

The impact of surgery on quality of life, esophageal motility, and tracheal anatomy and airflow in patients with benign nodular goiter

Jesper Roed Sorensen

This review has been accepted as a thesis together with three previously published papers by University of Southern Denmark 29th of September 2017 and defended on 25th of January 2018.

Tutors: Helle Døssing, Laszlo Hegedüs, Steen Joop Bonnema and Christian Godballe

Official opponents: Allan Carlé, Antti Mäkitie and Peter Licht

Correspondence: Department of ORL Head & Neck Surgery, Odense University Hospital, J.B. Winsløws vej 4, 500 0 Odense C, Denmark

E-mail: jesper.roed.sorensen@rsyd.dk

Dan Med J 2018;65(4):B5472

THE THREE ORIGINAL PAPERS ARE:

I. The quality of life after thyroidectomy in patients with non-toxic nodular Goiter. A prospective cohort study
Sorensen JR, Watt T, Cramon P, Døssing H, Hegedüs L, Bonnema SJ, Godballe C. *Head & Neck*, 2017 Nov;39(11):2232-2240. doi: 10.1002/hed.24886.

II. Changes in swallowing symptoms and esophageal motility after thyroid surgery. A prospective cohort study
Sorensen JR, Markoew S, Døssing H, Hegedüs L, Bonnema SJ, and Godballe C. *World Journal of Surgery*, 2017 October 11. doi: 10.1007/s00268-017-4247-5. [E-pub ahead of print]

III. Thyroidectomy improves tracheal anatomy and airflow in patients with nodular goiter. A prospective cohort study
Sorensen JR, Lauridsen JF, Døssing H, Nguyen N, Hegedüs L, Bonnema SJ, and Godballe C. *European Thyroid Journal*, 2017 Nov;6(6):307-314. doi: 10.1159/000480348.

INTRODUCTION

HISTORY OF THYROIDECTOMY

The word goiter is derived from the Latin word "gutter" meaning throat or neck. Goiter is defined as an enlargement of the thyroid gland in the absence of autoimmune thyroid disease, inflammation, or malignancy [1]. Until the mid-nineteenth century seaweed was, because of its high iodine content, among the only feasible treatments for patients with goiter. Thyroid surgery was not possible due to the inadequate technical development of surgery [2, 3]. Early attempts at performing surgery were horrifying and associated with very high mortality from bleeding and postoperative sepsis [3]. However, with the development of ether

anesthesia (1846), antiseptics (1867), and artery forceps (1877), unhurried and safer dissections were now possible. Surgery for patients with large suffocating goiters was thereafter a risky, but achievable procedure [3, 4]. Theodor Billroth in Vienna (1877) published his leading work showing an unprecedented mortality rate of 8% after thyroidectomy by the use of antiseptic techniques [2]. His achievements were further advanced by his mentee Theodor Kocher in Bern who would later receive the Nobel prize for his pioneering work in understanding thyroid disease (1909) [2, 4, 5]. The hallmarks of Kocher's surgical technique were antiseptics, arterial ligation, and precise intracapsular thyroid dissection, which eventually reduced the mortality rate to below 1% [2, 4, 5]. Nonetheless, with the introduction of safer surgery, recognition of surgical complications (recurrent laryngeal nerve (RLN) paralyzes, thyroid storm, myxedema, and hypocalcemia/tetany) began to occur.

Subsequent surgical achievements included the introduction of subtotal thyroidectomy by Thomas Dunhill (1920) which saved the lives of many young patient with severe hyperthyroidism [4]. However, the pioneering work of these early surgeons was not well known across continental Europe, throughout the British Empire, or in the United States. In these territories, thyroid surgery was a dubious endeavor associated with high risks of mortality. Eventually, William Halsted of Baltimore (today John Hopkins hospital), George Crile from Cleveland (today Cleveland clinic), and Charles Mayo from Rochester (today Mayo Clinic) introduced the European advances to the United States and achieved a recorded RLN palsy rate of 0.3% and introduced parathyroid auto-transplantations [2, 4].

Alongside these surgical achievements, other medical specialties also progressed and introduced: commercially available thyroxine (1917), iodine prophylaxis (1920s), fine needle aspirations (1948), safe positive pressure ventilation (1965), better antithyroid drugs (1965), and B-mode ultrasound (US) (1970s). All this contributed to the pronounced progress, which has made most thyroid diseases manageable in the present era [4].

BACKGROUND

PATHOPHYSIOLOGY OF GOITER

The role of the thyroid gland is to produce thyroid hormones [triiodothyronine and thyroxine (T₄)], which are important for cellular activities such as cell growth and cell development [6, 7]. Approximately 12% of females in Denmark have an enlarged thyroid gland with nodules which is called nodular goiter [8]. Thyroid growth precedes the formation of nodules, and the nod-

ules may eventually become autonomic, causing increased levels of thyroid hormones. However, the thyroid growth is subject to large variation among individuals ranging from extremely slow (no significant change in five years [9]) to a 20% nodule growth per year [10].

The most important etiology for goiter worldwide is the direct correlation between insufficient iodine intake, iodine deficiency, and the formation of goiter. Iodine deficiency causes low blood iodine levels, which leads to a reduction in T₄ levels and increased secretion of thyroid stimulating hormone to restore T₄ concentration. The increased TSH secretion also stimulates thyroid follicular cell hyperplasia and hypertrophy leading to a diffuse enlargement of the thyroid gland [11-13]. Denmark has traditionally been mildly to moderately iodine deficient with a goiter prevalence of 15-23% [14, 15], with, however, differences in prevalence between the eastern (15%) and western (23%) parts of Denmark [16]. The Danish iodine fortification program to correct the iodine deficiency commenced in the year 2000 and consists of iodine-enriched household salt and iodine enriched salt in commercial bread. Urine iodine excretion levels have consequently risen [17], and subsequent the prevalence of goiter [18] and the known age-related increase in thyroid volumes have decreased in Denmark [19].

Other factors also contribute to goiter-formation as goiters are present in both iodine deficient and sufficient areas [20]. Genetic factors account for approximately 39% of the risk of developing goiter (95% confidence interval 0-79%) [21, 22]. In addition to iodine deficiency, tobacco smoking with the formation of thiocyanate (which inhibits iodine uptake) [23, 24], and insulin resistance [25-31] can also promote the formation of goiters. Further, environmental pollutants such as perchlorate and nitrate (inhibits iodine uptake), are rare causes for goiter formation [32].

As the goiter progresses in volume, patients may develop a wide range of pressure symptoms [33]. This thesis is limited to patients with goiter who have symptomatic non-toxic, subclinical hyper- and hypothyroid goiter, and thereby only the anatomical pathology of goiter and not the impact of thyroid toxicity are investigated.

COMPRESSIVE GOITER

The compressive symptoms of goiter have a small, but significantly negative impact on health-related quality of life (HRQoL), as shown in both generic and disease-specific quality of life questionnaires [34, 35].

Swallowing abnormalities (effortful swallowing, globus-, and foreign body sensation) are among the most prevalent symptoms reported in 47-83% of patients with goiter referred for thyroidectomy [36-38]. The underlying pathophysiology is insufficiently understood as only a few studies have used objective examinations to understand this phenomenon. An explanation could be esophageal compression which is reported in 8-27% of patients referred for surgical treatment, however, with no clear definition of compression in this study [39, 40]. In addition, nearly 40% of patients with goiter had increased esophageal transit time compared to healthy controls [41] and the esophageal transit time increased with increasing goiter volumes [42]. Another study found 50% of the patients undergoing total thyroidectomy had disturbances in esophageal motility [36]. However, these patients were examined by water perfusion systems, which makes it difficult to examine the upper esophagus related to the thyroid gland due to insufficient dynamic performance [43] and low spatial resolution of the system [44]. The interpretation of all above

studies is hampered by selection bias as patients have been recruited in different environments (primary care, local center, tertiary center, and surgical or medical hospital department) making it difficult to compare prevalence as the cohorts varies greatly.

Limitations in tracheal airflow are frequently reported in patients with goiter and are suspected to be severely underestimated [45]. Upper airway obstruction (UAO) is reported in 14-31% of patients referred to tertiary centers for evaluation of goiter [46, 47], and in 26-60% of patients referred for thyroidectomy [48-50]. Tracheal deviation is reported in 69-73% of patients with large goiters referred for thyroid surgery [39, 40, 51-57], whereas tracheal compression is seen in 9-58% of patients with cervical goiters [46, 51, 58], and in 35-73% of patients with substernal goiters [39, 40, 52, 54, 59]. Most studies in this area are hampered by ill-defined parameters of compression and inadequate information on cohort composition making generalization to the individual patient difficult.

A few rare and extreme findings of goiter associated tissue compression have been published. These reflect the close relationship of goiter to nearby structures other than the esophagus and the trachea. They include reversible recurrent laryngeal nerve palsy [60], postoperative tracheomalacia [61, 62], tracheal varices from superior vena cava syndrome [63], headache from carotid and vertebral compression [64], chylothorax from thoracic duct compression [65], brachiocephalic vein compression [66], acute upper limb ischemia from arterial compression [67], and internal jugular vein thrombosis [68].

As the symptoms of the goiter presented are essential in the decision to initiate treatment, the pre-treatment consultation and the presentation of all available treatment are of great importance to align expectations of patients and physicians.

THYROID SURGERY

In patients with small goiters without symptoms of compression or cosmetic complaints, the watch-and-wait strategy can be optimal as further thyroid growth is by no means certain. Only a little is known with respect to the changes in HRQoL over time in patients with goiter of any size. The comparison of the various treatment options is therefore difficult and primarily focuses on post-treatment complications and surrogate markers for the beneficial effects.

Surgery is by far the most widely used treatment option for goiter and thyroid lobectomy, subtotal thyroidectomy, and total thyroidectomy, are all acceptable strategies. Both thyroid lobectomy and total thyroidectomy have a low risk of secondary surgery. In contrast whereas subtotal thyroidectomy has a 3% - 42% risk of goiter recurrence [69] and a three- to ten-fold higher risk of postoperative complications if a secondary surgery is necessary [70]. Thus, thyroid lobectomy and total thyroidectomy are both effective treatment options, which remove pathological tissue, eliminates the risk of recurrence, and treat thyroid nodule autonomy [71-74].

Thyroid surgery leads to significant improvements in HRQoL after treatment [75-78]. However, the interpretation of these results is hampered by mixed populations of patients with toxic and non-toxic goiter, and malignant thyroid disease combined into single groups [75-77]. This makes it impossible to give strong conclusions on the benefits of surgery to patients with benign non-toxic goiter alone (*knowledge gap 1*). Swallowing symptoms are improved within six months after surgery in patients with goiter when examined by patient-reported outcomes

[37, 79-81], but only a little information is available regarding post-surgical objective swallowing function. A significant proportion of patients undergoing surgery persist in having disturbances in esophageal motility in the weeks after surgery [36], but no long term results are available (*knowledge gap 2*). A few studies have shown improved inspiratory airflow following thyroidectomy [58, 82]. However, the correlation between goiter volume, degree of tracheal compression, and/or deviation, and airflow is ambiguous. Further, the impact of these factors on HRQoL is unknown (*knowledge gap 3*) [83].

The main caveats for surgery are the costs, required hospitalization, risk of postoperative hypothyroidism, hypoparathyroidism, and unattended voice changes [61, 84, 85]. Most of these risks are specific for surgery, so while effective, surgery does have an unfavorable risk profile in comparison to non-surgical interventions [86-93]).

NON-SURGICAL TREATMENT MODALITIES

Non-surgical treatment modalities have been present for many years. As previously mentioned, seaweed has been used for treating goiter, however, at present primarily radioiodine (¹³¹I), laser thermal ablation (LTA), and radiofrequency ablation (RFA) are used, as Levothyroxine (LT₄) is no longer recommended by specialist societies due to poor effect and side-effects [20, 94]. ¹³¹I treatment works by causing damage to DNA, and thereby cellular necrosis and destruction of the follicular cells in the thyroid gland [95]. The first clinical results with ¹³¹I were published in 1946 [96, 97]. ¹³¹I treatment reduces the goiter volume by 30-40% within the first year and by 50-60% within the first five years after treatment [98-106]. Both goiter symptoms and tracheal anatomy improves after ¹³¹I treatment [107]. So far, no data on swallowing function are available.

¹³¹I treatment is hampered by the risk of goiter regrowth, which occurs in 10% of patients, whom needs a second dose for full effect [103, 108]. Recombinant human TSH can increase the uptake of ¹³¹I two- to four-fold [109, 110] and this allows for either a lower dose of ¹³¹I or an increased goiter shrinkage. However, this comes at a cost of an increased risk of side effects compared to the conventional ¹³¹I treatment. The acute side effects of ¹³¹I are thyroiditis, which occurs in 3-10% of patients, esophagitis, and an increase in thyroid volume from inflammation, thus increasing the risk of airway compression. Late developing side effects may occur and these include a 20-60% risk of hypothyroidism [109], a 1% risk of developing autoimmune hyperthyroidism [111], and lastly an unsettled risk of later malignancy, especially in younger individuals [112-114]. Nonetheless, for patients with contraindications to surgery or for those who reject surgical treatment, ¹³¹I is a relevant alternative.

US-guided interventions include both LTA and RFA in addition to other more experimental treatments such as microwave ablation, which will not be discussed any further. These are promising treatment options as they report significantly fewer side effects in comparison to surgery [115, 116].

LTA works by emission of photons from excited atoms within the thyroid tissue, thereby creating a coagulative necrosis around the tip of a laser fiber when inserted into the thyroid gland. This procedure is performed in the outpatient clinic using local anesthesia. A recent systematic review identified a 13-82% reduction in thyroid nodule volume 12 months after treatment [117]. A large retrospective multicenter study which included 1531 patients [115] had a mean nodular volume reduction of 72±11% from baseline [mean 27±24 mL (range 1.4-216 mL) to 8±8

mL (range: not described)] 12 months after treatment. Eighty-three percent of the patients were treated with a single session and small, medium, and large nodules had similar volume reductions. Further, both local symptoms and cosmetic complaints were improved after treatment. Only eight of the 1531 patients (0.5%) experienced major complications, such as transient post-operative voice changes, and no patients required surgical repair or overnight admission, which is a significant advantage in comparison to surgery [85, 115]. Nine patients experienced minor complications such as skin burns and hematoma [115].

