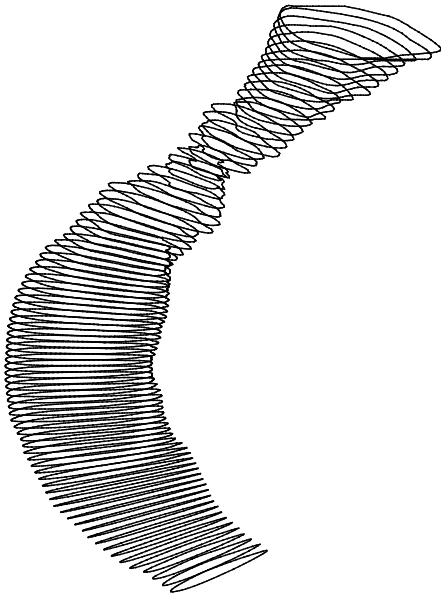
GOITRE-INDUCED VENTILATION IMPAIRMENT

BART HOEBEN



Exploring opportunities to improve management of goitre-induced trachea compression.

Department of Technical Medicine
Faculty of Science and Technology
University of Twente

November 6, 2015 – Final version

Bart Hoeben: *Goitre-induced Ventilation Impairment,* Exploring opportunities to improve management of goitre-induced trachea compression.
, Master's Thesis© November 6, 2015

SUPERVISORS:
Dr. H. de Boer
Prof. Dr. Ir. B.J. Geurts
Drs. N.S. Cramer Bornemann

LOCATION:
Arnhem

TIME FRAME:
November 2014 - November 2015

Whatever you do in this life,
it is not legendary unless your friends are there to see it.

— Barney Stinson

Dedicated to the loving memory of Jan Wagemans.

1932 – 2013

ABSTRACT

Goitre-induced trachea compression (GITC) may increase airway resistance to such an extent that ventilation capacity is impaired and physical performance is reduced. Objective assessment of these ventilation impairments may help to strengthen the indications for goire reduction.

The aim of this thesis was to devise a model for objective assessment of GITC and define how these results should be interpreted.

First an automated segmentation tool was devised for objective assessment of trachea dimension in CT-data. A novel model for the interpretation of these dimensions was devised using simplified fluid dynamics. Obtained trachea dimensions were used to calculate airway resistance and its effects on maximal achievable airflow. Ventilation capacity of the trachea was expressed as maximal ventilation capacity (MVC), defined as the calculated peak flow in the diseased state divided by the calculated peak flow in the healthy state. The MVC reflects the patient's potential ventilation capacity in the absence of cardiac or pulmonary disease, and this is the key to assess the limitations in daily life activities exclusively causes by a compressive goitre.

First this model was applied to pre- and post-operative CT-scan data from 10 patients with varying degrees of trachea narrowing to test its applicability. Preliminary results of this pilot study suggested the MVC might be a valuable tool for the assessment of clinical impact of GITC.

Thereafter a retrospective study was initiated, this study identified potential other useful Radiological Parameters such as the minimal cross-sectional area (CSA_m) and percentage trachea constriction (%-TC). Reference values for these additional RP and MVC were established using CT-data of 55 control patients. Thereafter the correlation between these RP and reported sensations dyspnea was tested in 72 patients with varying degrees of GITC. It was recognised that quality of these reported symptoms were suboptimal, therefore a exploratory prospective study was initiated.

This exploratory study identified other potential measures which could be used to validate the RP proposed during the retrospective study and how they should be applied.

Based on the results of this exploratory study a follow-up prospective study was initiated aimed to identify the most relevant measures for assessment of GITC. Due to the low sample size this study has not been able to generate conclusive statements regarding the assessment

of GITC. However, preliminary results suggest that will change with the conclusion of additional patients.

D'n iene di rent veur zien leave,
d'n andere wandelt hiel rustig vurbeej.
Heej zuj d'r alles vur geave,
en heej zet: ze mooge 't hebben van meej.
Woar ge ok loept en wat ge ok bint,
niemand die zet ow wat good is of slecht,
niemand die wet wie verluust of wie wint,
ge komt op 't end beej ow zelf terecht.

— Jack Poels

ACKNOWLEDGEMENTS

My journey in the wonderful world of goitres started in November of 2013 at the Rijnstate Hospital Group in Arnhem. At this time I started a 3-month clinical internship to research the effects of Goitre Induced Trachea Compression. My biggest thanks I owe to dr. de Boer who devoted a great amount of attention and time this project. E. Groot Jebbink MSc. has also helped me by providing essential assistance in the execution of our project. Dr. F. Joosten has helped me a lot with collecting my data, which was essential for my research. Dr. C. Blanken who allowed to study thyroid surgery in practice.
After this initial period of three months I continued the project for another year. Hereby we were able include other people to help realise this project. These people include but are not limited to prof. dr. ir. B.J. Geurts who helped me with his technical insights and provided me with a critical point of view, F. Borm which helped to conduct anamneses on various patients and put a lot of effort in the execution of the research, dr. B. Tonino and dr. M. van Gorp who helped to evaluate and execute the research using their clinical expertise, the staff of department of Internal Medicine which helped to include patients for my research as well as moral support, the staff of the department of Radiology which helped me with obtaining my CT-data, my dialogue group consisting of M. Kamphuis MSc., P. Bos BSc., B. Boti BSc., R. Wesselink BSc. and drs. N.S. Cramer Bornemann which supported me on my personal goals, my friends and family with a special mention of F.V. de Ruiter BSc. which has been my tower of strength for years.

Arnhem, November 6, 2015

Bart Hoeben

Bart Wilhelmus Hoeben

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

CC-point Closest to Centre-point

CFD Computational Fluid Dynamics

CHAD Consolidated Human Activity Database

CI Confidence Intervals

CLL Centre Lumen Line

CT Computed Tomography

DICOM Digital Imaging and Communications in Medicine

FVL Flow-Volume Loops

HAV Healthy Anatomical Variation

I^{131} Iodine

IOS Impulse Oscillometry Spirometry

GITC Goitre Induced Trachea Compression

GUI Graphical User Interface

LLN Lower Limit of Normal

MVC Maximal Ventilation Capacity

PACS Picture Archiving and Communications System

PD Positional Dyspnea

PFV Pulmonary Function Values

PFT Pulmonary Function Tests

RFA Radio Frequency Ablation

RA Radiological Assessment

RP Radiological Parameters

RR Respiratory Rate

SGRQ St. George's Respiratory Questionnaire

SRS Self Reported Symptoms

TC Trachea Constriction

TD Trachea Deviation

UAO Upper Airway Obstruction

QUANTITIES

Notation	Description	Unit
A	Area	m^2
C_f	Darcy friction-factor	-
A	Area	m^2
CSA	Cross-Sectional Area	mm^2
CSA_m	Minimal Cross-Sectional Area	mm^2
d	diameter	m
$d_{disease}$	Pre-surgery diameter	mm
d_{min}	Minimal diameter	mm
d_{max}	Maximal diameter	mm
d_{ref}	Reference diameter	mm
f_0	Fundamental frequency	Hz
FEV_1	Forced Expiratory Volume	L
FIF	Forced Inspiratory Flow	$L\ s^{-1}$
k	Roughness	-
l	Length	m
MMV	Maximal Minute Ventilation	$L\ min^{-1}$
μ	Dynamic viscosity	$N\ s\ m^{-2}$
P	Perimeter	m
Δp	Pressure drop	Pa
ϕ	Flow	$m^3\ s^{-1}$
ρ	density	$kg\ m^{-3}$
R	Resistance	$kPa\ L^{-1}\ s$
$\frac{\Delta R}{\Delta \dot{V}}$	Flow dependence of resistance	$kPa\ L^{-2}\ s^2$
Re	Reynolds number	-
t	Time	s
t_0	Start of inhalation	s
t_1	Start of exhalation	s
t_2	End of exhalation	s
V	Volume	m^3
$\dot{V}$	Velocity	$m\ s^{-1}$
$\dot{V}_e$	Ventilation	$L\ min^{-1}$
x_{min}	Smallest x-coordinate of the trachea wall	-
x_{max}	Biggest x-coordinate of the trachea wall	-
y_{min}	Smallest y-coordinate of the trachea wall	-
y_{max}	Biggest y-coordinate of the trachea wall	-

