (13.0–17.0)	0.34
F-A-S[a]	48	41.0 (32.0–49.0)	48	45.0 (38.0–55.0)	<0.01
Reading speed	50	3.0 (2.5–3.6)	50	3.2 (2.5–3.6)	0.82

[a]F-A-S is a test that measures the total number of words starting with F, A, and S that the test subject can produce in 1 minute.
Abbreviation: IQR, interquartile range.

29.1.1 ThyPRO

The ThyPRO questionnaire was completed by 63 patients and 54 controls at inclusion and by 54 patients at follow-up. Results are presented in Figure 11. Hyperthyroid patients had more symptoms than controls in all dimensions at inclusion ($p<0.001$). All dimensions, except cosmetic complaints, had improved after 15 months ($p<0.001$); however, patients still had more goiter, hyperthyroid, hypothyroid, and eye symptoms as well as more tiredness and cognitive problems after 15 months compared to controls ($p<0.05$ for all comparisons). The dimension cosmetic complaints at 15 months was inversely correlated to the patient's weight change between inclusion and 15 months ($\rho=-0.33$, $p<0.05$).

Figure 11. Mean Thyroid-related Patient-Reported Outcome (ThyPRO) questionnaire scores in patients at inclusion (Pat 0) and after 15 months (Pat 15), and in controls (Con). Items for four scales (impaired social life, impaired daily life, impaired sex life, and appearance) are asked with attribution to thyroid disease and cannot be answered by controls. Median values, interquartile ranges and levels of significance for group comparisons are specified in Paper II, page 18. Scale scores range from 0 to 100, with higher scores indicating worse health status.

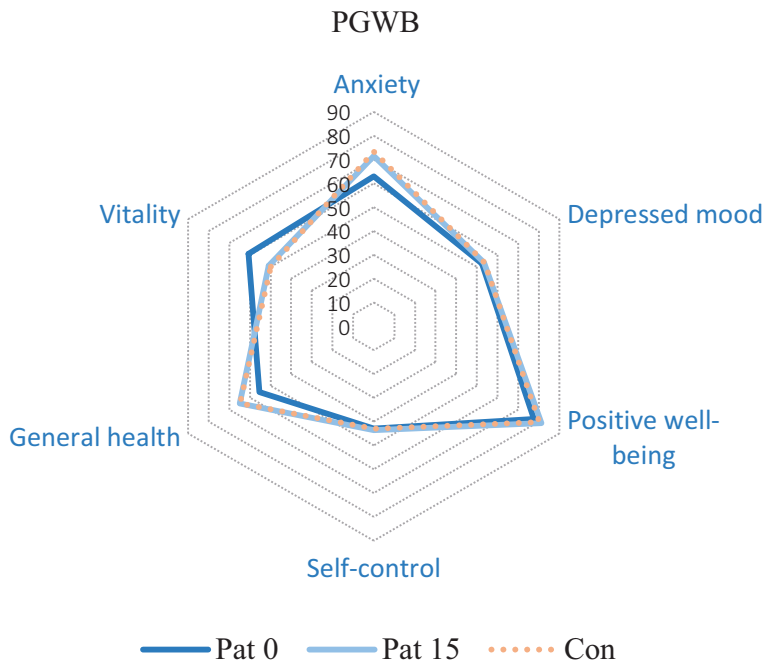


Patients with a previous psychiatric condition scored higher on ThyPRO (hyperthyroid and eye symptoms, cognitive problems, depressive symptoms, and impaired social life) than those without such a history ($p<0.05$ for each dimension). At follow-up, there was still a difference in depression and impaired daily life scores between patients with and without previous psychiatric conditions ($p<0.05$). For patients without psychiatric burden and without eye symptoms, the ThyPRO score in the dimension cognitive problems at follow-up was higher in patients than controls ($p<0.05$).

30.1.1 PGWB

PGWB was completed by 51 patients and 57 controls at inclusion and by 52 patients at follow-up. Results are presented in Figure 12. At inclusion, patients reported lower total score ($p<0.01$) as well as lower scores in the anxiety and general health dimensions ($p<0.001$), and higher score in the vitality dimension ($p<0.001$) compared to both the controls and to themselves at follow-up. At follow-up, there was no statistically significant difference in any dimension between the patients and controls.

Figure 12. Mean Psychological General Well-Being (PGWB) questionnaire scores in patients at inclusion (Pat 0) and after 15 months (Pat 15), and in controls (Con). Median values, interquartile ranges and levels of significance are specified in Paper II, page 19. Scale scores range from 0 to 100, with lower scores indicating worse health status.



At inclusion, patients with a previous psychiatric condition had lower score in the anxiety dimension compared to those without such a history ($p<0.05$).

31.1.1 ASSOCIATIONS BETWEEN EYE AND MENTAL SYMPTOMS

Patients with CAS ≥1 at inclusion had higher CPRS depression score ($p<0.05$), anxiety score ($p<0.001$), and MFS score ($p<0.01$) than those with CAS 0. At inclusion, patients with elevated CAS also had higher scores on the ThyPRO subscales for eye symptoms, depressivity, and impaired social life ($p<0.05$ for each subscale). At follow-up, patients with elevated CAS had higher scores in CPRS anxiety ($p<0.01$) and on the ThyPRO subscales for hyperthyroid symptoms, hypothyroid symptoms, eye symptoms, tiredness, depressivity, sex life, and cosmetic complaints ($p<0.05$).

At inclusion, patients with previous psychiatric conditions had significantly higher CAS score compared to patients without such a history (chi^2-test, $p<0.05$). This difference was not present at follow-up.

32.1.1 BRAIN MORPHOLOGY

At inclusion, the bilateral volumes of amygdala and hippocampus were smaller in patients compared to controls (mean difference in volumes in percent: left amygdala –10.4%, $p<0.001$; right amygdala –13.3%, $p<0.001$; left hippocampus –4.4%, $p<0.05$; and right hippocampus –4.8% $p<0.01$) [Figure 13 and Table 4]. With treatment, the patients' amygdala and hippocampus volumes increased significantly (mean difference in volumes in percent: left amygdala +6.7%, $p<0.001$; right amygdala +11.1%, $p<0.001$; left hippocampus +5.6%, $p<0.001$; and right hippocampus +5.8 %, $p<0.001$), but at follow-up, the left amygdala remained significantly smaller in patients than in controls (mean difference in volumes in percent: –5.0%, $p<0.05$) [Figure 13 and Table 4].

At follow-up, there was no difference in any of the volumes between the patients with a normal TSH and those with a TSH outside the reference range.

Figure 13. Box and whisker plots of amygdalar and hippocampal volumes in premenopausal women with newly diagnosed Graves' disease at inclusion (Pat 0, dark blue) and at 15 months (Pat 15, light blue), and in controls (Con, orange). Segmentation was performed with the MAPER automatic method. Intracranial volume (ICV)-normalized volumes were used.