RFA works by inducing thermal energy (electromagnetic energy) into the thyroid nodule. This procedure takes advantage of a "moving shot" technique, internally cooled needles, and a transisthmic approach and is considered safe and feasible in solid "cold" nodules. It was performed in the outpatient clinic using local anesthesia and post-treatment nodule volume reductions ranged from 69%-93% 12 months after treatment [117]. However, larger nodules required repeated treatments to secure sufficient volume reductions [118]. RFA is not efficient in cystic nodules due to the need for multiple repeated treatments [119]. A systematic review including 2421 patients with primarily benign non-autonomic nodules with a mean volume of 10.3 mL (range 0.05 – 27.7 mL) [116], reported that 42 patients [1.35 % (95% CI 0.89% - 1.81%)] experienced major complications. Of these, 31 patients had transient voice changes and four experienced permanent voice changes (0.2%). Nodule rupture occurred in four patients, permanent hypothyroidism in one patient, and one patient had a temporary brachial plexus injury.

US-guided interventions are especially relevant for smaller goiters accessible by US and cutaneous needles due to a simpler setup in comparison to surgery and a more favorable risk profile. Nonetheless, one should keep in mind that these results were hampered by an unknown selection bias, and a lack of long-term data. Despite their relevance, no study has investigated the impact of US-guided interventions on essential outcomes after treatment such as HRQoL or surrogate markers such as swallowing function, tracheal anatomy or tracheal airflow.

PROSPECT THEORY

From psychological investigations, humans are known to be risk adverse (Bernoulli [120]), meaning that humans in general avoid excessive risks. However, Kahneman et al. have modified this in their prospect theory [121] by showing that people's perception of risks changes with their present standpoint, e.g. the weight patient places on decisions are not linear with in relation to the actual probabilities of events/complications. Low probabilities are weighed higher than they should be, and vice versa. This is expected to have consequences for patients in their choice of treatment for thyroid disease.

Patients with malignant goiter may have a standpoint, which includes risk of death and this makes them more prone to accept the risks associated with surgical treatment to avoid the risk of death. However, the choice is by no means simple for patients with benign nodular goiter, as further progression of the disease is unknown for the individual patient, making it difficult to balance the risk of treatments to the risks associated with the disease. From prospect theory, patients with benign goiter would naturally chose a treatment option with the least risks and put less weight on treatment effectiveness as the non-surgical interventions do not exclude a later option for surgery.

With the current level of evidence regarding the available treatment options for benign thyroid disease, it is difficult to help

patients make any evidence-based choice of treatment. Surgery is the most used modality with a well-known empirically based effect, but it is also the treatment modality with the highest rates of major complications. Only little is known on the effect of the other treatment modalities on HRQoL. However, as non-surgical treatments are less risky procedures, they are indeed of relevance to patients despite their reduced effectiveness in comparison to surgery.

It is essential for all treatment options, including surgery, to provide relevant information on the benefits and risks associated with the specific treatment. This would allow patients, physicians and surgeons to make unbiased decisions between the available treatment modalities.

KNOWLEDGE GAPS, AIMS, AND OBJECTIVES

The primary indication for treatment of benign nodular goiter is HRQoL. However, as there is very limited information on improvements in HRQoL after surgical and non-surgical treatments, the optimal choice of treatment is a daily challenge for patients, physicians, and surgeons. In addition, how HRQoL is affected by changes in surrogate markers such as esophageal motility, and tracheal anatomy and airflow is also unknown.

We therefore aim to limit the previous knowledge gaps (page 20) regarding impact of thyroidectomy on various aspects in patients with benign nodular goiter. Through three studies, we seek to investigate the following objectives:

Knowledge gap 1 - Objective no. 1:

- To investigate changes in disease-specific QoL following surgical treatment in patients with benign non-toxic nodular goiter.

Knowledge gap 2 - Objective no. 2:

- To examine the influence of thyroid surgery on swallowing symptoms and esophageal motility in patients with benign nodular goiter.

Knowledge gap 3 - Objective no. 3:

- To assess the effect of thyroidectomy on tracheal anatomy and tracheal airflow and correlate this to changes in HRQoL in patients with benign nodular goiter.

METHODS

STUDY DESIGN

This project was a single-center, prospective, longitudinal, self-controlled case-series made in a routine clinical care setting [122]. It was conducted at a tertiary university hospital, Odense University Hospital (OUH), using an unmasked design.

For the first objective, investigating changes in HRQoL, a previously used control group from the general population was included [35]. However, for the remaining objectives an appropriate and feasible control group could not easily be defined. Hence, no control group was recruited, thus making these studies self-controlled case-series. Retrospectively, the optimal design would have been a comparative design with a control group of patients undergoing parathyroid surgery. As a consequence of the design, we will not give conclusions regarding the causal relationships between surgery and outcome.

INCLUSION AND EXCLUSION CRITERIA

The inclusion and exclusion criteria for the three studies are listed below.

Inclusion criteria

- Symptomatic benign nodular goiter undergoing thyroid surgery
- Sufficient ability to read and understand Danish

Exclusion criteria

- Previous surgery to the neck
- Present cancer or suspicion of cancer
- Age below 20 or above 80 years
- Neuromuscular disease including diabetes mellitus
- Preoperative overt hypo- or hyperthyroidism
- Paper I (HRQoL): Preoperative subclinical hypo- or hyperthyroidism
- Paper II (MRI): Isthmectomy, pacemaker, or pregnancy

The impact of the inclusion and exclusion criteria will be further described and discussed in section 4.4 Generalizability.

PATIENT RECRUITING STRATEGY

We designed the patient recruiting procedure according to best practice reports for randomized clinical trials (RCTs) [123-128]. The inclusion period ran from November 1 2014 to April 30 2016 and was of sufficient length to adjust for seasonal and temporal factors. Recruitment of patients followed these principles:

- The information of patients took place in the outpatient clinic immediately after the consultation with the thyroid surgeon. This was undertaken by the PhD fellow to minimize disruption to the normal clinic, a well-known barrier to inclusion [125, 126].
- The PhD fellow was chosen as the main recruiter of patients, as senior personnel would increase cost, but not the recruitment rate [127].
- A monetary incentive of Dkr 150.00 per additional visit was provided to increase the recruitment rate [124, 127]. In addition, transportation costs were covered to minimize selection bias according to socioeconomic status. Substantial effort was put into creating an atmosphere for which enrolment did not require payment [123, 127].
- Patients who could not be informed in the outpatient clinic were contacted by phone after the consultation with the thyroid surgeon [128]. Before the phone-call, information material including a health specific questionnaire was mailed to the patient, as this is known to increase the recruitment rate [127]. A "teaser" was written on the envelope front including mentioning of university sponsorship as this increases recruitment rate and reduced loss to follow up [129].
- All examinations were scheduled around patient preferences and patients received reminders for upcoming visits. Waiting time between examinations was kept at a minimum, but, if necessary, it was used for the completion of questionnaires [123].
- We maintained a consistent focus on developing a strong patient – principal investigator relationship to increase the likelihood of return-to-follow up [123].

IMPLICATIONS OF STUDY FLOW

Patients had the possibility to participate in four different studies (Figure 1). These included critical outcomes (post-operative voice changes), essential outcomes (change in HRQoL), and surrogate

outcomes (changes in esophageal motility, tracheal anatomy and airflow).

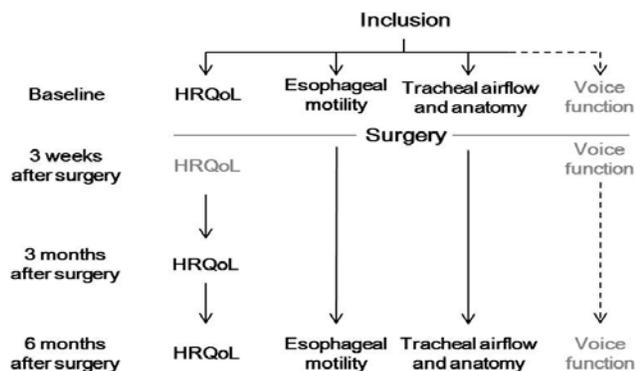


Figure 1. Flow diagram includes all four studies originally intended in the PhD project. Study 1 investigated changes in health-related quality of life (HRQoL) before and after thyroidectomy. Study 2 examined esophageal motility in relation to thyroidectomy. Study 3 assessed changes in tracheal anatomy and airflow after surgery. Study 4 examined voice function before and after surgery. Items marked with grey are not included in the thesis.

This thesis only covers three of the studies, as we await the remaining data on postoperative voice changes following thyroidectomy. If patients participated in all studies this involved a two-to-three hour examination program at baseline, a one-to-two hour program three weeks after surgery, answering a HRQoL questionnaire by mail three months after surgery, and a two-to-three hour examination six months after surgery. The examinations were scheduled late in the afternoon to minimize waiting time. Clinical data including biochemical measurements were obtained by chart review.

Patients followed the regular post-thyroidectomy follow up program including a consultation with a surgeon 10-12 days after surgery and an endocrinologist 2-3 months after surgery. In addition, to the study program and regular follow-up regime, patients were allowed to contact the PhD fellow at all times. All participants received a more intensive follow-up regime, including faster referrals to additional specialists than non-participating patients. A follow-up period of six months was chosen to account for the majority of the post-operative tissue healing phase, but balanced to reduce loss-to follow up [130].

HEALTH-RELATED QUALITY OF LIFE

Only a few options existed for assessing changes in HRQoL after thyroidectomy. The Short Form (36) Health Survey and the Euro-QoL five dimensions questionnaire were the most widely known generic questionnaires. These instruments covered broad categories of functioning and well-being and could be used to compare the relative burden of disease across various conditions. However, both questionnaires have limited sensitivity for evaluating changes in HRQoL after treatment of goiter [131, 132]. By contrast, disease specific questionnaires emphasize symptoms, functioning, and the patient’s perception of a narrowly-defined disease. Only the thyroid-specific patient-reported outcome measure (ThyPRO) was validated for the investigation of HRQoL in goiter patients [133-138]. This was more sensitive for examining changes in HRQoL than the generic questionnaires in this patient group [35] and was therefore used solely in this study.

The ThyPRO was a thoroughly validated questionnaire including evaluation of validity, reliability [133], responsiveness [139], cross-cultural effectiveness [135], and item misfits [136, 137]. It consisted of 85 items covering physical symptoms specific

to benign thyroid disease, but also non-specific effects related to impaired daily life. It measured HRQoL in the most recent four weeks [140-142]. The items were combined into 14 scales including one single item scale covering Overall QoL (Figure 2). Each scale was linearly transformed to a scale from 0-100 points, with a higher score indicating more severe symptoms. After the initiation of the project, an electronic version of the ThyPRO have become available [143].

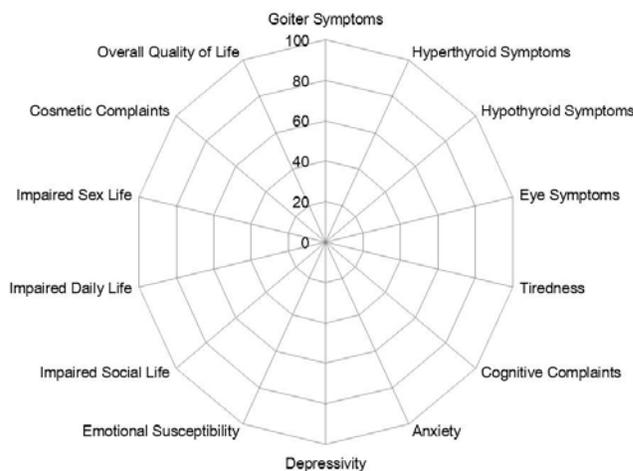


Figure 2. The results from the thyroid-specific patient-reported outcome (ThyPRO) scale are often presented in a radar plot to allow faster interpretation of results and changes in scores over time. Scale scores range from 0-100 points, with 100 points indicating most severe symptoms.

In cooperation with the developer of the ThyPRO Torquil Watt (TW) the eight items covering Graves orbitopathy (Eye Symptoms Scale) were removed to increase relevance for patients and to reduce answer times (to approximately 10 mins.), thereby seeking to increase the return rate of the questionnaire [144]. After our study was initiated, a shorter version of the ThyPRO questionnaire with only 39 questions have been developed based on the presence of mis-fitting items [145] This had a shorter completion time. However, the new version was not validated according to the standards of the original ThyPRO.

We chose to utilize the baseline questionnaire immediately after patients had consented to thyroid surgery. It is unknown whether this had an impact on the individual scale scores of the ThyPRO. The patients might have answered differently, had the questionnaire been answered at the referral to surgery or included in the letter with the time for the outpatient surgical consultation. Only a new study could determine this impact on the ThyPRO scale scores.

Patients completed the questionnaire before surgery, three weeks after surgery, three months after surgery, and six months after surgery in order to assess various time points in the post-operative healing phase. The ThyPRO questionnaire three weeks after surgery was excluded after the completion of the study, as we failed to recognize the ThyPRO specifically measures the HRQoL in last four weeks, wherefore we have data-overlap at three weeks after surgery with the week before surgery.

During the study we assessed if there was difference in scores between patients participating only in the ThyPRO questionnaire study and patients participating in additional studies (esophageal motility, tracheal anatomy and airflow, and voice function). Such differences would have implied a caring bias; however, we could not identify such an effect.

SUBJECTIVE SWALLOWING COMPLAINTS

The persistent motility disturbances observed after thyroidectomy might be explained by postoperative fibrotic tissue formation or damage to the vagal nerve [36]. A patient-reported outcome in combination with objective analyses would be optimal to sufficiently assess the impact of these factors on patients.

Only a few questionnaires have been validated to assess the impact of swallowing dysfunction on HRQoL. For oropharyngeal dysphagia the Deglutitive Handicap Index, Dysphagia Handicap Index, M.D. Anderson Dysphagia Inventory, and SWAL-QOL were available [146]. Of these only the SWAL-QOL had been extensively validated [146], however, it had not yet been translated into Danish. Hence, we therefore used the Goiter Symptom scale of the ThyPRO questionnaire as an alternative swallowing outcome as it included five items specific for swallowing (Figure 3).