1 BACKGROUND

In iodine-sufficient parts of the world the prevalence of goitre, defined as a palpable enlargement of the thyroid, ranges from 1% in men to 5% in women [1]. Benign, non-toxic goitres usually grow slowly and many years will elapse before symptoms of trachea compression arise [2]. Eventually, goitre-induced trachea compression (GITC) may develop in up to 10% of patients diagnosed with euthyroid goitre [3]. Some of these patients may develop ventilation impairments that reduce physical performance and quality of life, also known as Upper Airway Obstruction (UAO).

1.1 TREATMENT

The two most common therapies for goitre volume reduction include surgery and radioactive I^{131}. Recently a new technique called Radio Frequency Ablation (RFA) has been introduced, which has proven to be a valid and safe tool for goitre volume reduction.
For I^{131} therapy, thyroid volume reductions usually range from 15 - 60%, whereas for RFA volume reductions range from 51-85% and surgery can achieve volume reductions up to 90 - 95% [4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14]. Adverse effects of I^{131} such as radiation induced thyroditis, transient hyperthyroidism, Graves-like hyperthyroidism are reported in 3%, 5%, and 22 - 58% of cases respectively [7]. An increase in thyroid size during treatment, which increases discomfort, also occurs occasionally [7]. Failure of treatment with I^{131} is observed in 20% of patients [15]. Complications of RFA such as pain, voice change, hematoma, skin burn at the puncture site, thyrotoxicosis, hypothyroidism, edema and fever are reported in 6%, 0.2%, 1.6%, 0.2%, 1.4%, 0.6% and 1.1% of cases respectively [16]. Goitre surgery has a very low mortality rate, but can be associated with recurrence nerve palsy, tracheomalacia, transient hypocalcemia, permanent hypocalcemia and hypoparathyroidism in 0.8-2.3%, 0 to 1.5%, 0.3 to 49%, 0 to 13% and 1.5-12.5% of cases respectively [17, 18, 19]. In view of these varying efficacies and potential adverse effects it will be crucial to establish the degree of discomfort caused by GITC as accurately as possible to make a balanced decision for the most optimal treatment strategy in individual cases.

1.2 ASSESSMENT OF GITC

Although symptom reduction after goitre treatment has been well documented, it remains a major challenge to predict the benefits of surgery based on pre-operative symptoms alone [20, 21, 22, 23, 24, 25]. A recent retrospective study by Stang *et al.* including 1081 patients suggests that trachea compression with a reduction in transverse diameter of 35% or more might be a good predictor of symptomatic relieve after volume reduction surgery [24]. However, an objective and quantitative post-operative validation of this proposed predictor of success of surgery is currently not available.
It is commonly recognized that the criteria to perform goitre-volume reduction are not very precisely defined [26, 27]. Dyspnea and dysphagia are often used as indicators for thyroid surgery, however, these symptoms may have many other causes and thus are not very specific. Ideally, the decision to perform volume reduction should be based on a quantitative assessment of the limitations caused by a compressive goitre and by the degree of improvement that can be expected after an intervention.

1.2.1 *Pulmonary Function Tests*

According to literature, the current 'golden standard' for UAO due to goitre is whole-body-plethysmography combined with spirometry [28, 29].

1.2.1.1 *Spirometry*

Spirometry is the simplest method for diagnosis of UAO [30, 31, 32]. Hereby, the evaluation of Flow-Volume Loops (FVL) by an expert has been shown the most sensitive and specific test in Pulmonary Function tests (PFT) determining UAO [31]. However, the quality of such loops is dependent on the patient's effort and cooperation and the investigators familiarity with the device, thereby making its interpretation rather subjective [33, 34, 35, 36].
Additionally, spirometry can quantify discrete Pulmonary Function Values (PFV). Usually upper airway obstructions such as GITC will hinder inspiratory flow prior to significant changes in expiratory flow, because inspiration is completely effort-dependent whereas expiration is only partially effort-dependent [37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47]. Additionally the negative pressure gradient generated during inspiration can cause the trachea to compress even further, whereas the positive pressure gradient generated by expiration prevents this collapse. As a result of this, inspiratory PFV, will detect UAO prior to expiratory PFV. Ideally, these inspiratory PFV are compared to their reference values. Unfortunately, no universal applicable model for the prediction of these reference values is available, apart from some in-

cidental cases on small populations [45, 48, 49]. As a result of this, sensitivity of these PFV for the detection of UAO is low. Furthermore, optimal spirometry is often not feasible, because many patients have difficulty in executing maximal effort respiratory tests, thus causing effort-dependency of the test itself [35, 34, 36]. Moreover, concomitant pathologies, such as cardiac or pulmonary diseases, might confound the results of spirometry [34].

1.2.1.2 *Whole-body Plethysmography*

Plethysmography can be used to assess airway resistance. However, this concerns total airway resistance and is not specific for UAO. Additionally, this method has not been sufficiently standardised and therefore results between institutions are poorly comparable and reproducibility is limited [28].

1.2.1.3 *Impulse Oscillometry*

Impulse Oscillometry (IOS) is an alternative to traditional spirometry. Its advantage over traditional spirometry is the absence of effort dependency contrary to traditional spirometry [50] Additionally, an enhanced sensitivity over spirometry for detecting airflow obstruction has been observed [51]. Although IOS does not require patient cooperation, its results still depends greatly on the investigator's level of familiarity with the device [52, 53, 54, 55, 56, 57, 58, 59, 60]. As a result of this, its correlation with spirometry varies greatly [51]. Nevertheless, assessment of airway resistance using IOS shows similar repeatability compared to traditional assessment of airway resistance using plethysmography [51].
IOS uses pressure waves generated at different frequencies to measure respiratory impedance from which resistance (R) and reactance can be derived [51]. Several studies showed that R for oscillations at 4 or 5 Hz increases with decreasing tracheal luminal area, regardless of shape and length of the stenotic area[32]. Nevertheless, there still remains a lack of specificity for the cause of trachea stenosis. It has been suggested resistance corrected flow ($\frac{\Delta R}{\Delta \dot{V}}$) may be a more promising parameter [32]. In a study by Verbanck *et al*, comparing 10 healthy individuals and 10 patients with tracheal stenosis, R did not differ for controls and patients, whereas $\frac{\Delta R}{\Delta \dot{V}}$ was significantly higher in patients compared to controls. Moreover, after treatment of the stenotic areas $\frac{\Delta R}{\Delta \dot{V}}$ values returned to normal.
$\frac{\Delta R}{\Delta \dot{V}}$ measured by IOS seems like a more sensitive measure for UAO compared to traditional measures in PFT. Moreover, $\frac{\Delta R}{\Delta \dot{V}}$ is not affected by the limitations of traditional PFT such as confounding concomitant diseases, effort dependence and subjective interpretation.
Despite these clear advantages, $\frac{\Delta R}{\Delta \dot{V}}$ has not yet been used to assess

UAO due to GITC. Exploring its applicability and significance in these cases is therefore highly recommended.