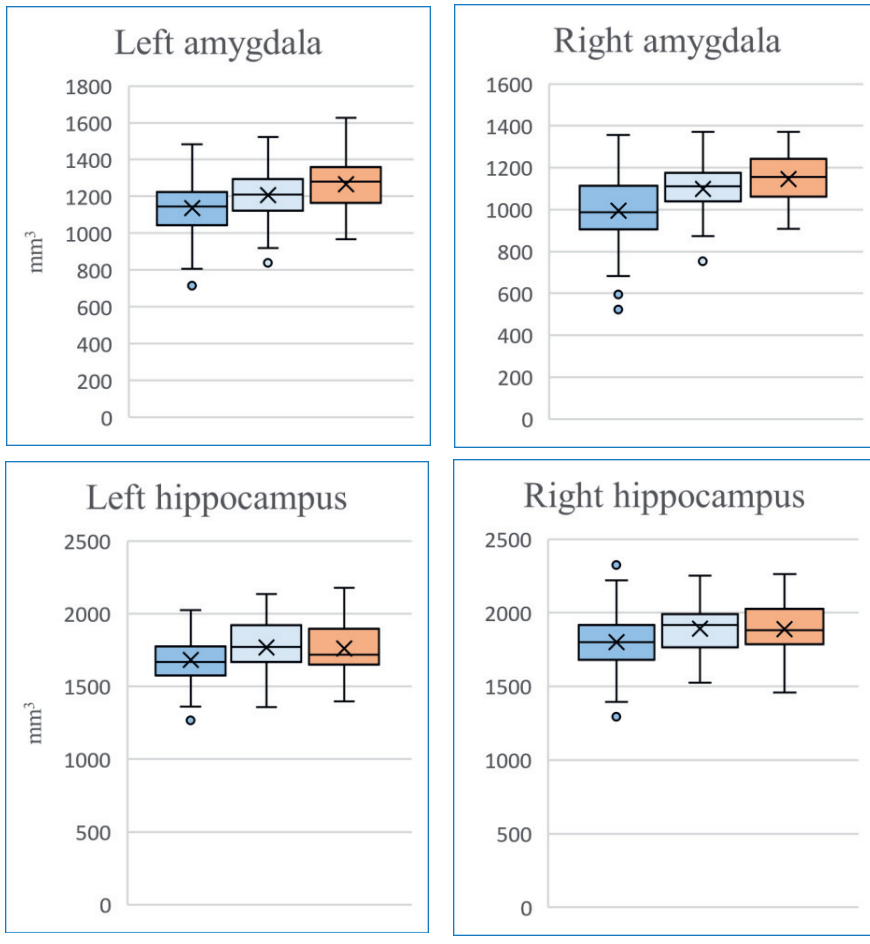


Table 4. Amygdala and hippocampus volumes in premenopausal women with newly diagnosed Graves' disease at inclusion and at 15 months, and in controls with statistical comparisons. Segmentation was performed with the MAPER automatic method. Intracranial volume-normalized volumes were used.

Region		*N*	Mean (±SD) ICV-normalized volume (mm^3)	*p*-value[a]		
				vs Pat 0	*vs* Pat 15	*vs* Con
Left amygdala	Pat 0	62	1135.2±148.4	–	<0.001	<0.001
	Pat	48	1207.1±132.2	<0.001	–	<0.05
	Con	56	1266.5±140.3	<0.001	<0.05	–
Right amygdala	Pat 0	62	994.1±161.5	–	<0.001	<0.001
	Pat	48	1099.9±126.5	<0.001	–	0.054
	Con	56	1146.1±114.9	<0.001	0.054	–
Left	Pat 0	62	1681.7±160.7	–	<0.001	<0.05
	Pat	48	1766.2±179.5	<0.001	–	0.858
	Con	56	1759.9±177.2	<0.05	0.858	–
Right	Pat 0	62	1799.0±193.7	–	<0.001	<0.01
	Pat	48	1892.6±173.5	<0.001	–	0.932
	Con	56	1889.6±176.8	<0.01	0.932	–

[a]Unpaired t-test.

Abbreviations: Con, control; Pat 0, patients at inclusion; Pat 15, patients at 15 months.

33.1.1 GD-RELATED FACTORS AND MTL VOLUMES

At inclusion, there was an inverse correlation between the TRAb level and the volumes of the right hippocampus (ρ=–0.29, p<0.05) and both amygdalae (left: ρ=–0.34, p<0.01, right: ρ=–0.35, p<0.01). There were no correlations between fT3, fT4, or TSH levels and the normalized amygdala and hippocampus volumes in patients.

The longitudinal results reported here are described in terms of delta correlations, i.e., correlations between the changes in different measures from inclusion to follow-up. There was an inverse correlation between the change in fT3 and the change in volume of both hippocampi (left: ρ=–0.40, p<0.01; right: ρ=–0.30, p<0.05) and the left amygdala (ρ=–0.43, p<0.01). There was also an inverse correlation between the change in TRAb and the change in the volumes of the left hippocampus (ρ=–0.30, p<0.05) and both amygdalae (left: ρ=–0.52, p<0.001; right: ρ=–0.37, p<0.05). There were no delta correlations between fT4 and MTL volumes.

34.1.1 MENTAL SYMPTOMS AND MTL VOLUMES

At inclusion, there was no significant difference in any of the MTL volumes between patients with previous psychiatric conditions and patients without. At inclusion, there was an inverse correlation between several ThyPRO dimensions (eye, depression, anxiety, and social life) and right hippocampus volume (data not shown). Apart from this there were no significant correlations between any of the other mental symptom scores and amygdala and hippocampus volumes at inclusion. At follow-up, there were no correlations between symptom scores and MTL volumes.

The duration of symptoms prior to inclusion did not correlate with MTL volumes. There were no correlations between the amygdala and hippocampus volume changes and the change in scores of any of the symptom questionnaires. Also, there was no significant correlations between the amygdala or hippocampus volume changes and the follow-up MFS or CPRS-S-A scores.

11 DISCUSSION

All the papers included in this thesis are about GD outcome. From a clinical perspective, seeing patients with newly diagnosed GD is different from handling recurrent disease, which is in turn different from dealing with its long-term consequences. I have therefore divided the discussion into three parts where each section represents one of these three clinical situations.

11.1 GD – THE HYPERTHYROID PHASE

One of the major concerns during clinical consultation with a patient with newly diagnosed GD is the choice of therapy. This requires the clinician to carefully consider both short- and long-term efficacy, and risks during discussion with the patient in order to achieve a well-informed choice.[8 11 12]

ATD treatment is the only modality that offers the patient a good chance of preserved thyroid function. Thus, informing patients of this option has a major impact on the patient's choice of treatment. In line with results from previous publications, the patients are usually informed that choosing ATD will result in remission in about 50% of cases. However, with a longer time perspective as in Paper I, and including the risk of change to ablative treatment and the risk of subsequent levothyroxine substitution, the actual percentage of patients with preserved thyroid function is only 40%.

According to the SCID interview, about half of patients and controls had a previous psychiatric condition of depression or anxiety, or both – a number much higher than the prevalence of depression and anxiety in the Swedish general population.[138] The high prevalence in our study may be a consequence of selection bias or a high inclusiveness of the SCID interview compared to other methods, but the possibility that GD patients have a higher prevalence of psychiatric morbidity compared to the general population even before the onset of GD has been previously suggested[55] and needs to be investigated further. When meeting a patient with a history of psychiatric conditions, the treating physician should be prepared to handle more severe symptomatology. Although a SCID interview is unlikely to be included in clinical care, other tools to assess previous psychiatric conditions, such as simpler questionnaires, may be considered.

The higher levels of occurrent depression and anxiety in hyperthyroid GD patients than in controls, as measured with CPRS-S-A, support previous

findings.[37 43 50 157] That previous psychiatric conditions predispose for depression and anxiety is also in line with a previous publication.[158]

The MFS has not been previously used in GD patients. Based on clinical experience, the scale was included to capture a specific syndrome, mental fatigue or astheno-emotional disorder which many patients report and which is difficult to capture with more general psychiatric or QoL questionnaires. The results show that mental fatigue is common in GD, even in the absence of previous psychiatric conditions and eye symptoms.

In line with previous publications, the ThyPRO score of cognitive complaints was higher in patients that in controls,[78 159] but the patients' scores on cognitive tests did not differ from those of controls or the general population.[160] A similar discrepancy between subjective and objective cognitive impairments has been noted before.[36] One explanation for this discrepancy might be that depressed patients have a tendency to underestimate their cognitive capacity. Another explanation might be that the "subjective" cognitive dysfunction is a record of everyday life, as opposed to the test situation. Mental fatigue entails that cognitive performance varies much with the situation.

The patients scored higher in the eye dimension of ThyPRO, which was expected and is in line with previous publications on untreated GD.[78 150 159 161] It is known that the presence of TAO has a negative effect on QoL[79 139 162] but the association between somewhat mild (CAS 1–3) eye symptoms and more depression and anxiety has not been previously reported.

There were no correlations between the patients' mental symptom scores and the levels of thyroid hormones or TRAb. One possible explanation for this lack of correlations is that only patients with severe hyperthyroidism were included and that biochemical differences within the very high range are not reflected in mental symptomatology. Another possibility is that the symptoms are caused by other brain related biochemical changes, e.g., diminished levels of TRH, that are not captured in standard GD laboratory measurements. Finally, the mental symptoms may be part of a psychological response to disease.