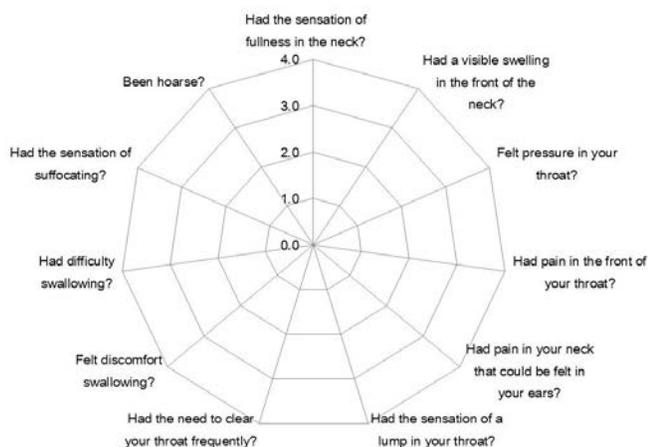


Figure 3. A radar plot of the individual items of the Goiter Symptom Scale in the thyroid-specific patient-reported outcome (ThyPRO) questionnaire. The items scores range from 0-4 points, with 4 point indicating severe symptoms. Each item starts with: "During the last four weeks have you:".

The Goiter Symptom scale was thoroughly validated for patients with goiter and consists of 11 items. Each item starts with the question "During the last four weeks have you". The five items specific for swallowing dysfunction are: 1) Pressure in your throat, 2) Sensation of a lump in your throat, 3) Need to clear your throat frequently, 4) Discomfort swallowing, and 5) Difficulty swallowing. The remaining six items covers other aspects of symptomatic goiter. The items were rated from 0 points (no symptoms) to 4 points (severe symptoms) on a Likert scale.

The choice of using the Goiter Symptom scale instead of the SWAL-QOL questionnaire saved us a large amount of time validating the SWAL-QOL questionnaire. However, this choice might have hampered the validity of our swallowing specific outcome. The ability of the Goiter Symptom scale to estimate swallowing difficulties was assessed by adding three Visual Analogue Scales (VAS) scores to all patients participating in the examination of swallowing dysfunction. The questions were specific for swallowing symptoms.

Question 1: How often do you have a sensation of a lump in your throat?

Question 2: How often do you feel discomfort while swallowing?

Question 3: How often do you have difficulties swallowing?

The VAS scores ranged from 0-100 points (0-100 mm) with a score of 0 points equaling "never" and one of 100 points

equaling "always". After completion of the study the VAS scores were correlated with the Goiter Symptom scale and we found a high correlation ($r > 0.60$) between these scores and the Goiter Symptoms scale ($p < 0.001$).

SWALLOWING FUNCTION

To objectively assess swallowing function in patients with goiter, barium esophagogram, impedance measurements, and high resolution esophageal manometry (HREM) were available. The most clinically relevant objective parameter was the bolus transit. This could be assessed by the barium esophagogram or by impedance measurements. The barium esophagogram was non-invasive, could be performed quickly, and illustrated esophageal anatomical deformities well [147]. Impedance measurement was invasive, but provided physiological information of the bolus transit in contrast to imaging alone [147]. However, neither of these could evaluate the esophageal motility which could be affected pre-surgically by external compression from the goiter, and post-surgically by vagal nerve affection or fibrotic tissue formation. Hence only HREM was considered for the present study. HREM was an invasive catheter based examination with 36 pressure sensors (each consisting of 12 individual micro transducers) capable of recording rapid pressure changes from the esophageal contraction around the sensors [148]. The sensors were spaced one cm apart, thereby covering the entire length of the esophagus from the pharynx to the stomach after insertion of the catheter through the nose. The data were plotted in a three-dimensional color coded topographic plot, with distance between sensors plotted on the y-axis, time on the x-axis (seconds), and pressure amplitude on the z-axis (marked with different colors) (Figure 4). The esophageal motility was measured from the esophageal contraction around individual sensors. This allowed for rapid real-time visual interpretations of the swallowing function.

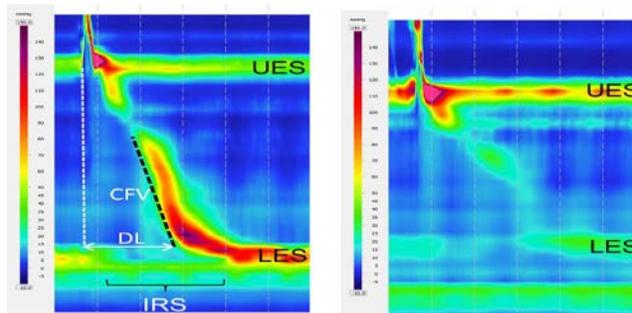


Figure 4. Left image shows a wet swallow with normal esophageal motility (A). Right image shows a wet swallow in a patient with a minor disorder of peristalsis characterized by $\geq 50\%$ ineffective swallows (B). Upper esophageal sphincter (UES), lower esophageal sphincter (LES), integrated relaxation pressure (IRS), contractile front velocity (CFV), and distal latency (DL)

Patients fasted for five hours before examination as the insertion of the catheter was associated with some discomfort and a risk of aspirating gastric content.

With the catheter correctly placed, patients completed 15 wet swallows, each of 5 mL water with 20-30 seconds between the swallows. For standardization, all examinations were performed in the supine positions [43].

A few parameters were of special relevance to esophageal manometry and for evaluating the impact of thyroidectomy (Figure 4). Important esophageal landmarks were the upper esophageal sphincter (UES) of which both the resting and contraction pressure were of interest. Also, the lower esophageal sphincters

(LES) resting pressure and relaxation were essential parameters. Relaxation could not be measured directly, so it was measured by the integrated relaxation pressure (IRP) (based on the LES pressure after initiation of a swallow in the UES [149]). To evaluate the propagation of the esophageal pressure waves, both the contraction front velocity (CFV), a measure of the peristaltic velocity in the smooth muscle esophagus, and the distal latency (DL), a measure of the time from opening of the UES to termination of the esophageal pressure wave, were of relevance [149]. Lastly, the mean contraction amplitude of the smooth muscle esophagus was measured by the distal contractile integral (DCI) [149].

The PhD fellow performed all HREM procedures independently. Before initiation of the study, the fellow shadowed consultant Søren Kruse-Andersen (SKA), who was experienced in the procedure, for four days. All procedures were thereafter performed independently, with the possibility of guidance by SKA. A few of the initial patients were excluded due to inadequately performed procedures e.g. insufficient saved files, missed calibration, and equipment malfunction.

After completion of all examinations the data were analyzed blindly in cooperation with MD Simone Markøw (SM), who previously had performed HREM examinations and interpretations independently as a laboratory technician. Initially, the ten most representative swallows were identified. Then each swallow was examined individually and relevant structures marked. Finally the software calculated mean values based on the ten swallows. The interpretation of the esophageal motility was conducted according to the Chicago classification v. 3.0 an international classification system of esophageal motility disorders [150, 151].

The HREM catheters have shown some challenges with reproducibility [44] as a pressure drift of sensors have been observed in the upper esophagus [152]. The pressure drift was seen during the catheter's life span of 200 examinations. We performed 75 examinations over a two-year period, not near to the catheter's maximum capacity, but it is unclear if the pressure drift affected the results at the six months follow-up. To limit the impact, the same catheter was used in all patients.

AIRFLOW LIMITATIONS

Limitation in tracheal airflow is considered an important physiological parameter to assess in patients with goiter as UAO might be underreported [45]. However, how to assess this was worth a few thoughts. A large goiter might cause UAO affecting the larynx and extra- and intrathoracic portion of the trachea. According to Poiseuille's law, air flow is fourfold proportional to tracheal radius, thus any compression of the trachea would lead to a substantial change in airflow. However, this can be modified, as the tracheal rings in particular are at risk of collapsing during inspiration from the negative transmural pressure gradient across the tracheal wall. This is less critical during expiration with pre-stenotic (i.e. intrathoracic) air pressure being above atmospheric pressure.

A regular pulmonary function test was vital for assessing airflow in patients with goiter. However, it needed to be extended with measurements of inspiratory airflow to fully assess the impact of goiter. A flow-volume loop (FVL) covered both expiratory and inspiratory airflow and consisted of a rapid full inspiration, immediately followed by an expiration of maximum force until no more air could be expired. This was followed by a quick maximum inspiration [153] (Figure 5). Measures of expiration [peak flow (PEF), forced expiratory volume in one second (FEV_1), and forced expiratory flow at 50% of forced vital capacity (FVC) ($FEF_{50\%}$)] could be assessed in combination with a measure of inspirations

[forced inspiratory flow at 50% of FVC ($FIF_{50\%}$)]. It had been suggested that the ratios $FEF_{50\%}/FIF_{50\%}$ and FEV_1/PEF could reflect the presence of UAO in patients with goiter.

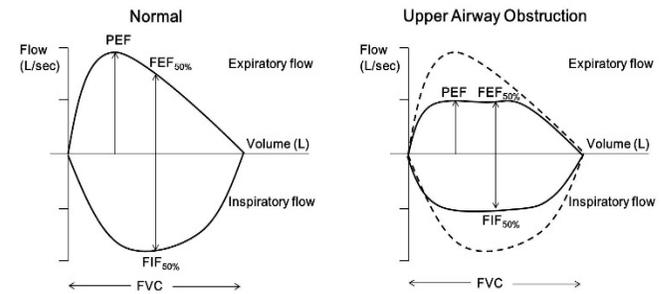


Figure 5. The left illustration is a normal flow volume loop curve, whereas the right illustration shows a patient with upper airway obstruction with flattening of expiratory flow and inspiratory flow. Peak flow (PEF), forced vital capacity (FVC), Forced expiratory flow at 50% of FVC ($FEF_{50\%}$), and Forced inspiratory flow at 50% of FVC ($FIF_{50\%}$).

This was a standardized procedure with patients positioned comfortably in a chair with a straight back and a slightly extended neck. The procedure required careful instructions and some training to perform it correctly. Each patient performed the procedure a maximum of six times, and the three best procedures were used for further analyses. As the examination required maximum effort from the patient there was a limit on how many times the patient could perform the examination to avoid exhaustion and reduced data quality.

To secure sufficient training of the PhD fellow, he was paired with MD Jeppe Faurholdt Lauridsen (JFL) who was trained in the procedure. They performed the first 25 procedures in cooperation, after which the PhD fellow would perform the procedures independently. No analyses of inter- and intra-observer variation between JFL and the fellow were undertaken.

TRACHEAL ANATOMY

To understand the limitations seen airflow in patients with goiter, the goiter volume and its relation to the trachea were important anatomical parameters to assess. US, computed tomography (CT), and magnetic resonance imaging (MRI) could all be used for visualization of the thyroid and trachea [39, 40, 52, 54, 59, 154, 155]. Plain radiographic images of the trachea were considered obsolete, due to poor correlation to airflow and CT determined tracheal cross-sectional area [154].

Despite US being used for measuring thyroid volume, thyroid nodularity, and the thyroid relation to trachea in daily clinic, it was insufficient to show tracheal compression in comparison to both CT and MRI [156, 157]. Only MRI had a good delineation of surrounding structures. However, the disadvantages of MRI in comparison to the other modalities were: the confined space, high noise level, long scanning time, and forceful magnetic field which excluded patients with pacemakers.

Experienced radiographic technicians performed all the MRI scans, whilst the PhD fellow performed all measurements of thyroid and tracheal parameters. Prior to performing measurements, the PhD fellow had a training session with consultant Peter Bøgeskov Andersen (PBA) and they evaluated ten pre- and postoperative scans for incongruence before making the final measurements.

Each patient had thyroid volume, tracheal volume, greatest thyroid encirclement of trachea, greatest tracheal deviation from the midline, smallest cross-sectional area of the trachea

(SCAT), cross-sectional area of the trachea two cm above the carina (CCAT), and the tracheal narrowing in percent ((CCAT-SCAT)/CCAT) measured (Figure 6).

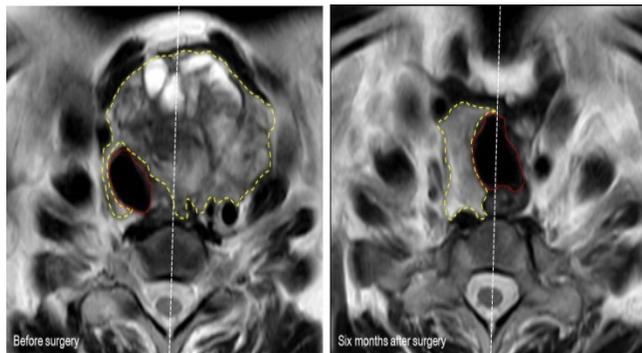


Figure 6. MRI images before and after surgery for a patient with a goiter of 125 mL. Red dotted line illustrates the tracheal cross-sectional area, white full line illustrates the midline, and yellow dotted line shows the thyroid cross-sectional area. Cross-sectional areas can be multiplied with slice thickness to reach thyroid and tracheal volume.

The PhD fellow was judged capable of assessing the tracheal parameters independently from his previous research using microCT for bone volume measurements [158]. No blinding of the MRI scans was applied, as the estimations of tracheal and goiter volume required continuous comparison of pre- and post-thyroidectomy images. This might have introduced bias in the results as the PhD fellow was by nature not blinded to the purpose of the study. In addition, inter- and intra-observer measurement analyses were not performed, as MRI had previously shown good properties for the assessment of goiter volume [155].

DATA PRESENTATION AND DATA RELIABILITY

Cohen's effect size (ES) was used to present the strength of post-thyroidectomy changes as a substantive effect on a scale with little variation, i.e. a small standard deviation was likely to have a greater clinical impact than on a scale with large existing variation [144, 159]. This was reflected in ES statistics (mean change of pre- and post-surgical observations divided by standard deviation at baseline) in contrast to statistical significance. According to Cohen et al. ES of 0.2-0.5 was defined as a small change, 0.5-0.8 as a moderate change, and a values > 0.8 as a large change [144, 159]. This standardized change could be compared across studies.