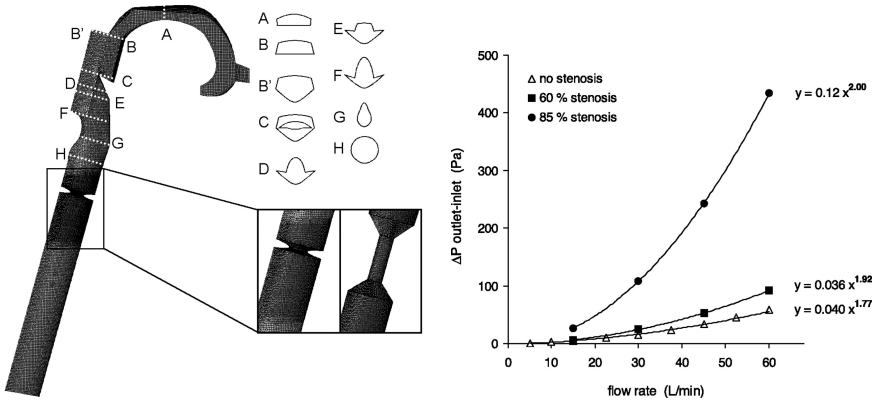
1.2.2 *Self Reported Symptoms*

Assessment of Self-Reported Symptoms (SRS), such as dyspnea, is common in clinical practice. It is a simple low-effort method to evaluate the consequences of UAO. Nevertheless, its reliability is suboptimal. Patients tend to adapt their physical behaviour to the goitre induced limitations and are often unaware to what extent their goitre has gradually reduced their physical capabilities [61]. Furthermore, it is well-known that dyspnea can have a wide variety of causes, and is therefore not specifically linked to UAO. Thyroid function disorders such as hyperthyroidism, are linked to goitre and can induce cardiovascular and respiratory diseases such as heart failure, pulmonary hypertension and respiratory muscle weakness, thereby causing dyspnea regardless of trachea compression [62, 63, 64, 65].

1.2.3 *Radiological Assessment*

Radiological Assessment (RA) can be done using traditional x-ray or Computed Tomography (CT). Hereby radiologist assess Radiological Parameters (RP) such as the magnitude of Trachea Constriction (TC) and Trachea Deviation (TD). However, several studies pointed out that TD is not relevant in terms of ventilation impairment [21, 30].
CT has demonstrated a superior sensitivity for assessment of GITC compared to other measures [66, 67]. Nevertheless these anatomical data alone are insufficient to assess the functional limitations of a compressive goitre. Another limitation is that the severity of trachea malformation is commonly described in rather subjective terms that lack a firm, objective basis [66, 67].
In 2007 Brouns *et al.* researched the relationship between airway resistance and TC by means of a Computational Fluid Dynamics (CFD) study on simulated human trachea's, see figure 1a [68]. They calculated pressure drops caused by airway resistance for different levels of TC. In their model they found that TC of 50% cross sectional area reduction has a similar effect as the glottic narrowing due to the vocal chords and is therefore insignificant. TC levels above 50% showed a strong increase in airway resistance. TC of 75% approximately doubled airway resistance. Further increases in TC will increase airway resistance dramatically due to the exponential relationship between TC and airway resistance. They computed that airway resistance was seen to dramatically increase only when TC was well over 70%, see figure 1b. Thus TC can initially increase without apparent ventilation impairment, however if TC reaches a certain critical value, airway

resistance will increase dramatically and become symptomatic. This is an apparent contradiction to the findings of the earlier mentioned study of Stang *et al* [24]. They found that TC levels over 35% are a reliable predictor for Positional Dyspnea (PD), whereby PD was defined as trouble breathing at rest that was improved or ameliorated by position change. This might be because TC can worsen when head position is altered. Thus ostensibly insignificant levels of TC, can impose significant constrictions by changes in posture.
It should be noted that Brouns *et al.* only researched flow levels up to 60 L min^{-1} [1]. However, several studies show that with increasing level of flow, pressure drop also increases (figure 1b) [68, 71, 72]. Since airflow with a magnitude 1 L s^{-1} is well below flow levels required for high levels of physical exertion, it is possible that constrictions below 70% will not hinder patients at rest, but will induce a limitation of their maximal physical capabilities. Furthermore, an individual with increased airway resistance might be able to reproduce normal flow levels during isolated manoeuvres (such as during spirometry), but will not be able to maintain prolonged bouts of increased ventilation due to fatigue of respiratory muscles. This might explain why patients may experience symptoms such as decreased exercise tolerance, with normal spirometry results.



(a) Simulated trachea used

(b) Pressure drop as a function of flow for various degrees of stenosis

Figure 1: Figures from Brouns *et al.* [68]

1.3 FUTURE PERSPECTIVE

Because physicians strive to intervene prior to severe escalation of UAO in order to keep patient discomfort to a minimum, GITC-management should be based on a reliable, sensitive and objective tool, therefore PFT falls short for use in clinical practice. Nevertheless, as mentioned earlier, SRS and RP are also suboptimal for assessment of UAO. SRS

1 Normal flow for adults at rest is 12 L min^{-1}, but can increase to 97±25 L min^{-1} in men and 69±22 L min^{-1} in women during maximal physical exertion [69, 70].

are subjective and might by confounded by concomitant diseases. Whereas the interpretation of RP has not been clearly defined.
CFD-studies such as the one performed by Brouns *et al.* hold a promise. These studies use radiological data to predict functional consequences of the goitre induced altered flow dynamics, and are thus a means for the interpretation of radiological data. However, their high computational demands makes the assessment of individual cases costly. A solution would be to devise an algorithm which uses highly simplified CFD, and is therefore less costly and more suitable for analysis of individual patients. It is evident such simplifications will induce some margin of error. The magnitude of this error will determine whether such an algorithm is useful in clinical practice. Quantification of this error can be problematic due to the lack of an efficient golden standard. Nevertheless, as explained in section 1.2, with currently available tools, assessment of very severe cases of UAO is possible. Thus analysis of these extreme cases can yield an preliminary conclusion regarding the potential of this simplified algorithm.
Additionally, exploring techniques such as IOS and improving on currently used methods, such as SRS and RP, may further assist to define the optimal strategy for GITC management.