The observation that hyperthyroid GD patients have a reduced hippocampus volume compared to controls agrees with the only previous work published on hippocampal volumes in hyperthyroidism by Zhang and colleagues.[134] Unlike these previous authors, we also observed reductions in amygdalar volumes. In contrast to Zhang and colleague*s*, no cross-sectional correlations between the thyroid hormone levels or duration of symptoms and the MTL volumes were found. MTL volumes were also not influenced by previous psychiatric

conditions. The absence of correlations between mental symptom scores, thyroid hormone levels, and MTL volumes makes the clinical relevance of the volumetric findings unclear.

Although no correlations were seen between symptom scores and the level of TRAb in patients, the finding of inverse correlations between TRAb and most MTL volumes supports the idea that the volume reductions are related to thyroid autoimmunity.

11.2 GD – AFTER 15 MONTHS OF TREATMENT

After 15 months of treatment, euthyroidism has usually been established in patients for many months. For most of those who have changed to ablative treatment, antithyroid treatment is also history. In the cohort in Papers II and III, almost one third of the patients were thyroidectomized during the first 15 months of treatment, a much larger proportion compared to previous publications and to Paper I; this is likely due to the high thyroid hormone levels required for inclusion. For the patients still on ATD therapy, 15 months of treatment usually mean reduced levels of antibodies and that the end of antithyroid treatment is approaching.

The observation that mental symptoms improved with treatment but did not return to the level of healthy controls despite 15 months of treatment is in line with several previous publications.[78 159 161 163]

At 15 months, thyroid hormone and TRAb levels had improved for most patients, but there were still differences between the patients and the controls regarding fT3, fT4, TSH, and TRAb. This raises the question whether the remaining symptoms might be a consequence of inadequate therapy. Not least, the relationship between TSH and fT3 levels and mental symptoms has been previously debated and a meta-analysis on T3 substitution among depressed patients has reported positive results.[164] Most of the patients at 15 months were on levothyroxine treatment and the T4/T3 quotient has been previously described as higher in patients on levothyroxine compared to patients with a normalized thyroid function.[165 166] In this thesis, we found no indication that the levels of any of the thyroid hormones or TRAb affected the degree of the symptoms remaining in patients.

After 15 months of treatment, the MFS score remained higher in patients compared to controls even in the absence of previous psychiatric diagnoses or eye symptoms. Hence, mental fatigue is an important and independent factor for the mental deficit experienced by patients and should be further explored.

At 15 months, the patients still had higher scores for ThyPRO cognitive complaints compared to controls. Despite this, there were no differences in cognitive test results between patients and controls. However, most of the cognitive test results improved with treatment. This may be an effect of repeated tests, but a real improvement cannot be excluded. One should also consider that paired comparisons have a higher statistical power than unpaired ones.

Hippocampal and amygdalar volumes increased with treatment and, at 15 months, only the left amygdala remained smaller in patients compared to controls. The finding of dynamic alterations in amygdala and hippocampus in GD reflects a previously unknown level of brain involvement both in the development of the condition and in its response to treatment. These results constitute the best evidence to date that GD is indeed associated with brain changes. A similar recovery in hippocampal volume has been previously reported in Cushing's syndrome and depression.[128-130]

The recovery of three out of the four measured MTL regions was inversely correlated with the restoration of fT3 and TRAb levels. These findings suggest that GD *per se* affects MTL volumes and that the volumetric changes are possibly related to GD autoimmunity. This warrants further study of GD as a neuroimmunological disorder.

Although no correlations between the severity of mental symptoms and brain volumes were found, nor between the improvement of symptoms and the recovery of MTL volumes, the robust finding of MTL volume reductions in hyperthyroid GD patients suggests that the mental symptomatology is an expression of GD-related brain changes rather than a psychological reaction to perceived somatic symptoms.

11.3 GD – 6–10 YEARS AFTER FIRST DISEASE MANIFESTATION

Six to 10 years after first disease manifestation is so many years after the onset of GD that most patients, even if they have had disease relapse, have finished antithyroid therapy a long time ago. In general, they have returned to a normal life with only on-demand contact with the healthcare system or they are followed in primary care. In cases of disease relapse, or laboratory or symptomatic complexity, they have been referred back to endocrine clinics for evaluation.

The main observations 6–10 years after treatment were that a substantial number of patients with GD will need life-long thyroxine substitution independently of the initial treatment modality. This rather high rate of hypothyroidism after ATD treatment is in agreement with an early study[167] but is not generally recognized. The reasons for levothyroxine treatment of these patients may be several. One possibility is that ongoing autoimmune thyroiditis in GD results in thyroid failure.[3] Second, a patient's clinician may be more willing to prescribe levothyroxine to prevent hypothyroidism and development or worsening of TAO.[16 168] Third, some patients may have developed blocking TRAb.

What is the best treatment option if there is recurrence of hyperthyroidism after ATD treatment? According to Paper I, a second course of ATD results in only three in ten patients achieving long-standing remission and only about one fifth of them will maintain a normal thyroid function. Alternatives are ablative treatment or daily long-term, low-dose ATD which, in recent years, has been reported as efficient and safe.[169 170]

One out of four patients did not feel fully recovered at follow-up, although their hyperthyroidism was in long-term remission. The presence of TAO has been associated with reduced QoL in previous studies.[79 80 139 140 162] However, in Paper I, patients reported tiredness as the most prevalent cause of not feeling fully recovered. The relationship between the mental symptoms remaining at 15 months and not feeling recovered several years later is not self-evident. However, previous publications[31 32 38 57] with both a shorter and longer follow-up time than 6–10 years have reported similar late symptoms. It is therefore likely that some symptoms at 15 months persist to become more or less chronic. Tiredness was reported in several previous publications.[38 78 161 171] Although tiredness does not equal mental fatigue, the prevalence of the latter should be evaluated in coming long-term studies on GD.

Patients on levothyroxine treatment were more likely to report non-recovery compared to those without levothyroxine. The causality of this finding is unclear. Life-long medical treatment may be a manifestation of impaired health, and not feeling well may lead to primary care visits and laboratory investigations that can increase the likeliness of a diagnosis of hypothyroidism.

The proportion of patients who felt fully recovered did not depend on treatment modality, which is in line with previous observations.[16 31 79]

The long-term mental health consequences of GD need further investigation, particularly in other more diverse populations. The effects of life-long

medication need to be considered when planning therapy for GD[172] and annual follow-up visits should be considered in line with European Thyroid Association guidelines.[173]

12 MAIN POINTS

- Patients with GD have more depression, anxiety, mental fatigue, and cognitive complaints than healthy persons. These symptoms improve with treatment but are still more severe in patients after 15 months, close to the end of specialty care, than in controls.

- Almost one in every four patients starting treatment with ATD will change to ablative treatment during their first treatment period.

- When seeing hyperthyroid GD patients with previous psychiatric conditions or eye symptoms, the clinician should be prepared to handle more severe mental symptoms.

- After 15 months of treatment, mental fatigue is equally severe in patients without eye symptoms or previous psychiatric disease as in a groups with either of these conditions. This points to an unknown mechanism that remains to be explored.

- GD is a condition that clearly affects the brain. Although the mechanism is unknown, the reduced MTL volumes in hyperthyroidism together with their almost complete recovery after 15 months of treatment indicate that the brain alterations follow the course of the disease.

- The inverse correlations between TRAb level and the MTL volumes in hyperthyroidism, as well as the correlations between the TRAb reduction and the increase in MTL volumes with treatment, points to autoimmunity as a possible cause of the MTL volume reduction.

- If a patient relapses after ATD treatment, there are better alternatives than a second course of ATD.

- A much higher proportion of patients will end up undergoing surgery than the proportion that selected this treatment option from the start.

- Although supposedly thyroid sparing, block-and-replace ATD treatment results in hypothyroidism and levothyroxine treatment in a large proportion of patients.

- When viewed from a long-term perspective, the chance of preserving a normal thyroid function with ATD therapy is lower than previously suggested.

- A high proportion of patients report not feeling fully recovered many years after diagnosis and having achieved euthyroidism. This supports the concept that GD is a chronic disease and that long-term follow-up should be routine.

- The most common reported cause for not feeling recovered many years after onset of GD is tiredness. It is plausible, but not known, that this tiredness mainly consists of mental fatigue.

13 FUTURE PERSPECTIVES

Symptoms and signs in GD give little guidance about who will suffer from long-standing symptoms. Continued research is therefore needed regarding both mechanistic explanations, predictions models, and treatment approaches.