To increase data reliability of the 338 variables collected for this study, data collection was set-up in the Research Electronic Data Capture (REDCap) [160] program developed by the Vanderbilt University Medical Center. REDCap was a web-based application designed to support data capture for research studies. We used both value limitations and double data entry on all variables to secure a reliable dataset reflecting the actual study results [161].

RESULTS

STUDIES AT A GLANCE

The results will be introduced as a summary with the full results presented in the individual papers. Initially, the study design, participants, and methods used are described (Table 1). This is followed by a presentation of internal and external validity and a Statistical Overview. Lastly selected results from Papers 1 - 3 are presented (Tables 2 - 4).

Participants	Cohort	Methods
<p>Paper I HRQoL</p> <p>Study group: 106 patients with benign nodular goiter (86% females). Age: 51 ± 13</p> <p>Control group: 739 participants from the general population (81% females). Age: 50 ± 16</p>	<p>Prospective cohort study</p> <p>Study group: Continuously included patients with non-toxic goiter undergoing isthmectomy, thyroid lobectomy or total thyroidectomy.</p> <p>Control group: Participants from the Danish civic registry.</p>	<p>Study group: ThyPRO questionnaire before, three months after, and six months after surgery.</p> <p>Control group: ThyPRO questionnaire sent by mail and answered once [35].</p>
<p>Paper II HREM</p> <p>33 patients with benign nodular goiter (74% females). Age: 60 ± 12</p>	<p>Self-control-led case-series</p> <p>Continuously included patients with non-toxic, subclinical hypothyroid, and subclinical hyperthyroid goiter undergoing isthmectomy, thyroid lobectomy or total thyroidectomy.</p>	<p>Goiter Symptom scale of the ThyPRO questionnaire and HREM performed at baseline and six months after surgery.</p>
<p>Paper III MRI</p> <p>65 patients with benign nodular goiter (64% females). Age: 55 ± 13</p>	<p>Self-controlled case-series</p> <p>Continuously included patients with non-toxic, subclinical hypothyroid and subclinical hyperthyroid goiter undergoing thyroid lobectomy or total thyroidectomy.</p>	<p>MRI and FVL and selected scales from the ThyPRO questionnaire at baseline and six months after surgery.</p>

Table 1. Overview of study design, participants, and methodology in the included papers.

Health-related quality of life (HRQoL), Thyroid-specific patient-reported outcome (ThyPRO), Grade of recommendation according to Centre for Evidence-Based Medicine (*), Oxford [162], High resolution esophageal manometry (HREM), Magnetic resonance imaging (MRI), and Flow volume loop curve (FVL).

INTERNAL VALIDITY

Paper I HRQoL

- HRQoL assessed by a thoroughly validated patient reported outcome (ThyPRO).
- Study group compared with a large control group from the general population.
- All questionnaires answered without supervision from health care personnel.
- A linear mixed model used to account for patient specific variation.
- Caring bias analyzed, and not present.

Paper II HREM

- The Goiter Symptom scale is not specifically validated for assessing swallowing symptoms.
- No control group included.
- Manometric data analyzed blind by experienced personnel.

- No use of a supplementary technique to confirm HREM findings.

Paper III MRI

- Appropriate choice of image modality.
- No control group recruited.
- No blinding or inter-/intra observer variation performed.
- FVL examinations performed in cooperation with experienced personnel.
- Inclusion of a validated HRQoL questionnaire

EXTERNAL VALIDITY

Paper I HRQoL

- 106 of 243 patients (44%) seen in the outpatient clinic with euthyroid goiter completed the study.
- 739 of 1200 participants (62%) from the control population answered the ThyPRO once.
- The inclusion rate of the available patients after application of exclusion criteria was 89% with a drop-out rate of 10% six months after surgery.

Paper II HREM

- 33 of 271 patients (12%) with nodular goiter seen in the outpatient clinic completed the study.
- The inclusion rate after exclusion criteria was 29% of available patients with a drop-out-rate of 20% at six months after surgery.

Paper III MRI

- 65 of 267 patients (24%) with nodular goiter seen in the outpatient clinic completed the study.
- The inclusion rate after exclusion criteria was 52% with a drop-out rate of 13% six months after surgery.

SAMPLE SIZE CALCULATION AND STATISTICAL METHODOLOGY

Paper I HRQoL

- The sample size of the study group was calculated at 104 patients based on a paired sample. This is based on a smallest clinically relevant difference between ThyPRO scale scores of five points with a standard deviation of 18 points [35].
- A non-paired t-test was used to compare ThyPRO scores between the study group and control group at baseline and six months after surgery.
- A linear mixed model was used to assess the impact of thyroidectomy on ThyPRO scores in the study group with each patient having three measurements (baseline, three months, and six months after surgery).
- A multiple regression analysis was used to evaluate the impact of age, sex, weight of the removed goiter, and previous radioiodine treatment on ThyPRO score changes in the study group.

Paper II HREM

- The power calculation was based on improvements in UES dysmotility after surgery, as examined by water manometry [36]. Preoperatively 25% of patients had UES dysmotility, and this was reduced to 6% 30-45 days after surgery [36]. Using this difference 30 patients were required to complete the investigation.

- The Wilcoxon signed rank test was used to assess change in the Goiter Symptom scale of the ThyPRO questionnaire due to non-transformable skewness.

- A paired t-test was used to compare the HREM data before and six months after surgery.
- A multiple linear regression analysis was used to analyze the impact of age, sex, weight of the removed goiter, and surgical procedure on motility parameters and to assess the motility parameters' influence on Goiter Symptoms.

Paper III MRI

- The sample size was calculated at 58 patients, based on a paired sample with changes in SCAT of $0.15 \pm 0.4 \text{ cm}^2$ after treatment with ^{131}I [100, 104, 163].
- A paired t-test was used to compare tracheal, respiratory, and HRQoL parameters before and six months after surgery.
- A multiple linear regression analysis was used to evaluate the impact of age, sex, and thyroid volume on preoperative tracheal parameters.
- A multiple regression analysis was used to evaluate the correlations between tracheal narrowing, FVL, and HRQoL parameters using age and sex as independent variables.

Paper I (HRQoL)

Study group	Before surgery	Six months after surgery	p-value ^a	Effect size
Goiter Symptoms	40 ± 21	8 ± 10	<0.001	1.52
Hyperthyroid Symptoms	20 ± 17	12 ± 13	<0.001	0.48
Hypothyroid Symptoms	21 ± 20	14 ± 17	<0.001	0.35
Tiredness	50 ± 26	33 ± 24	<0.001	0.60
Cognitive Complaints	21 ± 21	14 ± 19	<0.001	0.36
Anxiety	20 ± 21	8 ± 13	<0.001	0.54
Depressivity	27 ± 21	17 ± 15	<0.001	0.42
Emotional Susceptibility	29 ± 23	20 ± 21	<0.001	0.37
Impaired Social Life	6 ± 13	5 ± 13	0.52	0.08
Impaired Daily Life	12 ± 19	7 ± 17	0.008	0.28
Impaired Sex Life	13 ± 24	12 ± 23	0.38	0.08
Cosmetic Complaints	20 ± 22	9 ± 14	<0.001	0.48
Overall Quality of Life	35 ± 31	12 ± 25	<0.001	0.74
Control group	Cross-sectional score		p-value ^b	p-value ^c
Goiter Symptoms	5 ± 9		<0.001	<0.001
Hyperthyroid Symptoms	12 ± 14		<0.001	0.76
Hypothyroid Symptoms	14 ± 16		<0.001	0.99
Tiredness	35 ± 21		<0.001	0.35
Cognitive Complaints	14 ± 17		<0.001	0.90

Anxiety	13 ± 16	<0.001	0.005
Depressivity	21 ± 18	0.01	0.03
Emotional Susceptibility	23 ± 19	0.005	0.17

Table 2: Selected results (mean ± SD) from Paper I (HRQoL). Study group (n=106) and control group (n=739). [164] Study group baseline vs. 6 months after surgery using a linear mixed model (a). Control group vs. study group baseline using a non-paired t-test (b). Control group vs. study group at 6 months after surgery using a non-paired t-test (c). Changes with an effect size of 0.2-0.5 were defined as small, values 0.5-0.8 as moderate, and values > 0.8 as large. Health-related quality of life (HRQoL).

Short description of results

The study group had a reduced HRQoL at baseline in all assessed ThyPRO scales in comparison to the control group (p-values and ES seen in Table 4). The study group experienced pronounced improvements in HRQoL in the first three months after surgery. A large improvement was observed in the Goiter Symptom scale at both three months and six months after surgery. Overall QoL improved moderately at three months, but improved additionally at six months after surgery. The remaining scores for Hyperthyroid Symptoms, Hypothyroid Symptoms, Tiredness, Cognitive Complaints, Anxiety, Depressivity, and Cosmetic Complaints demonstrated only small improvements at three months after surgery. The Tiredness and Anxiety scales progressed to moderate improvements six months after surgery.

Six months after surgery, the study group had more Goiter Symptoms, but less Anxiety in comparison to the control group. There was no difference between the remaining scales and the control group.

Using multiple linear regression analyses, increasing age was associated with less relief in Goiter Symptoms after surgery (-0.45 points/year, p=0.003), but older individuals also had less Goiter Symptoms at baseline (-0.35 points/year, p=0.03). Female patients experienced more relief in Goiter Symptoms (11 points more than males, p=0.048). However, female patients had more Goiter Symptoms than males prior to surgery (15 points, p=0.007). A goiter above the median size of 35 grams was associated with a greater relief of Cosmetic Complaints (11 points, p=0.003), but these patients also had more Cosmetic Complaints prior to surgery (11 points, p=0.01). No other scales were affected by the goiter weight.

PAPER II (HREM)

Parameters	Before surgery	Six months after surgery	p-value
Goiter Symptom scale	39 (2-61)	5 (0-52)	<0.001
UES basal pressure (mmHg)	70.6 ± 27.7	87.7 ± 43.2	0.04
UES IRP (mmHg)	5.3 ± 7.3	2.8 ± 6.2	0.12
LES basal pressure (mmHg)	23.4 ± 17.0	25.4 ± 18.4	0.59
LES IRP (mmHg)	8.7 ± 11.0	10.5 ± 9.6	0.37
DL (s)	3.5 ± 0.6	3.6 ± 0.8	0.15
CFV (cm/s)	4.1 ± 1.4	4.1 ± 1.9	0.94
DCI (mmHg-cm-s)	1652 ± 1063	1750 ± 1660	0.25

Table 3. Selected results [median (range) and mean ± SD] from paper II (HREM) (n=33). [165] High resolution esophageal manometry (HREM), Upper esophageal sphincter (UES), Integrated relaxation pressure (IRP), Lower esophageal sphincter (LES), Distal latency (DL), Contractive front velocity (CFV), and Distal contractile integral (DCI). Effect sizes of 0.2-0.5 were defined as small changes, 0.5-0.8 as moderate changes, and values > 0.8 as large changes.

Short description of results

The Goiter Symptom scale showed a large improvement six months after surgery compared to baseline (p-values and ES seen in Table 5). All motility parameters were within the limits of normal swallowing before surgery [166], and only the UES basal pressure increased after surgery (p=0.04). The remaining motility parameters, i.e. DL, UES IRP, CFV, and DCI showed no significant change.

One patient (3%) fulfilled the criteria for having a minor peristaltic disorder, defined as >50% weak or failed swallows before surgery, while two patients (6%) fulfilled these criteria after surgery. No patient had major peristaltic disorders either before or after surgery.

With regression analyses we demonstrated a significantly larger reduction in UES basal pressure (9.6 mmHg, p=0.047), peak pressure 4 cm below the UES (32.4 mmHg, p=0.03), and of the DCI (1306 mmHg-cm-s, p=0.04), in patients undergoing total thyroidectomy in comparison to patients undergoing hemithyroidectomy or isthmectomy after surgery.

No correlations between the pre- and postoperative Goiter Symptom scores and motility parameters, presence of motility disturbances, or goiter weight, were present.

PAPER III (MRI)

Parameters	Before surgery	Six months after surgery	p-value	Effect size
Tracheal narrowing (%)	39 (0-90)	29 (0-54)	<0.001	0.67
Tracheal deviation (mm)	12 ± 7	8 ± 4	<0.001	0.61
SCAT (cm ²)	1.5 ± 0.5	1.8 ± 0.5	<0.001	0.73
PEF (L/min)	7.5 ± 2.2	8.0 ± 1.9	0.03	0.19
FIF _{50%} (L/min)	4.8 ± 1.4	5.3 ± 1.5	<0.001	0.32
Goiter Symptoms	40 ± 21	10 ± 11	<0.001	1.37
Tiredness	48 ± 25	35 ± 24	<0.001	0.52
Anxiety	20 ± 21	7 ± 12	<0.001	0.59
Impaired Daily Life	13 ± 19	7 ± 16	0.48	0.29

Table 4. Selected results [median (range) and mean ± SD] from paper III (MRI) (n=65). [167] Magnetic resonance imaging (MRI), Smallest cross-sectional area of the trachea (SCAT), Peak flow (PEF), and Forced inspiratory flow at 50 % of forced vital capacity (FIF_{50%}). Effect sizes of 0.2-0.5 were defined as small changes, 0.5-0.8 as moderate changes, and values > 0.8 as large changes.

Short description of results

Tracheal narrowing, tracheal deviation, and SCAT experienced moderate improvements six months after surgery (p-values and ES seen in Table 6). Tracheal narrowing, tracheal deviation, and SCAT at baseline correlated strongly with the initial goiter volume (p<0.001). After surgery, tracheal narrowing, tracheal deviation, and SCAT improved moderately, and tracheal volume underwent a small increase. Improvements in tracheal parameters all correlated strongly with the reduction in thyroid volume.

Postoperatively, a small, but highly significant, improvement was found in FIF_{50%}. Also small significant declines in the FEF_{50%}/FIF_{50%} (ES=0.29, p=0.001) and FEV₁/PEF (ES=0.39, p=0.002) ratios were observed which might reflect UAO. Tracheal narrowing correlated with the initial respiratory parameters. A 1% increase in tracheal narrowing led to a decrease of 0.03 L/min in PEF (p=0.003) and a 0.02 L/min decrease in FIF_{50%} (p=0.01).