1.4 AIM AND OUTLINE OF THIS THESIS

Patients are often unaware about the limitations imposed by GITC because they unconsciously adapt to the gradually increasing limitations. This phenomenon not only causes unnecessary delay of intervention, it also reduces the diagnostic reliability of reported symptoms, and raises difficulties in convincing patients about the potential benefits of an intervention. A tool that would be able to predict the expected functional improvement might be of great help to improve decision making. The aim of the present study is to devise and test a novel and simple algorithm for the assessment of the functional implications of GITC and the potential benefits of intervention, to generate objective data for decision-making in patients with GITC.
Chapter 2 describes the establishment of such a novel algorithm whilst demonstrating it in a small sample of patients and controls. In chapter 3 this algorithm is tested further and demonstrated on a larger population by means of a retrospective study. Thereafter, chapter 4 shows a prospective study on a single patient aimed to explore and identify potentially useful methods for the validation of this novel algorithm and assessment of GITC. Next, chapter 5 provides a more in-depth analysis of procedure introduced in chapter 2 for the analysis of trachea dimensions. Subsequently, chapter 6 analyses the reliability of assessment of RP by radiologists compared to automated segmentations methods. Thereafter, chapter 7 includes a prospective study

which aims to verify some of the hypothesis devised in preceding chapters. After that chapter 8 provides insights in how the algorithm devised in chapter 2 may be used in clinical practice. Finally chapter 9 wraps up this thesis by summarizing the most important conclusions of the preceding chapters and generating recommendations for future research based on these findings.

2

PILOT-STUDY

Goitre is a enlargement of the thyroid. This goitre may cause trachea compression, which reduced trachea dimensions and raises airway resistance, thereby causing ventilatory impairment. This phenomenon is also known as Goitre Induced Trachea Compression GITC [3].
A commonly used method for the assessment if the assessment of radiological parameters. However, these parameters are often describes in rather subjective terms, making them difficult to interpret [33, 34, 35, 36].
Chapter 1 motivates that simplified computational fluid dynamics might circumvent some of these issues by objectively describing RP and how they should be interpreted.
The aim of this study is to test the practicability and applicability of a simplified CFD-algorithm for the assessment of GITC using CT-data.

2.1 METHODS AND SUBJECTS

A method was developed to translate CT-derived trachea dimensions into limitations of maximal achievable airflow. This method used measurements of the areas unaffected by goitre were used to assess the individual's normal airflow characteristics, thus providing individualized normal ranges in size and function. Model performance was tested in a selection of 10 patients with varying degrees of trachea stenosis: five males and five females, ranging in age from 45 to 78 years. Non-contrast CT-data of these patients had been obtained three months before and three months after surgery to assess the efficacy of surgery. The indications for surgery were dyspnea, dysphagia and globus pharyngis. Two patients had a total thyroidectomy and eight subtotal. CT-images of ten age- and sex-matched subjects without goitre served as control data: 5 males and 5 females, ranging in age from 55 to 76 years with an indication for CT-scanning of the neck and mediastinum as part of the work-up for lung cancer staging, analysis of vocal cord paralysis or for parathyroid adenoma localization. Visual inspection of control-CT images performed by an experienced radiologist did not reveal abnormalities in thyroid size, trachea dimensions or any other anatomical aspect. Inclusion of these control data was performed according to the regulations of the local ethical committee.

2.2 OBTAINING TRACHEA DIMENSIONS

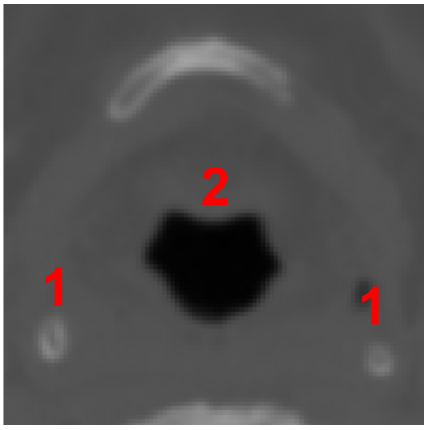
2.2.1 *Importing CT-data*

The CT-images were obtained from the Picture Archiving and Communications System (PACS) at the department of radiology. From this PACS, Digital Imaging and Communications in Medicine (DICOM) files were exported and subsequently imported in Matlab (version 7, The MathWorks Inc., Natick, MA) using the `dicomread` function from the image processing toolbox. This generated a $512 \times 512 \times n$-matrix, which thus contained n transversal grayscale images of 512 by 512 pixels. For adults trachea length is usually about 12 cm, so at a slice thickness of 2 mm, n will be approximately 60 [73].

2.2.2 *Semi-automated segmentation procedure*

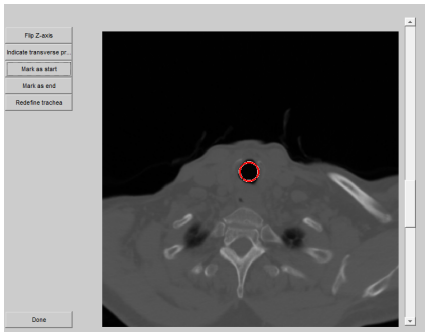
Using a specially devised Graphical User Interface (GUI) the user is able to scroll through the grayscale images. The area of interest was defined as the first slice caudal from the hyoid cartilage up until the carina of the trachea. This first slice, was selected when the tubercules of the hyoid bone were identified, see figure 2a. After the starting slice was identified the user clicks the *mark start* button, which translates this slice from a grayscale to a binary image. Research points out that the optimal threshold between air and tissue would be between -460 and -470 Houndfield Units (HU) [74]. Nevertheless, due to the close proximity of the trachea to the lungs, this caused leakage in some patients, i. e. lung tissue (which has a radio density of -500 HU) was also segmented. Therefore, through an iterative process, a more suitable threshold was found at -850 HU. Thus from the grayscale image all pixels below -850 HU were transformed to black (0-pixels) and all above -850 HU were transformed to white (1-pixels), see figure 2b. This lower threshold of -850 HU, might cause some pixels to be assigned to the trachea wall , whilst otherwise they would have been assigned to the lumen, thus thereby slightly shrinking the dimensions of the segmented lumen. Within this binary image Matlab's `imfindcircles` function tries to automatically locate the trachea by looking for the most circular object within the image. Subsequently the `imfindcircles` function generates a bright red circle in the original grayscale image at the presumed location of the trachea, as depicted in figure 2c. In most cases the `imfindcircles` function successfully identifies the location of the trachea, for the cases where the function fails to do so, the user is able to override the function and manually indicate the location of the trachea. Subsequently the user scrolls further through the grayscale slices until the first slice cranial from the trachea bifurcation, hereafter referred to as the ending slice, as depicted in figure 2d.

All slices following the starting slice, up onto the ending slice, are

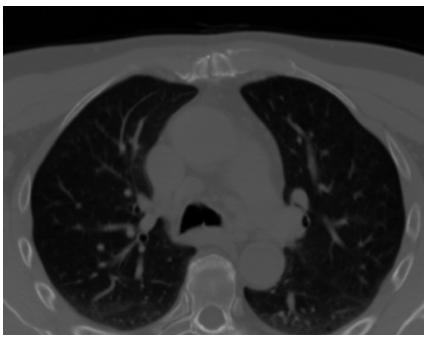


(a) Close-up of patients X's. trachea. 1: tubercules of the hyoid bone, 2: the trachea

(b) Binary version of the starting slice, threshold is -850 HU.



(c) Indicating the starting slice using the `imfindcircles` function.



(d) The endig slice.