Structurally, the MTL volume alterations should be more thoroughly described, where an obvious first task is to investigate whether these volume alterations also exist in men.

The automatic segmentation results should be consolidated by manual volumetry. Manual subregion segmentation will shed new light on which hippocampal and amygdalar subregions are affected. To further increase the detail, an MR scanner with a higher magnetic field strength than 3 Tesla can be used. The dynamics of the MTL volumes can also be captured in more detail by the introduction of more than two assessment time points.

Other regions of the brain have previously been described as affected in GD and further studies are needed in order to consolidate and extend these findings.

Previous functional MR imaging studies have found correlations between brain function and symptoms in the hyperthyroid state.[45 174] Functional MR and other functional techniques such as fNIRS may be feasible ways to further study the associations between the patients' mental symptoms and brain function.

In search of biomarkers, GD-related hormones such as TRH as well as GD-related antibodies such as TPOAb, thyroglobulin antibodies, and TRAb should be investigated. Also, possible causal mechanisms that include hormones such as cortisol or estrogen, neurotransmitters such as epinephrine or norepinephrine, and finally autoimmune factors outside of thyroid autoimmunity are possible future avenues for research.

Prediction of disease relapse is clinically important and there have been promising attempts such as the Graves' Recurrent Events After Therapy (GREAT) score.[175] Predictions regarding the risk of future hypothyroidism with ATD treatment and risk of long-standing symptoms are still in the bud and need to be developed and validated.

Finally, treatment. How should the remaining symptoms be treated? Adjustments of hormonal treatment? Antidepressant medication? Psychotherapy? Attempts have been made with all of these but we still lack evidence for what works and what does not.

14 FINAL THOUGHTS

Research is like walking on thin ice. Even if you have spent hours and hours of preparation on how to avoid slipping or falling into the cold water, you can never be sure that your cautions will lead you safely all the way to the other side. There are risks everywhere. From inclusion of participants that are not representative through methodologic flaws and incorrect calculations all the way to wrong interpretations of the results.

With this in mind, it is my belief that the most important contribution of this thesis to the field of thyroid diseases is that it may put an end to the conviction that GD does not affect the brain. Even if we still lack mechanistic explanations and clinical implications of these findings, the dynamic alterations in the volume of MTL structures in GD clearly indicates that GD is a disease with brain consequences.

There are many other topics related to the content of this thesis that interest me and that I feel are important to study in order to understand the complexity of mental symptoms in GD. An example of such a topic is the gender issue. From previous research in other diseases, it is known that both symptoms, reactions to chronic disease, and the clinical course can differ depending on sex. Despite this, surprisingly little research has been performed on the effect of sex in GD. Unfortunately, there was neither time nor space to cover this in this thesis.

Science depends on both beliefs and doubts, equally. Without beliefs, no hypotheses and without doubts, no critical evaluations of scientific results. The same balanced relationship exists between the human psyche and the body. It is my firm belief that giving both perspectives equal weight in future research on the outcome of GD is the best way to really improve the care of these patients.

15 ACKNOWLEDGEMENT

The patients. Thank you for offering your time and commitment. Without your support, clinical research would not exist.

The controls. Thank you for offering your time and commitment for nothing else than the belief in science.

Helena Filipsson Nyström, my main supervisor. Thank you for 8 inspiring years of discussions, drama, and comedy. For endless support, for warm friendship, and for always, always being constructive. If you ever need my support at any time in the future, I will be there.

Helge Malmgren, my co-supervisor. Thank you for offering infinite time, for warm friendship, and for being a superb mentor in psychiatry, philosophy, and the secret life of birds. Despite a life in science, nobody can beat you when it comes to enthusiasm.

Staffan Edén, my co-supervisor. Thank you for warm friendship, for always being ready for discussions throughout the whole process, and for providing direction at exactly the right moment.

Stefan Arver, my co-supervisor. Thank you for being an energetic, inspiring, and fun leader. For being relentlessly loyal to patients, for never hesitating in standing by my side, and for always supporting me both in my clinical work and as a researcher.

Gudmundur Johansson, my previous co-supervisor. Thank you for hiring me, for inspiring me, and for supporting me in registering as a PhD student.

Gabriel Sjölin, my fellow co-writer. Thank you for keeping in good spirits through the whole writing process and for always being ready to answer my questions, over and over again.

The TT-12 group: Göran Wallin, Jan Calissendorrf, Bengt Hallengren, Mikael Lantz, Selwan Khamisi, Ove Törring, Mirna Abraham Norling Thank you for years of good spirits, support, and continuous constructive feedback.

The CogThy Study Group: Erik Olsson, Niklas Klasson, Rolf Heckemann, Birgitta Johansson, Peter Berglund, Simon Skau, Göran Starck. Thank

you all for support, good times, and brilliant considerations regarding all aspects of the studies.

Jenny Tiberg. Thank you for consistent work on gathering and organizing all the patients in the CogThy study. And thank you for sharing my restlessness.

Dag Holmberg, my big son. Thank you for making me prouder than any dad can be and for always supporting your slightly lagging father.

Nils Holmberg, my smaller son. Thank you for being a true original being and for keeping my curiosity alive. You never cease to surprise me.

Radha Korsoski, my love. Thank you for everything. My life, my love, my future.

The Gothenburg Medical Society, the Swedish Society of Medicine, the regional funds of Västra Götaland, and the Göte Thyrings Fund. Thank you for funding my research.

Torquil Watt. Thank you for kind assistance in interpreting the ThyPRO results.

Everyone else. Summer workers, nurses, colleagues, and teachers from research school, workmates, bosses, friends, and others who have been so supportive in this work.

16 REFERENCES

1. Graves RJ. Clinical lectures delivered at the Meath Hospital during the session 1834–1835. Lecture XII. Newly observed affection of the thyroid. *London Med Surg* 1835:515–23.
2. Abraham-Nordling M, Bystrom K, Torring O, *et al*. Incidence of hyperthyroidism in Sweden. *Eur J Endocrinol* 2011;165:899-905.
3. Smith TJ, Hegedus L. Graves' disease. *New Eng J Med* 2016;375:1552-65.
4. Nyström HF, Jansson S, Berg G. Incidence rate and clinical features of hyperthyroidism in a long-term iodine sufficient area of Sweden (Gothenburg) 2003–2005. *Clin Endocrinol* 2013;78:768-76.
5. Harvey CB, Williams GR. Mechanism of thyroid hormone action. *Thyroid* 2002;12:441-6.
6. Silva JE, Bianco SD. Thyroid-adrenergic interactions: physiological and clinical implications. *Thyroid* 2008;18:157-65.
7. Mullur R, Liu YY, Brent GA. Thyroid hormone regulation of metabolism. *Physiol Rev* 2014;94:355-82.
8. Burch HB, Cooper DS. Management of Graves disease: a review. *JAMA* 2015;314:2544-54.
9. Laurberg P, Berman DC, Bulow Pedersen I, *et al*. Incidence and clinical presentation of moderate to severe graves' orbitopathy in a Danish population before and after iodine fortification of salt. *J Clin Endocrinol Metab* 2012;97:2325-32.
10. Fatourechi V. Thyroid dermopathy and acropachy. *Expert Rev Dermatol* 2011;6:75-90.
11. Sundaresh V, Brito JP, Thapa P, *et al*. Comparative effectiveness of treatment choices for Graves' hyperthyroidism: a historical cohort study. *Thyroid* 2017;27:497-505.
12. Franklyn JA, Boelaert K. Thyrotoxicosis. *Lancet* 2012;379:1155-66.
13. Kim YA, Cho SW, Choi HS, *et al*. The second antithyroid drug treatment is effective in relapsed Graves' disease patients: a median 11-year follow-up study. *Thyroid* 2017;27:491-96.
14. Bartalena L, Burch HB, Burman KD, *et al*. A 2013 European survey of clinical practice patterns in the management of Graves' disease. *Clin Endocrinol (Oxf)* 2016;84:115-20.
15. Cooper DS. Antithyroid drugs. *N Eng J Med* 2005;352:905-17.
16. Torring O, Tallstedt L, Wallin G, *et al*. Graves' hyperthyroidism: treatment with antithyroid drugs, surgery, or radioiodine - a