From selected scales of the ThyPRO questionnaire, the Goiter Symptoms score experienced a large improvement six months after surgery, whereas the Tiredness scale changed moderately. The Anxiety scale also had a moderate improvement, but the change in the Impaired Daily Life scale was insignificant. However, using multiple regression analyses, the improvement in tracheal narrowing correlated positively with improvement in the Impaired Daily Life scale with a coefficient of 0.33 point improvement per 1% decrease in tracheal narrowing ($p=0.03$). The baseline tracheal narrowing and inspiratory flow did not correlate with pre-treatment HRQoL.

DISCUSSION

This PhD thesis addresses two overall themes not previously extensively investigated: first, the capability of thyroidectomy to relieve symptoms of goiter and second, the impact of thyroidectomy on anatomical/physiological parameters commonly affected by goiter i.e. the esophageal compression and swallowing function and trachea narrowing and airflow. Both themes are original and having news value for the thyroid scientific community.

This section will primarily focus on strengths and limitations to the individual studies. For a thorough discussion of the results in relation to existing literature it is referred to the individual papers. This is to avoid duplication of text in thesis and paper.

THYROIDECTOMY AND SYMPTOM RELIEF

Patients with goiter experienced profound improvements in symptoms of goiter both three and six months after thyroidectomy. These gains, ranged from small to large on most of the 13 ThyPRO scales. Only the Sex Life and Social Life scales did not improve after surgery. At six months after surgery, all scores in the study group were close to the scores of the general population.

Three factors support the view that these results are trustworthy. First, the majority of improvements took place during the first three months after surgery wherefore symptom relieve was closely related to the surgical procedure and not to any later unknown effect. Second, the patient group had scores similar to the general population six months after surgery. This indicates that it is the removal of goiter which was the cause for improvement and not an underlying placebo effect of surgery. Third, other studies with different cohort compositions have also identified improvements in HRQoL after thyroidectomy which does support our findings [75-77]. Further, the use of a thoroughly validated patient-reported outcome measure, answers without supervision from health care personnel, a large control group, and multiple examination points indicates good internal validity of the study. In addition, our finding is representative of the study population, from the inclusion rate of 89% and a loss to follow-up of only 10% six months after surgery.

When using multiple regression analyses, various subgroups scored differently on the ThyPRO questionnaire. Women experienced more symptoms from smaller goiters in comparison to men. The underlying cause for this effect is unknown. It could be speculated that a larger neck in men makes them able to accommodate a larger goiter before presenting symptoms. However, this is pure speculation and the findings could also be associated with an underlying selection bias where men may delay their treatment of goiter longer than women. Younger individuals had more symptoms when compared to older individuals. This might be explained by different perception of disease between age groups. However, there is an increase in tracheal calcification

with age making the trachea of an older individual able better able to withstand external compression of goiter compared to a non-calcified trachea of younger individuals which might explain some of this finding. In addition, the exclusion criteria requiring first time thyroid surgery did in fact lead to the exclusion of older individuals who had previously been operated upon, who might have differed from higher age population who were included. The relatively poor correlation between weight of the goiter and the symptoms presented by patients was also of interest. The size of the goiters was only associated with cosmetic complaints indicating that the symptoms from goiter are not directly related to goiter volume and might instead be caused by confounding factors such as local esophageal compression or individual differences in perception of disease.

No study has investigated the impact of ^{131}I , LTA or RFA treatment on HRQoL, hence it is not possible to compare the effects of these treatments to surgery. However, a study investigating the impact of treatment i.e. ethanol sclerotherapy, ^{131}I , and surgery combined into one group, identified much less improvement in symptoms than seen in our paper I (HRQoL) [35]. The authors concluded that patients with goiter had persistent health deficits after intervention. We found no evidence for such a conclusion after surgical intervention alone, indicating that surgery is a very effective treatment in relieving symptoms of benign non-toxic goiter.

Limitations

Despite this study being the largest to use the ThyPRO questionnaire to investigate changes in HRQoL following thyroidectomy, it was still underpowered with regard to investigating many interesting aspects of goiter. The results for individual surgeons, the impact of post-operative TSH adjustments, surgical complications, and thyroid volume would all be relevant for future studies and quality control.

Our control group was not specifically recruited for this study and thus differs from the study cohort. The control group had only malignancy as an exclusion criterion and included patients from the age of 18 years, whereas we had a lower age limit of 20 years and used multiple exclusion criteria. Additionally, the response rate of the control group was 62% compared to our rate of 89%, so many control persons were not investigated. This selection bias between cohorts might explain some of the minor differences we observed between the study cohort and the control group six months after surgery. Further, our control group contained no longitudinal information, hence it is difficult to interpret some of the small changes seen in the ThyPRO scores over time.

Another possible bias comes from the Hawthorne effect where the participant's knowledge of being observed has a positive impact on the results of the observed procedure [121]. There was no evidence of a caring bias in this study, but all patients participating in the HRQoL study might have received superior service and treatment throughout the study, as hospital personnel also felt observed. Our results might thus have been more positive than if they had been implemented using blinded hospital staff.

While the questionnaire has been thoroughly validated, there are still considerations relevant for the interpretation of our results. First, we put a major focus on the Goiter Symptoms scale in both Paper I (HRQoL) and Paper II (HREM). However, on this scale four of the 11 items are subjected to differential item func-

tioning, making the results dependent on cohort composition, thereby reducing the generalizability of the results [136]. In addition, 11 items of the ThyPRO were found potentially mis-fitting by confirmatory factor analyses [137]. Despite, these considerations only having a small impact on the final results, the compiler of the questionnaire has recommended these items be removed in a subsequent version (ThyPRO-39) [145].

Implications

- The study provides clinically relevant information for patients, physicians, and surgeons on how symptomatic non-toxic goiter affects patients referred for surgical intervention, and of the expected improvements in HRQoL after surgery. This might improve communication between patients, physicians, and surgeons before and after surgery and align the expectations.
- This study will serve as a relevant comparison group for future studies investigating the effect of ¹³¹I, LFA, and RFA treatment on HRQoL in patients with benign nodular goiter.

THYROIDECTOMY AND SWALLOWING DYSFUNCTION

In addition to the profound improvements in Goiter Symptoms of the HRQoL study we also identified large improvements in the Goiter Symptoms scale in the sub-group of patients examined with HREM. We found no evidence that motility disturbances could explain swallowing difficulties before and after surgery.

Problems with this arm of the study arose from the low number of patients and the lack of a control group, resulting in a poor quality prognostic cohort. The small number of patients and the missing control group were directly associated with the discomfort of the examination which made it difficult to recruit patients and it increased the loss to follow-up. Given the low number of patients, attempts to perform sub-group analyses were not feasible.

Regarding the internal validity of the HREM study, the blind interpretation of the motility examination supports the results. However, the Goiter Symptom scale was not validated specifically for the assessment of swallowing symptoms, giving some limitations in interpreting the findings. Only a direct comparison of this scale to other questionnaires specifically validated for the investigation of swallowing dysfunction could limit this uncertainty. The limited recruiting rate (28% after the exclusion criteria) and a high drop-out rate (20%) makes our study at risk of being underpowered and thus prevents us from identifying real changes in esophageal motility after surgery.

Other studies have also found improved swallowing function after surgery [37, 79-81]. This strengthens our belief in the relevance of the Goiter Symptom scale to measure changes in swallowing function. With respect to motility changes after surgery, only one other study has investigated esophageal motility before and 30-45 days after thyroidectomy using water manometry [36]. No change was found in esophageal motility after surgery, but there were some disturbances in the UES motility which improved after surgery. We found no signs of disturbances in the UES motility in our cohort, but this may have been because of the differences between water manometry and HREM. Our findings could preferably have been confirmed with other techniques investigating bolus transit directly. Although relevant, there is no information available on changes in esophageal motility after ¹³¹I, LTA or RFA treatment.

Limitations

Only 12% of patients with goiter referred for thyroidectomy, before exclusion criteria were applied, were included in the study, which severely limit our ability to generalize from our findings. Additionally, as HREM is a novel technique, many studies are continuously being published to validate the technique under various conditions and much information is still missing regarding the reliability of the examination. As an example, the reference values used for comparison were from a selected group of patients aged 20-49 years [166] which was not necessarily representative for our group of patients with a mean age of 59.9 ± 11.8 years.

We experienced a larger frequency of drop-outs in the first half of the study compared to the latter half. This most likely reflected a learning curve in patient handling and did not affect the results of the individual patients which were analyzed blind. A different setup with more supervision might have prevented some patients from dropping out, and thus increasing the power of the study. The learning curve also led to faster examinations towards the end of the study. This could explain the observed increase in UES basal pressure, as this pressure decreases when patients are allowed a long phase of getting familiar with the catheter. By the end of the study the "getting familiar" phase was shorter than at the beginning of the study. Alternatively there could be pressure drift in the sensors, as explained in the methods section, or on esophageal decompression reducing the need for a high closing pressure.

The use of a five mL water bolus in the supine position to assess swallowing is standardized practice for HREM analyses. This is not physiologically sound as patients rarely lie down when swallowing. However, as the swallowing complaints in patients with goiter often occur when patients are lying down, this might be suitable for examining our patient group. Nevertheless, we might have found more motility disturbances if we had used a more challenging bolus than water, such as apple sauce, cubed bread or marshmallows, but this is pure speculation.

Implications

- Swallowing dysfunction in patients with goiter is likely to be related to factors other than motility disturbances.
- Thyroidectomy appears to be a safe procedure which does not affect esophageal motility negatively.

THYROIDECTOMY AND AIRWAYS

We found substantial improvements in airflow and a large reduction in tracheal compression in patients with large goiters after thyroidectomy. Patients with smaller goiters, i.e. volumes below 100 mL, experienced much less impact on the trachea and airflow. The improvements in tracheal narrowing correlated with improvements in HRQoL.

The results are supported by the following considerations: first, they are biologically plausible as the cartilage rings of the trachea can withstand some external pressure and probably also the pressure from a small goiter. Second, the increased airflow after surgery followed the principles of Poiseuille's law, showing air flow to increase with reductions in tracheal narrowing. Third, other studies have similarly identified improvements in airflow after surgery [58, 82]. The features, in combination with an inclusion rate over 50% and a low drop-out rate of 13%, enable us able to generalize from the study population. The internal validity of the study is acceptable as we used validated measuring modalities such as MRI and the ThyPRO questionnaire.

Our patients experienced significant improvements in HRQoL after surgery. However, only the improvement in the Impaired Daily Life scale was associated with relief in tracheal narrowing, but interestingly not with improvements in airflow. This is puzzling, as the luminal area of the trachea is a surrogate marker of tracheal airflow. However, the ES related to changes in the airflow were small when compared to the moderate ES changes of tracheal anatomy. Hence, the greater variation in airflow parameters lead to smaller ES compared to trachea anatomy data. This might explain the lack of significant association between airflow and HRQoL parameters. These factors indicate that the UAO previously suggested to be severely underestimated [45] might not be that common in patients with goiter and certainly not of great significance to most patients. Many other factors besides the volume of the removed thyroid gland are likely to be involved in the improvements of HRQoL after thyroid surgery.

It is interesting that no patient experienced a substantial increase in tracheal narrowing after surgery indicating that neither fibrotic tissue formation nor tracheomalacia was of great concern in patients undergoing thyroidectomy. Previous reports mentioned an ambiguous correlation between goiter volume and the degree of tracheal compression [83]. However, our data show a clear logarithmical correlation between goiter volume and tracheal narrowing, and thereby provide new knowledge of the basic physiology of goiter.

No study has previously examined the impact of surgery on tracheal anatomy, but several studies using ¹³¹I have shown improvements in SCAT after reduction in goiter volume similar to our study [100, 104, 107, 163].

Limitations

A limitation to our data was the lack of a control group during FVL and MRI evaluations. The small reduction in FVC seen six months after surgery could have been an effect of changes in the FVL instruction across time and may therefore indicate bias. From the study design, these differences could not be controlled for in the analyses. In addition to the missing control group, the lack of blinding could also give rise to bias. In the MRI measurements, the fact that the examiner knew the outcome of the study could have led to overestimation of the study effects by as much as 17% [168-170]. However, as the observed differences in the study were larger than this percentage, this bias likely does not affect our results. Had the images been blinded it might have led to additional bias by making their evaluation less precise [171].

The flow volume loop curve examination was obtained with patients in a seated position. However, this does not reflect the actual compression of goiter very well as the compression is most prominent during activity or when patients lie in the supine position as with the MRI scanning's. Thus the changes in airflow might have been underestimated, thus explaining the small effect sizes of the airflow parameters and the lack of correlation to tracheal anatomy and HRQoL parameters.

Post-surgically, all patients with a total thyroidectomy received LT₄, while this was only the case for 13% of patients undergoing hemithyroidectomy. In a few patients serum TSH took months to normalize and 13% of the patients had serum TSH levels above or below the reference level six months after surgery. It has been speculated that overt and even sub-clinical hypothyroidism, with the need for LT₄ substitution after surgery, might affect respiratory function negatively [172]. We found no evidence to support this claim, as there was no correlation be-

tween pulmonary function parameters and TSH or LT₄ substitution.

Implications

- Patients with a small goiter do not seem to have tracheal compression. This is primarily seen in patients with goiters > 100 mL, which is quite rare in areas with good access to healthcare.
- As the majority of patients with symptomatic goiter referred for thyroidectomy at our facility do not suffer from UAO, it might not be as severely underestimated as previously expected.

GENERALIZABILITY

The inclusion and exclusion criteria limited the number of patients available for examination and thereby reduced the generalizability of our results. The first inclusion criterion requiring patients with symptomatic goiter prior to undergoing thyroidectomy excluded patients with goiters not causing symptoms, patients who rejected surgical treatment, and patients with thyroid malignancy. This introduced a selection bias into the cohort, which makes us unable provide strong conclusions with regard to the correlation between goiter volume and symptoms. The second inclusion criterion requiring patients to read and understand Danish, excluded most patients with an ethnic background other than Danish. This criterion was introduced a few months into the inclusion period as we experienced difficulties in explaining the study protocol and especially the nature of the informed consent for patients, to people with an ethnic background other than Danish. Despite its relevance, this problem is well known in research and difficult to address [173].