Figure 2: Segmentation procedure.

thereafter also translated to binary images using the same process used in the starting slice. This yields a $512 \times 512 \times o$ binary matrix, whereby o is the number of slices within the region of interest. The segmentation starts at the binary starting slice. As mentioned earlier, the trachea location within this slice has already been identified. Subsequently, the image is cropped to a window of $k \times k$ pixels, where k is the number of pixels that would 25 mm. This can be calculated using the pixelspacing, which is a constant located in the DICOM-tags which reflects the width of a single pixel in mm. This value of 25 mm was chosen because it was large enough to accommodate the entire trachea lumen, but also not too big, to keep artefacts (black pixels which do not belong to the trachea lumen, see figure 3a) to a minimum. Because the largest patch of black pixels is believed to represent the actual trachea, large patches of lung-tissue might interfere with the artefact removal process, hence the cropping. After cropping the image a binary image remains which consists of a large patch of black pixels (the trachea lumen) and occasional smaller patches of

black pixels (artefacts), see figure 3a. Using the `regionprops` function of the Matlab image processing toolbox the area of all the patches of connected pixels is calculated. Thereafter, the smaller patches are removed until only the largest (the trachea lumen) remains, see figure 3b. Subsequently, the coordinates of the centre of this patch are calculated and used as an indication for trachea location in the subsequent slice. Within the next slice, again a cropped image is created around the centre of trachea of the preceding slice. This process continues until the trachea lumen within all slices is identified and stored in a single binary matrix. Usually, slices in this matrix will have a size of approximately 5×5 pixels.

(a) Cropped binary trachea image

(b) Same image, after artefact removal

Figure 3: Removing artefacts.

2.2.3 *Obtaining Centre Lumen Line*

As mentioned earlier, the goal of the segmentation process is to obtain the diameter of the trachea throughout its course. Since the lumen of the trachea is generally curved, calculating the diameter of the transversal slices would cause an overestimation of true diameter of the trachea, especially in regions with high curvature, due to the parallax effect, illustrated in figure 4. This parallax basically entails the overestimation of trachea dimensions due to analysis of cross-sections obliquely cut to its longitudinal axis skeleton, rather than orthogonal to this axis. Therefore, the diameter needs to be determined perpendicular to the Centre Lumen Line (CLL) of the trachea.

2.2.3.1 *Generating isosurface*

After identifying the trachea lumen, the binary matrix is smoothed using the `smooth3` function using the default convolution kernel at size [5,5,5]. This convolution kernel is a small matrix useful for smoothing images. Thereafter, using the smoothed matrix, an isosurface is created, which is a triangulated mesh consisting of faces and vertices, see figure 5. The x- and y-coordinates of the patch are multiplied with the pixelspacing and the z-coordinates are multiplied with the slice

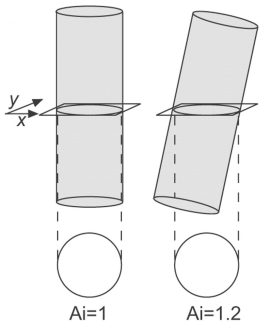


Figure 4: Illustration of the parallax effect [75] Analysing cross-sections obliquely cut to the longitudinal axis of a structure will cause an overestimation of dimensions.

thickness, so that the coordinates of the mesh reflect actual trachea dimensions in mm.

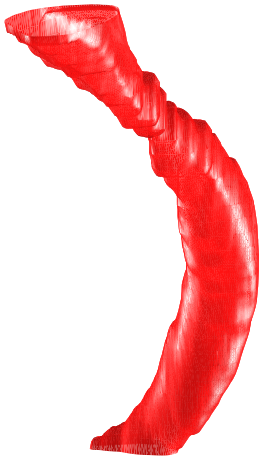


Figure 5: Isosurface of CT-data.

2.2.3.2 *Finding Centre of Lumen*

The CLL will be constructed using the coordinates of the centre of the lumen. Assuming the xy-plane as the transversal plane, it is possible to find these centres by generating a grid in the xy-plane at given z-coordinates, as depicted in figure 6. After removal of all coordinates of the grid outside the trachea lumen, using the `inpolyhedron` function, the distance for all of these coordinates to the vertices of the trachea wall are calculated. So if a grid-points lie inside the trachea lumen and the trachea wall consists of b vertices, $a \times b$ distances are calculated. Thereafter, for every grid-point the distance to the closest vertex in the trachea wall is calculated, the grid-point for which this distance is the largest is regarded to be the closest to the centre of the

lumen in that particular xy-plane. Logically, it follows that choosing a finer grid, will increase the accuracy of this process, but will also increase computational time. In order to reduce computational time a novel approach is introduced.

Firstly a 10×10 grid is constructed, whereby grid-points are spaced

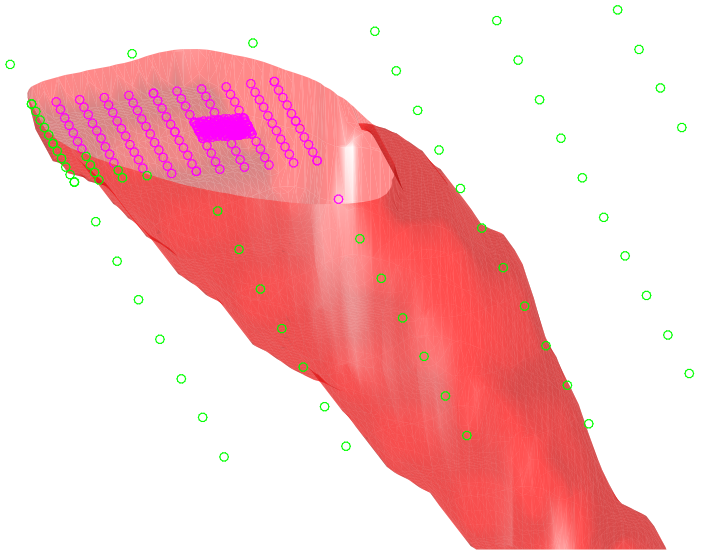


Figure 6: Finding the centre of lumen in xy-plane using increasingly finer grids. The lime coloured markers represent grid points outside of the lumen, and the magenta coloured markers represent grid points inside of the lumen.

at a distance Δ_x in the x-direction and Δ_y in the y-direction, thus the size of the grid in the xy-plane will be $10\Delta_x \times 10\Delta_y$, hereby $\Delta_x = \frac{x_{max}-x_{min}}{10}$ and $\Delta_y = \frac{y_{max}-y_{min}}{10}$, whereby:

x_{min} the smallest x-coordinate of the trachea wall

x_{max} the biggest x-coordinate of the trachea wall

y_{min} the smallest y-coordinate of the trachea wall

y_{max} the biggest y-coordinate of the trachea wall

From this grid, the grid-point closest to the centre of the lumen, hereafter referred to as the *Closest to Centre*-point (CC-point), is calculated using the above mentioned method. Thereafter a new 10×10 grid is constructed with the CC-point as centre, whereby grid-points are spaced at $\frac{1}{10}\Delta_x$ in the x-direction and $\frac{1}{10}\Delta_y$ in the y-direction. Using this new grid the CC-point is updated, due to the smaller grid-spacing this new CC-point is closer to the actual centre of the lumen. This process was repeated 10 times, usually this would yield a grid-spacing equal to 10^{-6} of its original value, see figure 6 for an illustration of this process. So for this process, for each slice $10 \times 10 \times 10 = 1000$ grid points would be assessed, whilst obtaining the same resolution with a single grid would require $10^6 \times 10^6 \times 100 = 10^{14}$ grid points. Since for all of these grid points the distance to each point

in the patch of the trachea wall is calculated, which is generally in the order of 10^4 points, the computational power required for each xy-plane is greatly reduced by using multiple grids.
The z-coordinates of the selected xy-planes are always 2 mm apart. So for a trachea wall of 100 mm in the longitudinal direction, 50 CC-points are calculated. Performing this procedure for the structure depicted in figure 5 took 188 seconds on a HP ProBook 4520s with a Intel® Core™ *i3* CPU M 370 @ 2.40GHz.