prospective, randomized study. *J Clin Endocrinol Metab* 1996;81:2986-93.
17. Laurberg P. Remission of Graves' disease during anti-thyroid drug therapy. Time to reconsider the mechanism? *Eur J Endocrinol* 2006;155:783-86.
18. Yang J, Zhu YJ, Zhong JJ, *et al*. Characteristics of antithyroid drug-induced agranulocytosis in patients with hyperthyroidism: a retrospective analysis of 114 cases in a single institution in China involving 9690 patients referred for radioiodine treatment over 15 years. *Thyroid* 2016;26:627-33.
19. Nakamura H, Miyauchi A, Miyawaki N, *et al*. Analysis of 754 cases of antithyroid drug-induced agranulocytosis over 30 years in Japan. *J Clin Endocrinol Metab* 2013;98:4776-83.
20. Sundaresh V, Brito JP, Wang Z, *et al*. Comparative effectiveness of therapies for Graves' hyperthyroidism: a systematic review and network meta-analysis. *J Clin Endocrinol Metab* 2013;98:3671-7.
21. Berglund J, Christensen SB, Dymling JF, *et al*. The incidence of recurrence and hypothyroidism following treatment with antithyroid drugs, surgery or radioiodine in all patients with thyrotoxicosis in Malmo during the period 1970-1974. *J Intern Med* 1991;229:435-42.
22. Ross DS, Burch HB, Cooper DS, *et al*. 2016 American Thyroid Association guidelines for diagnosis and management of hyperthyroidism and other causes of thyrotoxicosis. *Thyroid* 2016;26:1343-421.
23. Metso S, Jaatinen P, Huhtala H, *et al*. Increased cardiovascular and cancer mortality after radioiodine treatment for hyperthyroidism. *J Clin Endocrinol Metab* 2007;92:2190-6.
24. Franklyn JA, Maisonneuve P, Sheppard M, *et al*. Cancer incidence and mortality after radioiodine treatment for hyperthyroidism: a population-based cohort study. *Lancet* 1999;353:2111-5.
25. Franklyn JA, Sheppard MC, Maisonneuve P. Thyroid function and mortality in patients treated for hyperthyroidism. *JAMA* 2005;294:71-80.
26. Angusti T, Codegone A, Pellerito R, *et al*. Thyroid cancer prevalence after radioiodine treatment of hyperthyroidism. *J Nucl Med* 2000;41:1006-9.
27. Hall P, Berg G, Bjelkengren G, *et al*. Cancer mortality after iodine-131 therapy for hyperthyroidism. *Int J Cancer* 1992;50:886-90.
28. Ron E, Doody MM, Becker DV, *et al*. Cancer mortality following treatment for adult hyperthyroidism. Cooperative Thyrotoxicosis Therapy Follow-up Study Group. *JAMA* 1998;280:347-55.

29. Giesecke P, Frykman V, Wallin G, *et al*. All-cause and cardiovascular mortality risk after surgery versus radioiodine treatment for hyperthyroidism. *Br J Surg* 2017;Feb;105:279-86.
30. Traisk F, Tallstedt L, Abraham-Nordling M, *et al*. Thyroid-associated ophthalmopathy after treatment for Graves' hyperthyroidism with antithyroid drugs or iodine-131. *J Clin Endocrinol Metab* 2009;94:3700-7.
31. Abraham-Nordling M, Torring O, Hamberger B, *et al*. Graves' disease: a long-term quality-of-life follow up of patients randomized to treatment with antithyroid drugs, radioiodine, or surgery. *Thyroid* 2005;15:1279-86.
32. Torring O, Watt T, Sjolin G, *et al*. Impaired quality of life after radioiodine therapy compared with antithyroid drugs or surgical treatment for Graves' hyperthyroidism. A long-term follow-up with ThyPRO and SF-36. *Thyroid* 2019;29:322-31.
33. Smithson M, Asban A, Miller J, *et al*. Considerations for thyroidectomy as treatment for Graves disease. *Clin Med Insights Endocrinol Diabetes* 2019;12:1-5.
34. Nygaard B, Hegedus L, Nielsen KG, *et al*. Long-term effect of radioactive iodine on thyroid function and size in patients with solitary autonomously functioning toxic thyroid nodules. *Clin Endocrinol (Oxf)* 1999;50:197-202.
35. Bergenfelz A, Jansson S, Kristoffersson A, *et al*. Complications to thyroid surgery: results as reported in a database from a multicenter audit comprising 3,660 patients. *Langenbecks Arch Surg* 2008;393:667-73.
36. Vogel A, Elberling TV, Hording M, *et al*. Affective symptoms and cognitive functions in the acute phase of Graves' thyrotoxicosis. *Psychoneuroendocrinology* 2007;32:36-43.
37. Bove KB, Watt T, Vogel A, *et al*. Anxiety and depression are more prevalent in patients with graves' disease than in patients with nodular goitre. *Eur Thyroid J* 2014;3:173-8.
38. Watt T, Groenvold M, Rasmussen AK, *et al*. Quality of life in patients with benign thyroid disorders. A review. *Eur J Endocrinol* 2006;154:501-10.
39. Berg G, Michanek A, Holmberg E, *et al*. Clinical outcome of radioiodine treatment of hyperthyroidism: a follow-up study. *J Intern Med* 1996;239:165-71.
40. Elberling TV, Rasmussen AK, Feldt-Rasmussen U, *et al*. Impaired health-related quality of life in Graves' disease. A prospective study. *Eur J Endocrinol* 2004;151:549-55.

41. Miao Q, Zhang S, Guan YH, *et al*. Reversible changes in brain glucose metabolism following thyroid function normalization in hyperthyroidism. *AJNR Am J Neuroradiol* 2011;32:1034-42.
42. Schreckenberger MF, Egle UT, Drecker S, *et al*. Positron emission tomography reveals correlations between brain metabolism and mood changes in hyperthyroidism. *J Clin Endocrinol Metab* 2006;91:4786-91.
43. Trzepacz PT, McCue M, Klein I, *et al*. A psychiatric and neuropsychological study of patients with untreated Graves' disease. *Gen Hosp Psychiatry* 1988;10:49-55.
44. Chattopadhyay C, Chakrabarti N, Ghosh S. An assessment of psychiatric disturbances in Graves disease in a medical college in eastern India. *Niger J Clin Pract* 2012;15:276-9.
45. Li L, Zhi MM, Hou ZH, *et al*. Abnormal brain functional connectivity leads to impaired mood and cognition in hyperthyroidism: a resting-state functional MRI study. *Oncotarget* 2017;8:6283-94.
46. Zhang W, Liu X, Zhang Y, *et al*. Disrupted functional connectivity of the hippocampus in patients with hyperthyroidism: evidence from resting-state fMRI. *Eur J Radiol* 2014;83:1907-13.
47. Calissendorff J, Mikulski E, Larsen EH, *et al*. A prospective investigation of Graves' disease and selenium: thyroid hormones, auto-antibodies and self-rated symptoms. *Eur Thyroid J* 2015;4:93-8.
48. Tsymbaliuk I, Unukovych D, Shvets N, *et al*. Cardiovascular complications secondary to Graves' disease: a prospective study from Ukraine. *PLoS One* 2015;10:1-15.
49. Riguetto CM, Neto AM, Tambascia MA, *et al*. The relationship between quality of life, cognition, and thyroid status in Graves' disease. *Endocrine* 2019;63:87-93.
50. Bunevicius R, Velickiene D, Prange Jr AJ. Mood and anxiety disorders in women with treated hyperthyroidism and ophthalmopathy caused by Graves' disease. *Gen Hosp Psychiatry* 2005;27:133-39.
51. Hu LY, Shen CC, Hu YW, *et al*. Hyperthyroidism and risk for bipolar disorders: a nationwide population-based study. *PLoS One* 2013;8:1-5
52. Thomsen AF, Kessing LV. Increased risk of hyperthyroidism among patients hospitalized with bipolar disorder. *Bipolar disorders* 2005;7:351-7.
53. Fukao A, Takamatsu J, Arishima T, *et al*. Graves' disease and mental disorders. *J Clin Transl Endocrinol* 2020;19:1-8.