The exclusion criteria of: 1) previous surgery to the neck, 2) suspicion of thyroid cancer or present cancer, 3) age below 20 years or age above 80 years, 4) neuromuscular diseases including diabetes mellitus, and 5) pre-operative overt hypothyroidism and overt hyperthyroidism, all limited the generalizability of our results. The overt hyperthyroidism exclusion criterion excluded a few very large goiters with thyroid nodule autonomy, thereby reducing the number of patients with very large goiters. However, this was intended, as thyroid dysfunction can itself have a profound impact on HRQoL [174, 175]. The exclusion criterion of neuromuscular disease including diabetes mellitus was chosen for the HREM examination, as neuromuscular disease can affect HREM results and approximately 51% of patients with diabetes are known to have esophageal motility disturbances. This excluded many patients with multi-morbidity from participating and led to a healthier study population than would otherwise have been recruited [176]. Further, it precluded us from extrapolating our results to patients with diabetes. This was especially problematic as insulin resistance is an independent risk factor for developing goiter [25-31]. Retrospectively, this criterion should have been an exclusion criterion only for the HREM study or preferably not have been used at all, as the study objectives investigated changes in disturbances and not the prevalence.

A few of our patients had insufficiently adjusted TSH six months after surgery. How this influenced the results is unknown, but the improvements in HRQoL might have been underestimated compared to a patient cohort with well-adjusted TSH levels. As thyroid hormones affect all tissue this could have had unknown implications for the HREM and MRI study. However, this is pure speculation and our data does not indicate any impact of this limitation. Throughout the study all TSH variables were collected

from chart reviews, meaning that the day of blood sampling was not necessarily close to the study visits. This limits us from making a strong conclusion on the impact of TSH and post-operative changes in symptoms and physiological parameters.

Lastly our data cannot be transferred to other groups of patient with e.g. toxic nodular goiter, Grave's disease or autoimmune hypothyroid disease, as these were not covered by our inclusion criteria. A separate study could determine the impact of thyroidectomy on these diseases.

Despite the above limitations our study has some distinct advantages. First, we used a group of patients selected for their disease (benign nodular goiter) instead of the procedure (thyroidectomy), which made it easier for patients and doctors to interpret the results in relation to the individual patient. Second, by consecutive recruitment of patients with thyroid nodularity with no regard to the volume of the goiter, our results reflected the actual situation in the clinic with many patients having rather small goiters and some even thyroid nodularity as their main complaint. A larger effect might have been observed if patients with large goiters or patients undergoing total thyroidectomy had been exclusively recruited. Lastly, the study was implemented in a daily clinic without selecting surgeons or changing the follow-up program, hence the results accurately reflected the clinical situation.

CONCLUSIONS

Patients with benign non-toxic nodular goiter experienced a lower HRQoL compared to the general population before surgery. We found strong evidence for the beneficial effect of thyroid surgery for relieving compressive symptoms and returning HRQoL to a level close to the general population. This effect was of a magnitude not easily attributed to confounding factors.

The swallowing difficulties of patients with goiter seemed to be caused by other factors than motility disturbances, as evident from the normal esophageal motility observed in the study cohort before and after surgery. The reasons for the pronounced swallowing symptoms should be sought in other mechanisms affecting the esophagus.

Previously, UAO has been suggested to be severely underestimated in patients with goiter. From our data, UAO seems to be uncommon, especially in patients with smaller goiters. However, in patients with very large goiters the surgical removal of goiter has a profound and positive impact on tracheal compression and airflow, and additionally improves some aspects of HRQoL.

From this project, patients, physicians, and surgeons should gain better knowledge of the impact of goiter and surgical treatment on HRQoL, swallowing dysfunction, and affection of tracheal anatomy and airflow, all of which could be used in the daily consultations.

FUTURE RESEARCH

Future studies requiring recruitment of new patients are indeed of great interest. Here a direct comparison of ¹³¹I, LTA or RFA to thyroidectomy would be a clear benefit for both patients and clinicians. In addition, the impact of non-surgical interventions RFA and LTA needs more attention as these technologies have the potential to disrupt the present practice of performing surgery on patients with very small goiters. Such as study, should be performed with the ThyPRO questionnaire as the primary endpoint, but should preferably also include information on major complications such as post-treatment voice changes. We would suggest

performing this study in patients with goiters below 100 ml, as this is the most prevalent patient group, which rarely suffers from tracheal compression.

A future project constructing an evidence-based shared decision making tool presenting the benefits and risks associated with the various treatments for patients with benign nodular goiter, would provide relevant information to clinicians.

This study did not identify any changes in esophageal motility despite patient complaints of swallowing dysfunction. An obvious next project is to investigate the impact of local esophageal compression from goiter on the Goiter Symptom scale of the ThyPRO questionnaire. The study can be performed on our existing MRI scans, collected ThyPRO data, and would be a novel addition to the literature. The outcome would be a secondary outcome utilizing previous data, but nonetheless relevant for publication

Previous studies have suggested that UAO may be overlooked in patients with goiter if patients have no respiratory complaints [45]. Our findings suggest that UAO is rare in patients with symptomatic goiter. Based on our findings, the previously suggested method to identify UAO in patients with goiter by applying cut-off limits on FVL parameters appears to be an inappropriate way forward [153]

Suggested cut-off limit for identifying UAO:

$$\frac{FEV_1}{PEF} > 8 \frac{mL}{L/\min}$$

[153]

The suggested cut-off limit is based on an old study using mouth-pieces with very small orifices (0.5 cm²) to simulate UAO [177]. These mouth pieces are only comparable to the trachea the very few patients with goiters >300 mL in our study. The existing data from the MRI study might therefore be used to challenge the validity of this cut-off limit, preferably, in corporation with other researchers in possession of tracheal imaging data of patients with goiter.

SUMMARY

Surgery: Is it any Good for Goiter? In patients with goiter the benefits of thyroid surgery have previously rarely been investigated, as only few alternatives existed. However, the increasing evidence of the advantages with non-surgical substitutes with lower costs and preferable risk profiles prompted us to investigate the evidence base for thyroid surgery thoroughly. This thesis consists of three published studies investigating the impact of thyroidectomy on: 1) changes in disease-specific quality of life, 2) swallowing symptoms and esophageal motility, and 3) tracheal anatomy and airflow, in a cohort of patients with benign nodular goiter.

REFERENCES

1. *Werner & Ingbar's the thyroid: a fundamental and clinical text*. 10. ed: Philadelphia, Lippincott Williams & Wilkins, a Wolter Kluwer business; 2013.
2. Hannan SA. The magnificent seven: a history of modern thyroid surgery. *Int J Surg*. 2006;4:187-191.
3. Sakorafas GH. Historical evolution of thyroid surgery: from the ancient times to the dawn of the 21st century. *World J Surg*. 2010;34:1793-1804.
4. Giddings AE. The history of thyroidectomy. *J R Soc Med*. 1998;91 Suppl 33:3-6.
5. Dralle H, Machens A. European endocrine surgery in the 150-year history of Langenbeck's Archives of Surgery. *Langenbecks Arch Surg*. 2010;395 Suppl 1:43-55.

6. Oppenheimer JH, Schwartz HL, Mariash CN, Kinlaw WB, Wong NC, Freaque HC. Advances in our understanding of thyroid hormone action at the cellular level. *Endocr Rev*. 1987;8:288-308.
7. Brent GA. The molecular basis of thyroid hormone action. *N Engl J Med*. 1994;331:847-853.
8. Carle A, Krejbjerg A, Laurberg P. Epidemiology of nodular goitre. Influence of iodine intake. *Best Pract Res Clin Endocrinol Metab*. 2014;28:465-479.
9. Vadiveloo T, Donnan PT, Cochrane L, Leese GP. The Thyroid Epidemiology, Audit, and Research Study (TEARS): morbidity in patients with endogenous subclinical hyperthyroidism. *J Clin Endocrinol Metab*. 2011;96:1344-1351.
10. Hegedus L, Bonnema SJ, Bennedbaek FN. Management of simple nodular goiter: current status and future perspectives. *Endocrine Reviews*. 2003;24:102-132.
11. Rieu M, Bekka S, Sambor B, Berrod JL, Fombeur JP. Prevalence of subclinical hyperthyroidism and relationship between thyroid hormonal status and thyroid ultrasonographic parameters in patients with non-toxic nodular goitre. *Clin Endocrinol (Oxf)*. 1993;39:67-71.
12. Krohn K, Wohlgenuth S, Gerber H, Paschke R. Hot microscopic areas of iodine-deficient euthyroid goitres contain constitutively activating TSH receptor mutations. *J Pathol*. 2000;192:37-42.
13. Tonacchera M, Agretti P, Chiovato L, et al. Activating thyrotropin receptor mutations are present in nonadenomatous hyperfunctioning nodules of toxic or autonomous multinodular goiter. *J Clin Endocrinol Metab*. 2000;85:2270-2274.
14. Laurberg P, Pedersen KM, Vestergaard H, Sigurdsson G. High incidence of multinodular toxic goitre in the elderly population in a low iodine intake area vs. high incidence of Graves' disease in the young in a high iodine intake area: comparative surveys of thyrotoxicosis epidemiology in East-Jutland Denmark and Iceland. *J Intern Med*. 1991;229:415-420.
15. Knudsen N, Bulow I, Jorgensen T, Laurberg P, Ovesen L, Perrild H. Goitre prevalence and thyroid abnormalities at ultrasonography: a comparative epidemiological study in two regions with slightly different iodine status. *Clin Endocrinol (Oxf)*. 2000;53:479-485.
16. Pedersen KM, Laurberg P, Nohr S, Jorgensen A, Andersen S. Iodine in drinking water varies by more than 100-fold in Denmark. Importance for iodine content of infant formulas. *Eur J Endocrinol*. 1999;140:400-403.
17. Rasmussen LB, Jorgensen T, Perrild H, et al. Mandatory iodine fortification of bread and salt increases iodine excretion in adults in Denmark - a 11-year follow-up study. *Clin Nutr*. 2014;33:1033-1040.
18. Vejbjerg P, Knudsen N, Perrild H, et al. Effect of a mandatory iodization program on thyroid gland volume based on individuals' age, gender, and preceding severity of dietary iodine deficiency: a prospective, population-based study. *The Journal of Clinical Endocrinology & Metabolism*. 2007;92:1397-1401.
19. Krejbjerg A, Bjergved L, Pedersen IB, et al. Iodine fortification may influence the age-related change in thyroid volume: a longitudinal population-based study (DanThyr). *Eur J Endocrinol*. 2014;170:507-517.
20. Gharib H, Papini E, Garber JR, et al. American Association of Clinical Endocrinologists, American College of Endocrinology, and Associazione Medici Endocrinologi Medical Guidelines for Clinical Practice for the Diagnosis and Management of Thyroid Nodules--2016 Update. *Endocr Pract*. 2016;22:622-639.
21. Brix TH, Hansen PS, Kyvik KO, Hegedus L. Cigarette smoking and risk of clinically overt thyroid disease: a population-based twin case-control study. *Archives of Internal Medicine*. 2000;160:661-666.
22. Greig WR, Boyle JA, Duncan A, et al. Genetic and non-genetic factors in simple goitre formation: evidence from a twin study. *Q J Med*. 1967;36:175-188.
23. Zimmermann MB. Iodine deficiency. *Endocr Rev*. 2009;30:376-408.
24. Pearce EN, Braverman LE. Environmental pollutants and the thyroid. *Best Pract Res Clin Endocrinol Metab*. 2009;23:801-813.
25. Boas M, Hegedus L, Feldt-Rasmussen U, Skakkebaek NE, Hilsted L, Main KM. Association of thyroid gland volume, serum insulin-like growth factor-I, and anthropometric variables in euthyroid prepubertal children. *J Clin Endocrinol Metab*. 2009;94:4031-4035.
26. Volzke H, Friedrich N, Schipf S, et al. Association between serum insulin-like growth factor-I levels and thyroid disorders in a population-based study. *J Clin Endocrinol Metab*. 2007;92:4039-4045.
27. von Werder K. Acromegaly and thyroid. *J Endocrinol Invest*. 2002;25:930-931.
28. Clement S, Refetoff S, Robaye B, Dumont JE, Schurmans S. Low TSH requirement and goiter in transgenic mice overexpressing IGF-I and IGF-Ir receptor in the thyroid gland. *Endocrinology*. 2001;142:5131-5139.
29. Bidey SP, Hill DJ, Eggo MC. Growth factors and goitrogenesis. *J Endocrinol*. 1999;160:321-332.
30. Krohn K, Fuhrer D, Bayer Y, et al. Molecular pathogenesis of euthyroid and toxic multinodular goiter. *Endocr Rev*. 2005;26:504-524.
31. Rendina D, De Filippo G, Mossetti G, et al. Relationship between metabolic syndrome and multinodular non-toxic goiter in an inpatient population from a geographic area with moderate iodine deficiency. *J Endocrinol Invest*. 2012;35:407-412.
32. Knudsen N, Brix TH. Genetic and non-iodine-related factors in the aetiology of nodular goitre. *Best Pract Res Clin Endocrinol Metab*. 2014;28:495-506.
33. Al-Bazzaz F, Grillo H, Kazemi H. Response to exercise in upper airway obstruction. *Am Rev Respir Dis*. 1975;111:631-640.
34. Bianchi GP, Zaccheroni V, Solaroli E, et al. Health-related quality of life in patients with thyroid disorders. *Qual Life Res*. 2004;13:45-54.
35. Cramon P, Bonnema SJ, Bjorner JB, et al. Quality of life in patients with benign nontoxic goiter: impact of disease and treatment response, and comparison with the general population. *Thyroid*. 2015;25:284-291.
36. Scerrino G, Inviati A, Di Giovanni S, et al. Esophageal motility changes after thyroidectomy; possible associations with postoperative voice and swallowing disorders: preliminary results. *Otolaryngology--Head and Neck Surgery*. 2013;148:926-932.