2.2.3.3 *Constructing Centre Lumen Line*

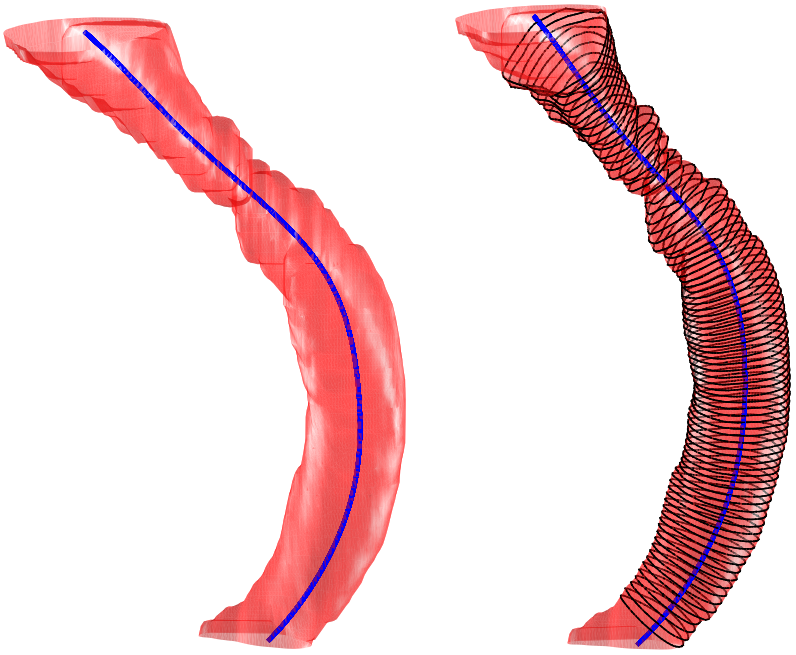
Using the `splinefit` function a direct spline interpolation is done of the generated CC-points. The spline works by calculating a function which follows these CC-points with minimal error. The higher the order of this function, the smaller the error between this spline and the CC-points will be. Nevertheless, if the order is too high, the spline will generate unwanted oscillations between the CC-points, which can be troublesome since the slope of the CLL is of importance. A way to circumvent this problem is to use multiple functions which each describe a portion of the line, i. e. breaking the line in several segments. Choosing to many breaks will also yield unwanted oscillations, whilst choosing too few breaks will increase error between the line and CC-points. Through an iterative process based on visual inspection, the best suitable number of breaks was chosen to be five. Such a CLL is depicted in figure 7a.

2.2.4 *Calculating diameters*

Perpendicular to this CLL, planes are generated spaced 2 mm apart, thereafter the locations at which the planes intersect with the trachea wall are calculated using the `intersectPlaneSurf` function, as depicted in figure 7b. Using a rotation matrix the 3D-coordinates from these intersection sites are translated to a 2D-coordinate system. This will generate closed 2D-polygons, as seen in figure 8 these polygons are generally not circular. Since the mathematics in section 2.3 assume circular dimensions, a circular equivalent diameter is used, which is calculated using equation 1 [76].

$$d = 1.55\frac{A^{0.625}}{P^{0.25}} \tag{1}$$

Whereby P is the perimeter in m and A is the cross sectional area in m^2. This equivalent diameter is the diameter of a circular tube that gives the same pressure loss as an equivalent oval tube and was devised by Heyt *et al.* based on experiments in oval ducts [76].



(a) Constructing the Centre Lumen Line (blue line) within the trachea lumen (red) using the `splinefit` function.

(b) Cross-sections (black lines) of the trachea lumen (red structure) perpendicular to the CLL (blue line)

Figure 7

2.2.5 *Estimating healthy trachea dimensions*

After obtaining the A and P of these polygons, it is possible to calculate d throughout the course of the trachea, as depicted in figure 9, showing trachea diameters of a single patients plotted as a function of trachea length. The diameters which were obtained from the pre-surgery CT-scan are referred to as $d_{disease}$. In this patient, the diameter dropped from about 12.3 to nearly 5.6 mm. The reference point is chosen at the first peak in diameter after the vocal chords. In healthy trachea's the vocal chords will usually impose a constriction of approximately 40%, thereafter the diameter stays fairly constant[68]. Since

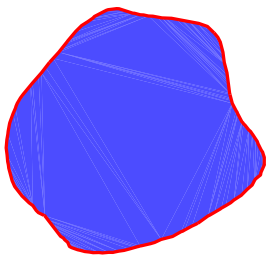


Figure 8: Perimeter (red) and area (blue) of a 2D-polygon representing an cross-section of the trachea perpendicular to the CLL

the correlation for inter-individual dimensions is rather weak but for intra-individual trachea dimension is quite strong, it is assumed that this peak in cross sectional diameter after the vocal chords (d_{ref}) is a good estimate for cross sectional diameter in the rest of the healthy trachea [77, 73, 78, 79].
However, despite strong correlations for intra-individual dimensions, variation is expected to a certain extent as a consequence of Healthy Anatomical Variation (HAV). The limits of these HAVs are important to assess which anatomies are pathological. Determination of these limits is done by assessing the confidence intervals for HAV in the control subject, using the 95%-percentile above and under d_{ref}.

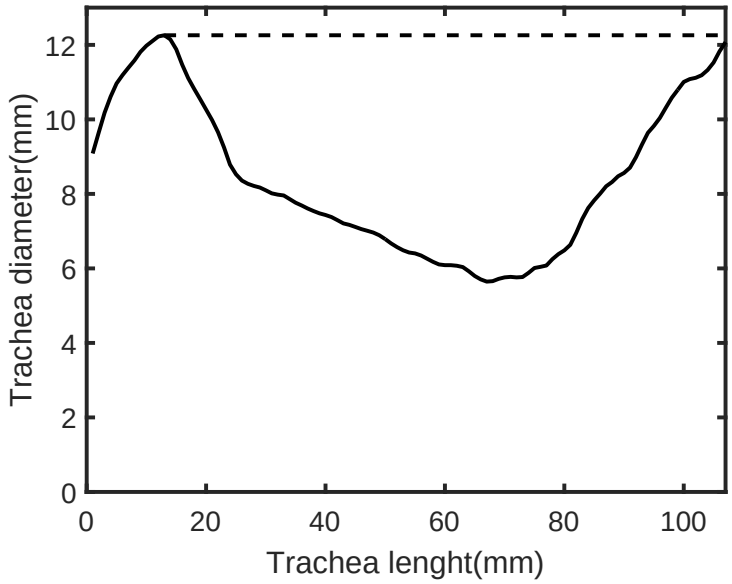


Figure 9: This figure illustrates the trachea diameter of patient 2 throughout the course of his trachea. The first peak (d_{ref}) is used as a estimation for $d_{healthy}$ and extrapolated over the course of the rest of the trachea (interrupted line), whereas the actual calculated trachea diameters represent $d_{diseased}$ (solid line).