54. Winsa B, Adami HO, Bergstrom R, *et al*. Stressful life events and Graves' disease. *Lancet* 1991;338:1475-9.
55. Brandt F, Thvilum M, Almind D, *et al*. Hyperthyroidism and psychiatric morbidity: evidence from a Danish nationwide register study. *Eur J Endocrinol* 2014;170:341-8.
56. Kathol RG, Turner R, Delahunt J. Depression and anxiety associated with hyperthyroidism - response to antithyroid therapy. *Psychosomatics* 1986;27:501-05.
57. Fahrenfort JJ, Wilterdink AM, van der Veen EA. Long-term residual complaints and psychosocial sequelae after remission of hyperthyroidism. *Psychoneuroendocrinology* 2000;25:201-11.
58. Perrild H, Hansen JM, Arnung K, *et al*. Intellectual impairment after hyperthyroidism. *Acta Endocrinol (Copenh)* 1986;112:185-91.
59. Stern RA, Robinson B, Thorner AR, *et al*. A survey study of neuropsychiatric complaints in patients with Graves' disease. *J Neuropsychiatry Clin Neurosci* 1996;8:181-5.
60. Radosavljević VR, Janković SM, Marinković JM. Stressful life events in the pathogenesis of Graves' disease. *Eur J Endocrinol* 1996;134:699-701.
61. Vita R, Lapa D, Trimarchi F, *et al*. Stress triggers the onset and the recurrences of hyperthyroidism in patients with Graves' disease. *Endocrine* 2015;48:254-63.
62. Ittermann T, Volzke H, Baumeister SE, *et al*. Diagnosed thyroid disorders are associated with depression and anxiety. *Soc Psychiatry Psychiatr Epidemiol* 2015;50:1417-25.
63. Fukao A, Takamatsu J, Kubota S, *et al*. The thyroid function of Graves' disease patients is aggravated by depressive personality during antithyroid drug treatment. *Biopsychosoc Med* 2011;5:9.
64. MacCrimmon DJ, Wallace JE, Goldberg WM, *et al*. Emotional disturbance and cognitive deficits in hyperthyroidism. *Psychosom Med* 1979;41:331-40.
65. Pop VJ, Maartens LH, Leusink G, *et al.* Are autoimmune thyroid dysfunction and depression related? *J Clinl Endocrinol Metab* 1998;83:3194-97.
66. Carta MG, Loviselli A, Hardoy MC, *et al*. The link between thyroid autoimmunity (antithyroid peroxidase autoantibodies) with anxiety and mood disorders in the community: a field of interest for public health in the future. *BMC Psychiatry* 2004;4:25-25.
67. Engum A, Bjoro T, Mykletun A, *et al*. An association between depression, anxiety and thyroid function-a clinical fact or an artefact? *Acta Psychiatr Scand* 2002;106:27-34.

68. Grabe HJ, Volzke H, Ludemann J, *et al*. Mental and physical complaints in thyroid disorders in the general population. *Acta Psychiatr Scand* 2005;112:286-93.
69. Yudiarto FL, Muliadi L, Moeljanto D, *et al*. Neuropsychological findings in hyperthyroid patients. *Acta Med Indones* 2006;38:6-10.
70. Kumar M, Rana P, Modi S, *et al*. Aberrant intra and inter network resting state functional connectivity in thyrotoxicosis. *J Neuroendocrinol* 2019;31:11.
71. Moller M, Bartfai A, Floter Radestad A, *et al*. The relationship between thyroid hormone levels and cognitive functions in Graves' disease might be influenced by a history of mild traumatic brain injury. *Brain Nerves* 2017;1:1-5
72. Lindqvist G, Malmgren H. Organisk Psykiatri (Organic Psychiatry). Stockholm: Almqvist & Wiksell 1990.
73. Rodholm M, Starmark JE, Svensson E, *et al*. Astheno-emotional disorder after aneurysmal SAH: reliability, symptomatology and relation to outcome. *Acta Neurol Scand* 2001;103:379-85.
74. Johansson B, Ronnback L. Long-lasting mental fatigue after traumatic brain injury – a major problem most often neglected diagnostic criteria, assessment, relation to emotional and cognitive problems, cellular background, and aspects on treatment. *IntechOpen* 2014: DOI: 10.5772/57311
75. Johansson B, Starmark A, Berglund P, *et al*. A self-assessment questionnaire for mental fatigue and related symptoms after neurological disorders and injuries. *Brain Inj* 2010;24:2-12.
76. Manual of the International Statistical Classification of Diseases, Injuries, and Causes of Death, Tenth Revision. Geneva; Switzerland: World Health Organization 1992.
77. Diagnostic and Statistical Manual of Mental Disorders (5th ed.). Arlington, VA.: American Psychiatric Association 2013.
78. Cramon P, Winther KH, Watt T, *et al*. Quality-of-life impairments persist six months after treatment of Graves' hyperthyroidism and toxic nodular goiter: a prospective cohort study. *Thyroid* 2016;26:1010-8.
79. Ljunggren JG, Torring O, Wallin G, *et al*. Quality of life aspects and costs in treatment of Graves' hyperthyroidism with antithyroid drugs, surgery, or radioiodine: results from a prospective, randomized study. *Thyroid* 1998;8:653-9.
80. Nexo MA, Watt T, Pedersen J, *et al*. Increased risk of long-term sickness absence, lower rate of return to work, and higher risk of unemployment and disability pensioning for thyroid patients: a

Danish register-based cohort study. *J Clin Endocrinol Metab* 2014;99:3184-92.
81. Stepien BK, Huttner WB. Transport, metabolism and function of thyroid hormones in the developing mammalian brain. *Front Endocrinol (Lausanne)* 2019;10:209.
82. de Escobar GM, Obregon MJ, del Rey FE. Maternal thyroid hormones early in pregnancy and fetal brain development. *Best Pract Res Clinical Endocrinol Metab* 2004;18:225-48.
83. McLeod DS, McIntyre HD. Subclinical hypothyroidism and related biochemical entities in pregnancy: implications and management. *Obstet Med* 2010;3:139-44.
84. Zoeller RT. Transplacental thyroxine and fetal brain development. *J Clinical Invest* 2003;111:954-57.
85. Jansen TA, Korevaar TIM, Mulder TA, *et al*. Maternal thyroid function during pregnancy and child brain morphology: a time window-specific analysis of a prospective cohort. *Lancet Diabetes Endocrinol* 2019;7:629-37.
86. Schroeder AC, Privalsky ML. Thyroid hormones, t3 and t4, in the brain. *Front Endocrinol (Lausanne)* 2014;5:40.
87. Kapoor R, Fanibunda SE, Desouza LA, *et al*. Perspectives on thyroid hormone action in adult neurogenesis. *J Neurochem* 2015;133:599-616.
88. Gereben B, Zeold A, Dentice M, *et al*. Activation and inactivation of thyroid hormone by deiodinases: local action with general consequences. *Cell Mol Life Sci* 2008;65:570-90.
89. Peeters RP, Visser TJ. Metabolism of Thyroid Hormone. In: De Groot LJ, Chrousos G, Dungan K, et al., eds. Endotext. South Dartmouth (MA): MDText.com, Inc. 2000.
90. Peeters R, Fekete C, Goncalves C, *et al*. Regional physiological adaptation of the central nervous system deiodinases to iodine deficiency. *Am J Physiol Endocrinol Metab* 2001;281:E54-61.
91. de Escobar GM, Obregón MJ, del Rey FE. Iodine deficiency and brain development in the first half of pregnancy. *Public health nutr* 2007;10:1554-70.
92. Ferreiro B, Bernal J, Goodyer CG, *et al*. Estimation of nuclear thyroid hormone receptor saturation in human fetal brain and lung during early gestation. *J Clin Endocrinol Metab* 1988;67:853-6.
93. Cook CB, Koenig RJ. Expression of erbA alpha and beta mRNAs in regions of adult rat brain. *Mol Cell Endocrinol* 1990;70:13-20.
94. Wallis K, Dudazy S, van Hogerlinden M, *et al*. The thyroid hormone receptor alpha1 protein is expressed in embryonic postmitotic