37. Lombardi CP, Raffaelli M, De Crea C, et al. Long-term outcome of functional post-thyroidectomy voice and swallowing symptoms. *Surgery*. 2009;146:1174-1181.
38. Lombardi CP, Raffaelli M, D'Alatri L, et al. Video-assisted thyroidectomy significantly reduces the risk of early postthyroidectomy voice and swallowing symptoms. *World Journal of Surgery*. 2008;32:693-700.
39. Shin JJ, Grillo HC, Mathisen D, et al. The surgical management of goiter: Part I. Preoperative evaluation. *Laryngoscope*. 2011;121:60-67.
40. Nettekville JL, Coleman SC, Smith JC, Smith MM, Day TA, Burkey BB. Management of substernal goiter. *Laryngoscope*. 1998;108:1611-1617.
41. Glinoe D, Verelst J, Ham HR. Abnormalities of esophageal transit in patients with sporadic nontoxic goitre. *European Journal of Nuclear Medicine*. 1987;13:239-243.
42. Jorgensen F, Hesse B, Gronbaek P, Fogh J, Hainso S. Abnormal oesophageal function in patients with non-toxic goiter or enlarged left atrium, demonstrated by radionuclide transit measurements. *Scandinavian Journal of Gastroenterology*. 1989;24:1186-1192.
43. Bredenoord AJ, Hebbard GS. Technical aspects of clinical high-resolution manometry studies. *Neurogastroenterol Motil*. 2012;24 Suppl 1:5-10.
44. Bogte A, Bredenoord AJ, Oors J, Siersema PD, Smout AJ. Reproducibility of esophageal high-resolution manometry. *Neurogastroenterol Motil*. 2011;23:e271-276.
45. Bonnema SJ. Upper airway obstruction due to goiter: an overlooked problem! *J Postgrad Med*. 2008;54:82-83.
46. Menon SK, Jagtap VS, Sarathi V, et al. Prevalence of upper airway obstruction in patients with apparently asymptomatic euthyroid multi nodular goitre. *Indian Journal of Endocrinology and Metabolism*. 2011;15:S127-131.
47. Miller MR, Pincock AC, Oates GD, Wilkinson R, Skene-Smith H. Upper airway obstruction due to goitre: detection, prevalence and results of surgical management. *The Quarterly Journal of Medicine*. 1990;74:177-188.
48. Thusoo TK, Gupta U, Kochhar K, Hira HS. Upper airway obstruction in patients with goiter studies by flow volume loops and effect of thyroidectomy. *World Journal of Surgery*. 2000;24:1570-1572.
49. Jauregui R, Lilker ES, Bayley A. Upper airway obstruction in euthyroid goiter. *The Journal of the American Medical Association*. 1977;238:2163-2166.
50. Albareda M, Viguera J, Santiveri C, et al. Upper airway obstruction in patients with endotracheal goiter enlargement: no relationship between flow-volume loops and radiological tests. *European Journal of Endocrinology*. 2010;163:665-669.
51. Komolafe F. Place of tracheal configuration in the radiological assessment of goitres. *African Journal of Medicine and Medical Sciences*. 1984;13:183-187.
52. Abboud B, Sleilaty G, Mallak N, Abou Zeid H, Tabchy B. Morbidity and mortality of thyroidectomy for substernal goiter. *Head & Neck*. 2010;32:744-749.
53. Wax MK, Briant TD. Management of substernal goitre. *The Journal of Otolaryngology*. 1992;21:165-170.
54. Shen WT, Kebebew E, Duh QY, Clark OH. Predictors of airway complications after thyroidectomy for substernal goiter. *Archives of Surgery (Chicago, Ill. : 1960)*. 2004;139:656-659; discussion 659-660.
55. Mackle T, Meaney J, Timon C. Tracheoesophageal compression associated with substernal goitre. Correlation of symptoms with cross-sectional imaging findings. *The Journal of Laryngology and Otology*. 2007;121:358-361.
56. Shaha AR, Alfonso AE, Jaffe BM. Operative treatment of substernal goiters. *Head & Neck*. 1989;11:325-330.
57. Gardiner KR, Russell CF. Thyroidectomy for large multinodular colloid goitre. *Journal of the Royal College of Surgeons of Edinburgh*. 1995;40:367-370.
58. Geraghty JG, Coveney EC, Kiernan M, O'Higgins NJ. Flow volume loops in patients with goiters. *Annals of Surgery*. 1992;215:83-86.
59. Cooper JC, Nakielny R, Talbot CH. The use of computed tomography in the evaluation of large multinodular goitres. *Annals of the Royal College of Surgeons of England*. 1991;73:32-35.
60. Mizaki AP, Araujo-Filho VJ, Brandao LG, de Araujo-Neto VJ, Matos LL, Cernea CR. Vocal fold mobility alteration reversed after thyroidectomy. *Autops Case Rep*. 2016;6:53-57.
61. Nakadate Y, Fukuda T, Hara H, Tanaka M. Tracheomalacia after reoperation for an adenomatous goiter located in a unique position. *Journal of Anesthesia*. 2011;25:745-748.
62. Peric I, Paladin I, Lozo Vukovac E, Vela Ljubic J, Gudelj I, Lozo M. Tracheomalacia, to stent or not to stent. *Respir Med Case Rep*. 2015;16:137-139.
63. Lucchini R, Santoprete S, Triola R, et al. Tracheal varices caused by mediastinal compression of a large intrathoracic goiter: report of a case. *G Chir*. 2015;36:26-28.
64. Evans RW, Timm JS. New Daily Persistent Headache Caused by a Multinodular Goiter and Headaches Associated With Thyroid Disease. *Headache*. 2017;57:285-289.
65. Hunt JP, Wilson M, Buchmann LO. Chylothorax associated with substernal goiter treated with transcervical thyroidectomy. *Thyroid*. 2011;21:551-553.
66. McNeill CJ, Sinnott JD, Howlett D. Bilateral brachiocephalic vein compression: an unusual and rare presentation of multinodular goitre. *BMJ Case Rep*. 2016;2016.
67. Chafik A, Benjelloun A, Bouchentouf R, Mikou I, Ehriouchi A. Left laterotracheal retrosternal goiter with left upper limb ischemia. *Asian Cardiovasc Thorac Ann*. 2010;18:579-580.
68. Lonnebakken MT, Pedersen OM, Andersen KS, Varhaug JE. Incidental detection of internal jugular vein thrombosis secondary to undiagnosed benign substernal goiter. *Case Rep Med*. 2010;2010.
69. Moalem J, Suh I, Duh QY. Treatment and prevention of recurrence of multinodular goiter: an evidence-based review of the literature. *World J Surg*. 2008;32:1301-1312.
70. Thomusch O, Machens A, Sekulla C, et al. Multivariate analysis of risk factors for postoperative complications in benign goiter surgery: prospective multicenter study in Germany. *World J Surg*. 2000;24:1335-1341.
71. Bahn Chair RS, Burch HB, Cooper DS, et al. Hyperthyroidism and other causes of thyrotoxicosis:

- management guidelines of the American Thyroid Association and American Association of Clinical Endocrinologists. *Thyroid*. 2011;21:593-646.
72. Agarwal G, Aggarwal V. Is total thyroidectomy the surgical procedure of choice for benign multinodular goiter? An evidence-based review. *World J Surg*. 2008;32:1313-1324.
 73. Barczynski M, Konturek A, Golkowski F, Hubalewska-Dydejczyk A, Cichon S, Nowak W. Five-year follow-up of a randomized clinical trial of unilateral thyroid lobectomy with or without postoperative levothyroxine treatment. *World J Surg*. 2010;34:1232-1238.
 74. Vaiman M, Nagibin A, Olevson J. Complications in primary and completed thyroidectomy. *Surg Today*. 2010;40:114-118.
 75. Mishra A, Sabaretnam M, Chand G, et al. Quality of life (QoL) in patients with benign thyroid goiters (pre- and post-thyroidectomy): a prospective study. *World J Surg*. 2013;37:2322-2329.
 76. Bukvic BR, Zivaljevic VR, Sipetic SB, Diklic AD, Tausanovic KM, Paunovic IR. Improvement of quality of life in patients with benign goiter after surgical treatment. *Langenbecks Arch Surg*. 2014;399:755-764.
 77. Zivaljevic VR, Bukvic Bacotic BR, Sipetic SB, et al. Quality of life improvement in patients with Hashimoto thyroiditis and other goiters after surgery: A prospective cohort study. *Int J Surg*. 2015;21:150-155.
 78. Watt T, Cramon P, Frendl DM, et al. Assessing health-related quality of life in patients with benign non-toxic goitre. *Best Practice & Research Clinical Endocrinology & Metabolism*. 2014;28(4):559-575.
 79. Greenblatt DY, Sippel R, Levenson G, Frydman J, Schaefer S, Chen H. Thyroid resection improves perception of swallowing function in patients with thyroid disease. *World Journal of Surgery*. 2009;33:255-260.
 80. Maung KH, Hayworth D, Nix PA, Atkin SL, England RJ. Thyroidectomy does not cause globus pattern symptoms. *The Journal of Laryngology and Otology*. 2005;119:973-975.
 81. Tae K, Kim KY, Yun BR, et al. Functional voice and swallowing outcomes after robotic thyroidectomy by a gasless unilateral axillo-breast approach: comparison with open thyroidectomy. *Surgical Endoscopy*. 2012;26:1871-1877.
 82. Pradeep PV, Tiwari P, Mishra A, et al. Pulmonary function profile in patients with benign goiters without symptoms of respiratory compromise and the early effect of thyroidectomy. *Journal of Postgraduate Medicine*. 2008;54:98-101.
 83. Sorensen JR, Hegedus L, Kruse-Andersen S, Godballe C, Bonnema SJ. The impact of goitre and its treatment on the trachea, airflow, oesophagus and swallowing function. A systematic review. *Best Pract Res Clin Endocrinol Metab*. 2014;28:481-494.
 84. Geelhoed GW. Tracheomalacia from compressing goiter: management after thyroidectomy. *Surgery*. 1988;104:1100-1108.
 85. Chandrasekhar SS, Randolph GW, Seidman MD, et al. Clinical practice guideline: improving voice outcomes after thyroid surgery. *Otolaryngol Head Neck Surg*. 2013;148:S1-37.
 86. 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;393:667-673.
 87. Promberger R, Ott J, Kober F, Karik M, Freissmuth M, Herrmann M. Normal parathyroid hormone levels do not exclude permanent hypoparathyroidism after thyroidectomy. *Thyroid*. 2011;21:145-150.
 88. Dralle H, Lorenz K, Machens A. State of the art: surgery for endemic goiter--a plea for individualizing the extent of resection instead of heading for routine total thyroidectomy. *Langenbecks Arch Surg*. 2011;396:1137-1143.
 89. Barczynski M, Konturek A, Cichon S. Randomized clinical trial of visualization versus neuromonitoring of recurrent laryngeal nerves during thyroidectomy. *Br J Surg*. 2009;96:240-246.
 90. Ortega J, Cassinello N, Dorcarratto D, Leopaldi E. Computerized acoustic voice analysis and subjective scaled evaluation of the voice can avoid the need for laryngoscopy after thyroid surgery. *Surgery*. 2009;145:265-271.
 91. Sosa JA, Bowman HM, Tielsch JM, Powe NR, Gordon TA, Udelsman R. The importance of surgeon experience for clinical and economic outcomes from thyroidectomy. *Ann Surg*. 1998;228:320-330.
 92. Stavrakis AI, Ituarte PH, Ko CY, Yeh MW. Surgeon volume as a predictor of outcomes in inpatient and outpatient endocrine surgery. *Surgery*. 2007;142:887-899; discussion 887-899.
 93. Pieracci FM, Fahey TJ, 3rd. Effect of hospital volume of thyroidectomies on outcomes following substernal thyroidectomy. *World J Surg*. 2008;32:740-746.
 94. Biondi B, Bartalena L, Cooper DS, Hegedus L, Laurberg P, Kahaly GJ. The 2015 European Thyroid Association Guidelines on Diagnosis and Treatment of Endogenous Subclinical Hyperthyroidism. *Eur Thyroid J*. 2015;4:149-163.
 95. Bonnema SJ, Hegedus L. Radioiodine therapy in benign thyroid diseases: effects, side effects, and factors affecting therapeutic outcome. *Endocrine reviews*. 2012;33:920-980.
 96. Chapman EM, Evans RD. The treatment of hyperthyroidism with radioactive iodine. *J Am Med Assoc*. 1946;131:86-91.
 97. Hertz S, Roberts A. Radioactive iodine in the study of thyroid physiology; the use of radioactive iodine therapy in hyperthyroidism. *J Am Med Assoc*. 1946;131:81-86.
 98. Verelst J, Bonnyns M, Glinoyer D. Radioiodine therapy in voluminous multinodular non-toxic goitre. *Acta Endocrinol (Copenh)*. 1990;122:417-421.
 99. Nygaard B, Hegedus L, Gervil M, Hjalgrim H, Soe-Jensen P, Hansen JM. Radioiodine treatment of multinodular non-toxic goitre. *BMJ*. 1993;307:828-832.
 100. Huysmans DA, Hermus AR, Corstens FH, Barentsz JO, Kloppenborg PW. Large, compressive goiters treated with radioiodine. *Annals of Internal Medicine*. 1994;121:757-762.
 101. Huysmans D, Hermus A, Edelbroek M, Barentsz J, Corstens F, Kloppenborg P. Radioiodine for nontoxic multinodular goiter. *Thyroid*. 1997;7:235-239.