2.3 SIMPLIFIED CFD: THE MVC

To obtain a physiologically meaningful parameter for the impact of trachea compression, the goitre-induced changes in trachea dimensions were translated into impairments in peak airflow (ϕ in L s^{-1}) and maximal minute ventilation (MMV in L min^{-1}). If a goitre leads to local trachea compression, the reduced dimensions will increase local airway resistance and cause an increase pressure drop over the affected area. By viewing the trachea as a circular tube, the pressure drop for fully developed turbulent flow can be calculated according to equation 2 [80]:

$$\Delta p = 12 C_f \rho V^2 \tag{2}$$

Where Δp is the pressure drop in Pa, C_f is the Darcy friction factor which is a dimensionless quantity, ρ is the air density in kg m^{-3} and V is the air velocity in m s^{-1}. V is defined by equation 3 [81]:

$$V = \frac{\Phi}{A} \tag{3}$$

Where Φ is the air flow in L s^{-1} and A the cross sectional area in m^2 which is defined by equation 4:

$$A = d^2 \times \frac{\pi}{4} \tag{4}$$

Where d is the diameter in m. The establishment of C_f depends on the Reynolds number (Re) which is defined by equation 5 [80]:

$$Re = \frac{\rho \times V \times d}{\mu} \tag{5}$$

Hereby μ is the dynamic viscosity in N s m^{-2}. Re is a indication for the state of flow and generally Re > 4000 indicates a fully developed turbulent flow [80]. For fully developed turbulent flow equation 6 applies [81]:

$$C_f = \frac{k \times l}{d} \tag{6}$$

Hereby k is a dimensionless quantity which describes the roughness of the trachea wall and l is the length of the tube in m. By substituting equation 6 into equation 2 we obtain equation 7.

$$\Delta p = 12\frac{k \times l}{d}\rho V^2 \tag{7}$$

Subsequently, substitution of equation 4 in to equation 3 yields equation 8:

$$V = \frac{4\Phi}{d^2 \times \pi} \tag{8}$$

Thereafter, equation 8 is substituted into equation 7, which yields equation 9:

$$\Delta p = 12\frac{k \times l}{d}\rho\frac{16\Phi^2}{d^4 \times \pi^2} \tag{9}$$

Finally rewriting equation 9 yields equation 10:

$$\phi = \sqrt{\frac{\Delta p \times d^5 \times \pi^2}{192k \times l \times \rho}} \tag{10}$$

However, in practice the trachea diameter will most likely vary throughout its course, especially if its dimensions have been affected by goitre or other pathologies. To account for this the trachea will be subdivided in small sections. For each of these sections the diameter will

be calculated using a segmentation procedure. This diameter is thereafter used to calculate the pressure drop of the corresponding section of tracheal lumen. Similar to an electrical circuit with tandem connections, as illustrated in figure 10, these individual pressure drops will be summed up to approximate the pressure drop of the entire trachea.

In this model the lungs are viewed as a motor which needs to gen-

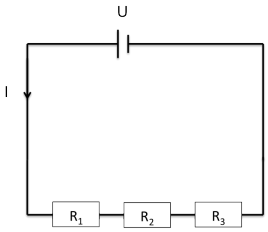


Figure 10: Example of an electrical circuit with tandem connections.

erate a certain pressure drop (Δp) at the caudal end of the trachea to overcome its airway resistance (R) and generate a flow (Φ) at the cranial end of the trachea. Diminishing the tracheal dimensions will most likely increase this airway resistance and thereby impair the amount of flow the lungs are able to generate. If the situation whereby trachea dimensions of a patient are diminished due to goitre are defined as the *diseased state* and the situation whereby trachea dimensions are not affected by any pathologies as the *healthy state*, we could introduce a novel measure which reflects this ventilation impairment due to GITC, as illustrated by equation 11:

$$MVC = \frac{\Phi_{disease}}{\Phi_{healthy}} \times 100\% \tag{11}$$

Whereby $\Phi_{disease}$ is the amount of airflow the lungs are able to generate in the diseased state in $m^3\ s^{-1}$, $\Phi_{healthy}$ is the amount of airflow the lungs are able to generate in the healthy state in $m^3\ s^{-1}$ and MVC is the Maximal Ventilation Capacity, a dimensionless quantity which illustrates the percentage of healthy ventilation capacity remaining. Note that it is assumed that the power of the lungs does not alter with development of GITC, thus Δp in the healthy state is assumed equal to Δp in the diseased state. By assuming $Re > 4000$, thus a turbulent flow, it is possible to rewrite equation 11 to equation 12[1]:

$$MVC = \frac{\sqrt{\frac{\Delta p \times \pi^2}{192 k \times l \times \rho}} \times d_{disease}^{\frac{5}{2}}}{\sqrt{\frac{\Delta p \times \pi^2}{192 k \times l \times \rho}} \times d_{healthy}^{\frac{5}{2}}} \times 100\% \tag{12}$$

Hereby $d_{healthy}$ reflects the presumed trachea diameter in the healthy state $d_{disease}$ the actual trachea diameter. However, the magnitude of

1 Re was found to be larger than 4000 for all patients and subjects throughout the entire trachea, even at flow levels as low as 1 L s^{-1}, where ρ = 1.1644 kg m^{-3} and $\mu = 1.8861 \times 10^{-5}$ N s m^{-2}

this diameter varies along the course of the trachea. To apply calculated trachea diameters to equation 12, the trachea diameters should be replaced by variable representing these dimensions. This variable is referred to as $d_{effective}$ and may be used to calculate the MVC according to equation 13

$$MVC = \frac{\sqrt{\frac{\Delta p \times \pi^2}{192k \times l \times \rho}} \times d_{effective}^{\frac{5}{2}}}{\sqrt{\frac{\Delta p \times \pi^2}{192k \times l \times \rho}} \times d_{healthy}^{\frac{5}{2}}} \times 100\% \tag{13}$$

Whereby $d_{effective}^{\frac{5}{2}}$ can be obtained by equation 14:

$$d_{effective}^{\frac{5}{2}} = \frac{1}{N} \sum_{i=1}^{N} d_{disease_i}^{\frac{5}{2}} \tag{14}$$

Whereby N is the number of slices for which a diameter is obtained. Removal of common factors yields equation 15:

$$MVC = \frac{d_{effective}{}^{\frac{5}{2}}}{d_{healthy}{}^{\frac{5}{2}}} = \frac{1}{N} \sum_{i=1}^{N} \left(\frac{d_{disease}}{d_{healthy}}\right)^{\frac{5}{2}} \tag{15}$$

Literature shows that in healthy humans flow patterns during rest are at least transitional between laminar and turbulent and during exercise flow patterns are turbulent [82, 83]. As seen in equation 5, diminishing trachea dimensions will most likely increase Re and thereby occurrence of turbulent flow, based on these observations it was concluded that equation 15 applied for healthy as well as pathological trachea's. Because equation 15 allows to calculate the MVC solely using d, constants such as k,ρ and μ do not have to be assessed.