neurons and persists in most adult neurons. *Mol Endocrinol* 2010;24:1904-16.
95. Bradley DJ, Young WS, 3rd, Weinberger C. Differential expression of alpha and beta thyroid hormone receptor genes in rat brain and pituitary. *Proc Natl Acad Sci U S A* 1989;86:7250-4.
96. Daimon CM, Chirdon P, Maudsley S, *et al*. The role of Thyrotropin Releasing Hormone in aging and neurodegenerative diseases. *Am J Alzheimers Dis (Columbia)* 2013;1:1-25
97. Naicker M, Naidoo S. Expression of thyroid-stimulating hormone receptors and thyroglobulin in limbic regions in the adult human brain. *Metab Brain Dis* 2018;33:481-89
98. Mouri A, Hoshino Y, Narusawa S, *et al*. Thyrotoropin receptor knockout changes monoaminergic neuronal system and produces methylphenidate-sensitive emotional and cognitive dysfunction. *Psychoneuroendocrinology* 2014;48:147-61.
99. Frohlich E, Wahl R. Thyroid Autoimmunity: Role of anti-thyroid antibodies in thyroid and extra-thyroidal diseases. *Front Immunol* 2017;8:521.
100. Molnar I, Szentmiklosi JA, Gesztelyi R, *et al*. Effect of antithyroid drugs on the occurrence of antibodies against type 2 deiodinase (DIO2), which are involved in hyperthyroid Graves' disease influencing the therapeutic efficacy. *Clin Exp Med* 2019;19:245-54.
101. Pritchard J, Han R, Horst N, *et al*. Immunoglobulin activation of T cell chemoattractant expression in fibroblasts from patients with Graves' disease is mediated through the insulin-like growth factor I receptor pathway. *J Immunol* 2003;170:6348-54.
102. Xu J, Zhao L, Xiang G, *et al*. Relationship between autoantibody to the angiotensin II-1 receptor and cardiovascular manifestations of Graves' disease. *Exp Clin Endocrinol Diabetes* 2014;226:254-58.
103. Stavrakis S, Yu X, Patterson E, *et al*. Activating autoantibodies to the beta-1 adrenergic and m2 muscarinic receptors facilitate atrial fibrillation in patients with Graves' hyperthyroidism. *J Am Coll Cardiol* 2009;54:1309-16.
104. Galloway A, Li H, Vanderlinde-Wood M, *et al*. Activating autoantibodies to the β1/2-adrenergic and M2 muscarinic receptors associate with atrial tachyarrhythmias in patients with hyperthyroidism. *Endocrine* 2015;49:457-63.
105. Bunevicius R, Prange AJ, Jr. Thyroid disease and mental disorders: cause and effect or only comorbidity? *Curr Opin Psychiatry* 2010;23:363-8.
106. Squire LR, Stark CE, Clark RE. The medial temporal lobe. *Annu Rev Neurosci* 2004;27:279-306.

107. McDonald AJ, Mott DD. Functional neuroanatomy of amygdalohippocampal interconnections and their role in learning and memory. *J Neurosci Res* 2017;95:797-820.
108. DeKraker J, Ferko KM, Lau JC, *et al*. Unfolding the hippocampus: An intrinsic coordinate system for subfield segmentations and quantitative mapping. *NeuroImage* 2018;167:408-18.
109. Ledoux J. The amygdala. *Current Biology* 2007;17:R868-R74.
110. Eriksson PS, Perfilieva E, Björk-Eriksson T, *et al*. Neurogenesis in the adult human hippocampus. *Nature medicine* 1998;4:1313.
111. Geerlings MI, Gerritsen L. Late-life depression, hippocampal volumes, and hypothalamic-pituitary-adrenal axis regulation: a systematic review and meta-analysis. *Biological Psychiatry* 2017;82:339-50.
112. Bauduin S, van der Wee NJA, van der Werff SJA. Structural brain abnormalities in Cushing's syndrome. *Curr Opin Endocrinol Diabetes Obes* 2018;25:285-89.
113. Bremner JD, Narayan M, Anderson ER, *et al*. Hippocampal volume reduction in major depression. *Am J Psychiatry* 2000;157:115-8.
114. Mueller SG, Schuff N, Yaffe K, *et al*. Hippocampal atrophy patterns in mild cognitive impairment and Alzheimer's disease. *Human brain mapping* 2010;31:1339-47.
115. Santos A, Granell E, Gomez-Anson B, *et al*. Depression and Anxiety Scores Are Associated with Amygdala Volume in Cushing's Syndrome: Preliminary Study. *Biomed Res Int* 2017;2017:1-7.
116. van Mierlo TJ, Chung C, Foncke EM, *et al*. Depressive symptoms in Parkinson's disease are related to decreased hippocampus and amygdala volume. *Mov Disord* 2015;30:245-52.
117. Kumar A, Pareek V, Faiq MA, *et al*. Adult neurogenesis in humans: a review of basic concepts, history, current research, and clinical implications. *Innov Clin Neurosci* 2019;16:30-37.
118. Spalding KL, Bergmann O, Alkass K, *et al*. Dynamics of hippocampal neurogenesis in adult humans. *Cell* 2013;153:1219-27.
119. Firth J, Stubbs B, Vancampfort D, *et al*. Effect of aerobic exercise on hippocampal volume in humans: A systematic review and meta-analysis. *NeuroImage* 2018;166:230-38.
120. Pletzer B, Harris T, Hidalgo-Lopez E. Subcortical structural changes along the menstrual cycle: beyond the hippocampus. *Sci Rep* 2018;8:16042.
121. Wang YP, Xu QF, Luo J, *et al*. Effects of age and sex on subcortical volumes. *Front Aging Neurosci* 2019;11:12.

122. Barnes J, Bartlett JW, van de Pol LA, *et al*. A meta-analysis of hippocampal atrophy rates in Alzheimer's disease. *Neurobiol Aging* 2009;30:1711-23.
123. Apostolova LG, Green AE, Babakchanian S, *et al*. Hippocampal atrophy and ventricular enlargement in normal aging, mild cognitive impairment (MCI), and Alzheimer disease. *Alzheimer Dis Assoc Disord* 2012;26:17-27.
124. O'Doherty DCM, Chitty KM, Saddiqui S, *et al*. A systematic review and meta-analysis of magnetic resonance imaging measurement of structural volumes in posttraumatic stress disorder. *Psychiatry Res Neuroimaging* 2015;232:1-33.
125. Roddy DW, Farrell C, Doolin K, *et al*. The hippocampus in depression: more than the sum of its parts? advanced hippocampal substructure segmentation in depression. *Biological Psychiatry* 2019;85:487-97.
126. Wise T, Radua J, Via E, *et al*. Common and distinct patterns of grey-matter volume alteration in major depression and bipolar disorder: evidence from voxel-based meta-analysis. *Molecular Psychiatry* 2017;22:1455-63.
127. Weintraub D, Doshi J, Koka D, *et al*. Neurodegeneration across stages of cognitive decline in Parkinson disease. *Arch Neurol* 2011;68:1562-68.
128. Bourdeau I, Bard Cl, Noël B, *et al*. Loss of brain volume in endogenous Cushing's syndrome and its reversibility after correction of hypercortisolism. *J Clin Endocrinol Metab* 2002;87:1949-54.
129. Cano M, Martinez-Zalacain I, Bernabeu-Sanz A, *et al*. Brain volumetric and metabolic correlates of electroconvulsive therapy for treatment-resistant depression: a longitudinal neuroimaging study. *Translational psychiatry* 2017;7:e1023.
130. Tendolkar I, van Beek M, van Oostrom I, *et al*. Electroconvulsive therapy increases hippocampal and amygdala volume in therapy refractory depression: a longitudinal pilot study. *Psychiatry Res* 2013;214:197-203.
131. Vriend C, Boedhoe PS, Rutten S, *et al*. A smaller amygdala is associated with anxiety in Parkinson's disease: a combined FreeSurfer--VBM study. *J Neurol Neurosurg Psychiatry* 2015;87:493-500.
132. Klein-Koerkamp Y, Heckemann RA, Ramdeen KT, *et al*. Amygdalar atrophy in early Alzheimer's disease. *Curr Alzheimer Res* 2014;11:239-52.