102. de Klerk JM, van Isselt JW, van Dijk A, et al. Iodine-131 therapy in sporadic nontoxic goiter. *J Nucl Med.* 1997;38:372-376.
103. Le Moli R, Wesche MF, Tiel-Van Buul MM, Wiersinga WM. Determinants of longterm outcome of radioiodine therapy of sporadic non-toxic goitre. *Clin Endocrinol (Oxf).* 1999;50:783-789.
104. Bonnema SJ, Bertelsen H, Mortensen J, et al. The feasibility of high dose iodine 131 treatment as an alternative to surgery in patients with a very large goiter: effect on thyroid function and size and pulmonary function. *The Journal of Clinical Endocrinology and Metabolism.* 1999;84:3636-3641.
105. Wesche MF, Tiel VBMM, Lips P, Smits NJ, Wiersinga WM. A randomized trial comparing levothyroxine with radioactive iodine in the treatment of sporadic nontoxic goiter. *J Clin Endocrinol Metab.* 2001;86:998-1005.
106. Bachmann J, Kobe C, Bor S, et al. Radioiodine therapy for thyroid volume reduction of large goitres. *Nucl Med Commun.* 2009;30:466-471.
107. Graf H, Fast S, Pacini F, et al. Modified-release recombinant human TSH (MRrhTSH) augments the effect of (131)I therapy in benign multinodular goiter: results from a multicenter international, randomized, placebo-controlled study. *The Journal of Clinical Endocrinology and Metabolism.* 2011;96:1368-1376.
108. Kang AS, Grant CS, Thompson GB, van Heerden JA. Current treatment of nodular goiter with hyperthyroidism (Plummer's disease): surgery versus radioiodine. *Surgery.* 2002;132:916-923; discussion 923.
109. Fast S, Nielsen VE, Bonnema SJ, Hegedus L. Time to reconsider nonsurgical therapy of benign non-toxic multinodular goiter: focus on recombinant human TSH augmented radioiodine therapy. *Eur J Endocrinol.* 2009;160:517-528.
110. Huysmans DA, Nieuwlaat WA, Erdtsieck RJ, et al. Administration of a single low dose of recombinant human thyrotropin significantly enhances thyroid radioiodide uptake in nontoxic nodular goiter. *J Clin Endocrinol Metab.* 2000;85:3592-3596.
111. Schmidt M, Gorbauch E, Dietlein M, et al. Incidence of postradioiodine immunogenic hyperthyroidism/Graves' disease in relation to a temporary increase in thyrotropin receptor antibodies after radioiodine therapy for autonomous thyroid disease. *Thyroid.* 2006;16:281-288.
112. Brown AP, Chen J, Hitchcock YJ, Szabo A, Shrieve DC, Tward JD. The risk of second primary malignancies up to three decades after the treatment of differentiated thyroid cancer. *J Clin Endocrinol Metab.* 2008;93:504-515.
113. Ron E, Doody MM, Becker DV, et al. Cancer mortality following treatment for adult hyperthyroidism. Cooperative Thyrotoxicosis Therapy Follow-up Study Group. *JAMA.* 1998;280:347-355.
114. Angusti T, Codegone A, Pellerito R, Favero A. Thyroid cancer prevalence after radioiodine treatment of hyperthyroidism. *J Nucl Med.* 2000;41:1006-1009.
115. Pacella CM, Mauri G, Achille G, et al. Outcomes and Risk Factors for Complications of Laser Ablation for Thyroid Nodules: A Multicenter Study on 1531 Patients. *J Clin Endocrinol Metab.* 2015;100:3903-3910.
116. Chung SR, Suh CH, Baek JH, Park HS, Choi YJ, Lee JH. Safety of Radiofrequency Ablation of Benign Thyroid Nodules and Recurrent Thyroid Cancers: A Systematic Review and Meta-Analysis. *Int J Hyperthermia.* 2017:1-35.
117. Mainini AP, Monaco C, Pescatori LC, et al. Image-guided thermal ablation of benign thyroid nodules. *J Ultrasound.* 2017;20:11-22.
118. Lim HK, Lee JH, Ha EJ, Sung JY, Kim JK, Baek JH. Radiofrequency ablation of benign non-functioning thyroid nodules: 4-year follow-up results for 111 patients. *Eur Radiol.* 2013;23:1044-1049.
119. Sung JY, Baek JH, Kim KS, et al. Single-session treatment of benign cystic thyroid nodules with ethanol versus radiofrequency ablation: a prospective randomized study. *Radiology.* 2013;269:293-300.
120. Bernoulli D. Exposition of a New Theory on the Measurement of Risk. *Econometrica.* 1954;22:23-36.
121. Kahneman D. *Thinking Fast and Slow.* 1st ed 2011.
122. Lu CY. Observational studies: a review of study designs, challenges and strategies to reduce confounding. *Int J Clin Pract.* 2009;63:691-697.
123. Sibai T, Carlisle H, Tornetta P, 3rd. The darker side of randomized trials: recruitment challenges. *J Bone Joint Surg Am.* 2012;94 Suppl 1:49-55.
124. Huynh L, Johns B, Liu SH, Vedula SS, Li T, Puhan MA. Cost-effectiveness of health research study participant recruitment strategies: a systematic review. *Clin Trials.* 2014;11:576-583.
125. Fletcher B, Gheorghe A, Moore D, Wilson S, Damery S. Improving the recruitment activity of clinicians in randomised controlled trials: a systematic review. *BMJ Open.* 2012;2:e000496.
126. Newington L, Metcalfe A. Researchers' and clinicians' perceptions of recruiting participants to clinical research: a thematic meta-synthesis. *J Clin Med Res.* 2014;6:162-172.
127. Caldwell PH, Hamilton S, Tan A, Craig JC. Strategies for increasing recruitment to randomised controlled trials: systematic review. *PLoS Med.* 2010;7:e1000368.
128. Treweek S, Pitkethly M, Cook J, et al. Strategies to improve recruitment to randomised controlled trials. *Cochrane Database Syst Rev.* 2010:MR000013.
129. Edwards PJ, Roberts I, Clarke MJ, et al. Methods to increase response to postal and electronic questionnaires. *Cochrane Database Syst Rev.* 2009:MR000008.
130. Reinke JM, Sorg H. Wound repair and regeneration. *Eur Surg Res.* 2012;49:35-43.
131. Promberger R, Hermann M, Pallikunnel SJ, Seemann R, Meusel M, Ott J. Quality of life after thyroid surgery in women with benign euthyroid goiter: influencing factors including Hashimoto's thyroiditis. *Am J Surg.* 2014;207:974-979.
132. Schmitz-Winnenthal FH, Schimmack S, Lawrence B, et al. Quality of life is not influenced by the extent of surgery in patients with benign goiter. *Langenbecks Arch Surg.* 2011;396:1157-1163.
133. 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;162:161-167.

134. Watt T, Bjorner JB, Groenvold M, et al. Establishing construct validity for the thyroid-specific patient reported outcome measure (ThyPRO): an initial examination. *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation*. 2009;18:483-496.
135. Watt T, Barbesino G, Bjorner JB, et al. Cross-cultural validity of the thyroid-specific quality-of-life patient-reported outcome measure, ThyPRO. *Qual Life Res*. 2014.
136. Watt T, Groenvold M, Hegedus L, et al. Few items in the thyroid-related quality of life instrument ThyPRO exhibited differential item functioning. *Qual Life Res*. 2014;23:327-338.
137. Watt T, Groenvold M, Deng N, et al. Confirmatory factor analysis of the thyroid-related quality of life questionnaire ThyPRO. *Health Qual Life Outcomes*. 2014;12:126.
138. Wong CK, Lang BH, Lam CL. A systematic review of quality of thyroid-specific health-related quality-of-life instruments recommends ThyPRO for patients with benign thyroid diseases. *J Clin Epidemiol*. 2016;78:63-72.
139. Watt T, Cramon P, Hegedus 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;99:3708-3717.
140. Wong CK, Lang BH, Lam CL. A systematic review of quality of thyroid-specific health-related quality-of-life instruments recommends ThyPRO for patients with benign thyroid diseases. *J Clin Epidemiol*. 2016.
141. Watt T, Barbesino G, Bjorner JB, et al. Cross-cultural validity of the thyroid-specific quality-of-life patient-reported outcome measure, ThyPRO. *Qual Life Res*. 2015;24:769-780.
142. Watt T, Groenvold M, Deng N, et al. Confirmatory factor analysis of the thyroid-related quality of life questionnaire ThyPRO. *Health Qual Life Outcomes*. 2014;12:126.
143. Rasmussen SL, Rejnmark L, Ebbelohj E, et al. High Level of Agreement between Electronic and Paper Mode of Administration of a Thyroid-Specific Patient-Reported Outcome, ThyPRO. *Eur Thyroid J*. 2016;5:65-72.
144. Fayers PM, Machin D. *Quality of Life: The assessment, analyses, and interpretation of patient-reported outcomes*. 2nd ed: John Wiley & Sons Ltd; 2007.
145. Watt T, Bjorner JB, Groenvold M, et al. Development of a Short Version of the Thyroid-Related Patient-Reported Outcome ThyPRO. *Thyroid*. 2015;25:1069-1079.
146. Timmerman AA, Speyer R, Heijnen BJ, Klijn-Zwijnenberg IR. Psychometric characteristics of health-related quality-of-life questionnaires in oropharyngeal dysphagia. *Dysphagia*. 2014;29:183-198.
147. Kuo P, Holloway RH, Nguyen NQ. Current and future techniques in the evaluation of dysphagia. *Journal of Gastroenterology and Hepatology*. 2012;27:873-881.
148. Ayazi S, Crookes PF. High-resolution esophageal manometry: using technical advances for clinical advantages. *J Gastrointest Surg*. 2010;14 Suppl 1:S24-32.
149. Conklin JL. Evaluation of Esophageal Motor Function With High-resolution Manometry. *J Neurogastroenterol Motil*. 2013;19:281-294.
150. Kahrilas PJ, Bredenoord AJ, Fox M, et al. The Chicago Classification of esophageal motility disorders, v3.0. *Neurogastroenterol Motil*. 2015;27:160-174.
151. Rohof WOA, Bredenoord AJ. Chicago Classification of Esophageal Motility Disorders: Lessons Learned. *Curr Gastroenterol Rep*. 2017;19:37.
152. Babaei A, Lin EC, Szabo A, Massey BT. Determinants of pressure drift in Manoscan() esophageal high-resolution manometry system. *Neurogastroenterol Motil*. 2015;27:277-284.
153. Miller MR, Hankinson J, Brusasco V, et al. Standardisation of spirometry. *Eur Respir J*. 2005;26:319-338.
154. Melissant CF, Smith SJ, Perlberger R, Verschakelen J, Lammers JW, Demedts M. Lung function, CT-scan and X-ray in upper airway obstruction due to thyroid goitre. *European Respiratory Journal*. 1994;7:1782-1787.
155. Bonnema SJ, Andersen PB, Knudsen DU, Hegedus L. MR imaging of large multinodular goiters: observer agreement on volume versus observer disagreement on dimensions of the involved trachea. *American Journal of Roentgenology*. 2002;179:259-266.
156. Brauer VF, Eder P, Miehle K, Wiesner TD, Hasenclever H, Paschke R. Interobserver variation for ultrasound determination of thyroid nodule volumes. *Thyroid*. 2005;15:1169-1175.
157. Choi SH, Kim EK, Kwak JY, Kim MJ, Son EJ. Interobserver and intraobserver variations in ultrasound assessment of thyroid nodules. *Thyroid*. 2010;20:167-172.
158. Sorensen JR, Koroma KE, Ding M, et al. Effects of a perfusion bioreactor activated novel bone substitute in spine fusion in sheep. *Eur Spine J*. 2012;21:1740-1747.
159. Cohen J. *Statistical power analyses for behavioral sciences. Chapter 2*. Second edition ed: Lawrence Erlbaum, Hillsdale, NJ; 1988.
160. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*. 2009;42:377-381.
161. Rieder H.L. LJM. Quality assurance of data: ensuring that numbers reflect operational definitions and contain real measurements *The International Journal of Tuberculosis and Lung Disease*. 2011;15:296-304.
162. Oxford Centre for Evidence-Based Medicine - Levels of Evidence, <http://www.cebm.net/oxford-centre-evidence-based-medicine-levels-evidence-march-2009/>. Vol 20172009.
163. Bonnema SJ, Nielsen VE, Boel-Jorgensen H, et al. Recombinant human thyrotropin-stimulated radioiodine therapy of large nodular goiters facilitates tracheal decompression and improves inspiration. *The Journal of Clinical Endocrinology and Metabolism*. 2008;93:3981-3984.
164. Sorensen JRW, T.; Cramon, P.; Døssing, H.; Hegedüs, L.; Bonnema, S.J.; Godballe, C. The quality of life after thyroidectomy in patients with nontoxic nodular goiter – a prospective cohort study. *Head & Neck*. 2017.
165. Sorensen JR, Markoew S, Dossing H, Hegedus L, Bonnema SJ, Godballe C. Changes in Swallowing Symptoms and Esophageal Motility After Thyroid Surgery: A Prospective Cohort Study. *World J Surg*. 2017.

166. Herregods TV, Roman S, Kahrilas PJ, Smout AJ, Bredenoord AJ. Normative values in esophageal high-resolution manometry. *Neurogastroenterol Motil.* 2015;27:175-187.
167. Sorensen JR, Lauridsen JK, Døssing H, et al. Thyroidectomy Improves Tracheal Anatomy and Airflow in Patients with Nodular Goiter: A Prospective Cohort Study. *Eur Thyroid J.* 2017.
168. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA.* 1995;273:408-412.
169. Balk EM, Bonis PA, Moskowitz H, et al. Correlation of quality measures with estimates of treatment effect in meta-analyses of randomized controlled trials. *JAMA.* 2002;287:2973-2982.
170. Juni P, Altman DG, Egger M. Systematic reviews in health care: Assessing the quality of controlled clinical trials. *BMJ.* 2001;323:42-46.
171. Karanicolas PJ, Farrokhyar F, Bhandari M. Practical tips for surgical research: blinding: who, what, when, why, how? *Can J Surg.* 2010;53:345-348.
172. Sorensen JR, Winther KH, Bonnema SJ, Godballe C, Hegedus L. Respiratory Manifestations of Hypothyroidism: A Systematic Review. *Thyroid.* 2016;26:1519-1527.
173. Glickman SW, Ndubuizu A, Weinfurt KP, et al. Perspective: The case for research justice: inclusion of patients with limited English proficiency in clinical research. *Acad Med.* 2011;86:389-393.
174. 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;26:1010-1018.
175. Winther KH, Cramon P, Watt T, et al. Disease-Specific as Well as Generic Quality of Life Is Widely Impacted in Autoimmune Hypothyroidism and Improves during the First Six Months of Levothyroxine Therapy. *PLoS One.* 2016;11:e0156925.
176. George NS, Rangan V, Geng Z, et al. Distribution of Esophageal Motor Disorders in Diabetic Patients With Dysphagia. *J Clin Gastroenterol.* 2017.
177. Empey D. Patterns of airflow in upper airways obstruction. *Thorax.* 1972;27:262.