133. Zavorotnyy M, Zollner R, Schulte-Gustenberg LR, *et al*. Low left amygdala volume is associated with a longer duration of unipolar depression. *J Neural Transm (Vienna)* 2018;125:229-38.
134. Zhang W, Song L, Yin X, *et al*. Grey matter abnormalities in untreated hyperthyroidism: a voxel-based morphometry study using the DARTEL approach. *Eur J Radiol* 2014;83:e43-8.
135. Ittermann T, Wittfeld K, Nauck M, *et al*. High thyrotropin is associated with reduced hippocampal volume in a population-based study from Germany. *Thyroid* 2018;11:1-9
136. Andersson M, Berg G, Eggertsen R, *et al*. Adequate iodine nutrition in Sweden: a cross-sectional national study of urinary iodine concentration in school-age children. *Eur J Clin Nutr* 2009;63:828-34.
137. Saravanan P, Chau WF, Roberts N, *et al*. Psychological well-being in patients on 'adequate' doses of l-thyroxine: results of a large, controlled community-based questionnaire study. *Clin Endocrinol (Oxf)* 2002;57:577-85.
138. Johansson R, Carlbring P, Heedman Å, *et al*. Depression, anxiety and their comorbidity in the Swedish general population: point prevalence and the effect on health-related quality of life. *PeerJ* 2013;1:e98.
139. Wiersinga WM. Quality of life in Graves' ophthalmopathy. *Best Pract Res Clin Endocrinol Metab* 2012;26:359-70.
140. Abraham-Nordling M, Wallin G, Traisk F, *et al*. Thyroid-associated ophthalmopathy; quality of life follow-up of patients randomized to treatment with antithyroid drugs or radioiodine. *Eur J Endocrinol* 2010;163:651-7.
141. Ponto KA, Pitz S, Pfeiffer N, *et al*. Quality of life and occupational disability in endocrine orbitopathy. *Deutsches Arzteblatt international* 2009;106:283.
142. Yeatts RP. Quality of life in patients with Graves ophthalmopathy. *Trans Am Ophthalmol Soc* 2005;103:368.
143. First MB, Spitzer RL, Gibbon M, *et al*. Structured clinical interview for DSM-IV-TR axis I disorders, research version, patient edition: SCID-I/P New York, NY, 2002.
144. First M, Spitzer R, Gibbon M, *et al*. Structured Clinical Interviews for DSM-IV Axis I Disorders-Clinical Version (SCID-CV). 1997. Herlofson J: Translation to Swedish language, Pilgrim Press: Danderyd, Sweden, 1998 ed: Washington, DC: American Psychiatric Press, 1998.
145. Svanborg P, Åsberg M. A new self-rating scale for depression and anxiety states based on the Comprehensive Psychopathological Scale. *Acta Psychiatr Scand* 1994;89:21-28.

146. Lindqvist G, Malmgren H. Classification and Diagnosis in Organic Psychiatry. *Acta Psychiatr Scand* 1993;88(Suppl 373)
147. Johansson B, Ronnback L. Evaluation of the mental fatigue scale and its relation to cognitive and emotional functioning after traumatic brain injury or stroke. *Int J Phys Med Rehabil* 2013;02:1-7
148. Wiklund I, Karlberg J. Evaluation of quality of life in clinical trials. Selecting quality-of-life measures. *Control Clin Trials* 1991;12:204S-16S.
149. Watt T, Hegedus L, Groenvold M, *et al*. Validity and reliability of the novel thyroid-specific quality of life questionnaire, ThyPRO. *Eur J Endocrinol* 2010;162:161-7.
150. Watt T. Cross-cultural validity of the thyroid-specific quality-of-life patient-reported outcome measure, ThyPRO. *Qual Life Res* 2015;24:769-80.
151. Cabezas M, Oliver A, Lladó X, *et al*. A review of atlas-based segmentation for magnetic resonance brain images. *Comput Methods Programs Biomed* 2011;104:e158-e77.
152. Heckemann RA, Keihaninejad S, Aljabar P, *et al*. Improving intersubject image registration using tissue-class information benefits robustness and accuracy of multi-atlas based anatomical segmentation. *NeuroImage* 2010;51:221-7.
153. Hammers A, Allom R, Koepp MJ, *et al*. Three-dimensional maximum probability atlas of the human brain, with particular reference to the temporal lobe. *Human brain mapping* 2003;19:224-47.
154. Wild HM, Heckemann RA, Studholme C, *et al*. Gyri of the human parietal lobe: Volumes, spatial extents, automatic labelling, and probabilistic atlases. *PLoS One* 2017;12:e0180866.
155. Klasson N. Intracranial volume in neuroimaging: estimation and use in regional brain volume normalization. Thesis. University of Gothenburg, Sahlgrenska Academy, 2019.
156. Bartalena L, Baldeschi L, Boboridis K, *et al*. The 2016 European Thyroid Association/European Group on Graves' Orbitopathy Guidelines for the Management of Graves' Orbitopathy. *Eur Thyroid J* 2016;5:9-26.
157. Demet MM, Özmen B, Deveci A, *et al*. Depression and anxiety in hyperthyroidism. *Arch Med Res* 2002;33:552-56.
158. Sprah L, Dernovsek MZ, Wahlbeck K, *et al*. Psychiatric readmissions and their association with physical comorbidity: a systematic literature review. *BMC Psychiatry* 2017;17:2.

159. Conaglen HM, Tamatea JAU, Conaglen JV, *et al*. Treatment choice, satisfaction and quality of life in patients with Graves' disease. *Clin Endocrinol (Oxf)* 2018;88:977-84.
160. Wechsler D, editor. *Wechsler Adult Intelligence Scale – third edition, WAIS-III, Swedish version*: Pearson Assessment, 2003.
161. Bukvic B, Zivaljevic V, Sipetic S, *et al*. Improved quality of life in hyperthyroidism patients after surgery. *J Surg Res* 2015;193:724-30.
162. Gerding MN, Terwee CB, Dekker FW, *et al*. Quality of life in patients with Graves' ophthalmopathy is markedly decreased: measurement by the medical outcomes study instrument. *Thyroid* 1997;7:885-9.
163. Watt T, Cramon P, Hegedüs L, *et al*. The thyroid-related quality of life measure ThyPRO has good responsiveness and ability to detect relevant treatment effects. *J Clin Endocrinol Metab* 2014;99:3708-17.
164. Altshuler LL, Bauer M, Frye MA, *et al*. Does thyroid supplementation accelerate tricyclic antidepressant response? A review and meta-analysis of the literature. *Am J Psychiatry* 2001;158:1617-22.
165. Jonklaas J, Davidson B, Bhagat S, *et al*. Triiodothyronine levels in athyreotic individuals during levothyroxine therapy. *JAMA* 2008;299:769-77.
166. Woeber KA. Levothyroxine therapy and serum free thyroxine and free triiodothyronine concentrations. *J Endocrinol Invest* 2002;25:106-9.
167. Wood LC, Ingbar SH. Hypothyroidism as a late sequela in patient with Graves' disease treated with antithyroid agents. *J Clin Invest* 1979;64:1429-36.
168. Tallstedt L, Lundell G, Torring O, *et al*. Occurrence of ophthalmopathy after treatment for Graves' hyperthyroidism. The Thyroid Study Group. *N Eng J Med*1992;326:1733-8.
169. Azizi F, Malboosbaf R. Long-term antithyroid drug treatment: a systematic review and meta-analysis. *Thyroid* 2017;27:1223-31.
170. Villagelin D, Romaldini JH, Santos RB, *et al*. Outcomes in relapsed Graves' disease patients following radioiodine or prolonged low dose of methimazole treatment. *Thyroid* 2015;25:1282-90.
171. Möller MC, Bartfai A, Nygren de Boussard C, *et al*. High rates of fatigue in newly diagnosed Graves' disease. *Fatigue: Biomedicine, Health & Behavior* 2014;2:153-62.
172. Turner J, Kelly B. Emotional dimensions of chronic disease. *West J Med* 2000;172:124-28.

173. Kahaly GJ, Bartalena L, Hegedüs L, *et al*. 2018 European Thyroid Association guideline for the management of Graves' hyperthyroidism. *Eur Thyroid Jl* 2018;7:167-86.
174. Zhi M, Hou Z, We Q, *et al*. Abnormal spontaneous brain activity is associated with impaired emotion and cognition in hyperthyroidism: A rs-fMRI study. *Behav Brain Res* 2018;351:188-94.
175. Vos XG, Endert E, Zwinderman AH, *et al*. Predicting the risk of recurrence before the start of antithyroid drug therapy in patients with Graves' hyperthyroidism. *J Clin Endocrinol Metab* 2016;101:1381-9.