2.3.1 *Ventilation to flow*

As stated in section 2.3, the MVC is believed to represent an individuals maximal ventilation capacity, but it is calculated using flow. Hereby it is assumed that flow is proportional to ventilation. This can be motivated by assuming inspiratory flow:

$$\Phi_i(t) = f(t) \tag{16}$$

For $t_0 \leqslant t \leqslant t_1$ and $f(t_1) = f(t_0) = 0$. Whereby Φ_i is the inspiratory flow in L s^{-1}, t_0 is the start of inspiration and t_1 is the end of inspiration in seconds. Additionally expiratory flow will be defined as:

$$\Phi_e(t) = g(t) \tag{17}$$

For $t_1 \leqslant t \leqslant t_2$ and $g(t_2) = g(t_1) = 0$. Whereby Φ_e is the expiratory flow in L s^{-1}, t_1 is the start of expiration and t_2 is the end of expira-

tion in seconds. Using equation 16 and 17 it is possible to determine the $\frac{I}{E}$-ratio:

$$\frac{I}{E}\text{-ratio} = \frac{t_1 - t_0}{t_2 - t_1} \tag{18}$$

Hereby the $\frac{I}{E}$-ratio indicates the ratio between the duration of inhalation and expiration. Additionally this yields the Ventilator Rate (VR):

$$VR = \frac{60}{t_2 - t_0} \tag{19}$$

Whereby VR is the number of breaths per minute. This makes that the volume inhaled during one breath is equal to $\int_{t_0}^{t_1} f(t)dt$. Multiplying this volume with the VR gives us the ventilation (V_e) in L min^{-1}:

$$V_e = \int_{t_0}^{t_1} f(t)dt \times \frac{60}{t_2 - t_0} \tag{20}$$

The maximal flow during inhalation is defined by the highest value of $f(t)$ for which holds $f'(t) = 0$ and $f''(t) > 0$. From equation 20 it is obvious that $f(t)$ is proportional to Ve, just as long as $\frac{60}{t_2 - t_0}$ is not altered over time, i. e. flow is proportional to ventilation if the $\frac{I}{E}$-ratio and VR do not change over time.

2.3.2 *Interpretation of the MVC*

In healthy subjects without any trachea narrowing the MVC should be close to 100%. However, anatomical variations, for example compression caused by the aorta may produce mild local narrowing, and this will reduce the MVC to a value less than 100%. To account for this variation, and to establish the MVC's lower limit of normal in patients without goitre, MVCs were also calculated from control data.
As illustrated in section 2.3 it can be proven mathematically that the MVC is equal to the sum of the square root of all individual slice diameters to the power of five in the healthy state divided by the sum of the square root of all individual slice diameters to the power of five in the diseased state. Based on the assumption that any impairment in flow causes a proportional decrease in minute ventilations (arguments outlined in section 2.3.1), the MVC quantifies the individual's remaining Maximal Minute Ventilation (MMV) capacity, expressed as percentage of normal.

2.4 RESULTS

CT scans of ten patients with goitres, five females and five males, obtained three months before and three months after surgery and ten non-goitre patients were used to perform an initial evaluation of

the potential of applicability and usefulness of the method. The slice by slice assessment of tracheal diameters in non-goitre and goitre patients is explained in section 2.2. The MVC's in control subjects ranged from 65.5 to 102.5%, and from 38.9 to 100.5% in patients with goitre, see figure 13. These MVC values over 100% were caused by portions of the trachea that were wider than the reference value, i. e. trachea dimensions exceeded the predicted healthy dimensions. Surgery led to an increase in trachea diameters and a higher MVC in all cases except for patients 7 and 10, see figure 12. Possible explanations for this are discussed in section 2.5. Using the trachea dimensions of the control patients (figure 11), confidence intervals for Healthy Anatomical Variation (HAV) are determined and also depicted in figure 12. The mean pre-surgery and post-surgery goitre MVC's were 65.2 $\pm$ 18 and 84.8 $\pm$ 20. Control subjects had a mean MVC of 82.8 $\pm$ 12.

2.5 DISCUSSION

Results of this study suggest that CT-scans of the trachea may become helpful to improve the evaluation of goitres. CT generated data not only permit visual inspection of the structual impact of goitre on trachea dimensions, they can also be transformed into data that quantify the degree of functional limitations. In this approach trachea dimensions are used to calculate the degree of ventilation impairment, relative to the individuals predicted normal ventilation capacity that is based on trachea width of unaffected areas. The degree of ventilation impairment is obtained by translation of trachea narrowing into limitations in peak airflow based on physical assumptions that describe the relation between structure and airflow (section 2.3). The end results of this calculation procedure is the maximal ventilation capacity (MVC), i. e. , a value representing an objective assessment of the degree of ventilation restriction caused by a compressive goitre in an individual patient. Preliminary results suggest that MVC's of 65% represent the lower limit of normal. A greater number of control subjects will be needed to establish this lower limit of normal more firmly.

Maximal trachea peak airflow is always limited. It is determined by trachea diameter and the power of ventilatory muscles to generate a pressure gradient for inspiration. The increased resistance induced by trachea narrowing will cause an additional pressure gradient over the affected area, and this will hinder the generation of airflow. Any increase in airway resistance will lead to a decrease in peak airflow, which results in a decline in minute ventilation and the capacity to perform physical work.

The proposed method may not only help to decide when to perform goitre volume reduction for ventilation impairment, it can also be

used to monitor the impact of goitre growth during follow-up, or to quantify the efficacy of an intervention. Another advantage is that results are not affected by the impact of concomitant pulmonary or cardiac disease. This is an important issue in daily practice, because GITC often occurs at an age when cardiac or pulmonary disease are also frequently present. Prediction of the potential benefit of goitre volume reduction, independent of cardiac or pulmonary status is valuable information often needed to make balanced decision in elderly subjects. In contrast, commonly used pulmonary function tests often cannot discriminate between the relative contributions of tracheal cardiac or pulmonary restrictions.

To date, results of goitre surgery are poorly documented. Although symptomatic improvement is generally recorded, it is not common practice to assess the improvement in trachea dimensions after surgery, and it is commonly assumed that goitre reduction will lead to trachea expansion [20, 21, 22, 61, 23, 24, 25]. Figure 12 might imply that this assumption is not valid, responses varied widely. Some patients demonstrated increase of trachea dimension, where others had no or little improvement in trachea diameter (patient 7 & 10). Irreversible damage of tracheal rings may hamper normalization of trachea dimension. However, a three-month observation period may also be too short to document full recovery. Another explanation could be that the trachea's which did not show improvement, were not pathological to begin with. Analysis of the control subjects suggests that these patients fall within the realm of benign anatomical variation. However, due to the quantity of data assessed it is not yet possible to support this statement from a statistical point of view. Therefore, further studies in larger groups of patients will be needed to reveal what is healthy anatomy and what is pathological and how different patients respond to thyroid surgery.

2.5.1 *Limitations*

The current model has been tested in a small group of subjects. Several aspects will need further evaluation and validation in a large group of subjects. A more accurate assessment of normal values by inclusion of larger numbers of control subjects, assessment of the reproducibility of the trachea measurements and its predicted normal values, accuracy assessment of predicted versus measured $MMV_{healthy}$, and studies comparing CT-trachea derived parameters and conventional pulmonary function tests will be needed to further examine the diagnostic value of this new approach.

At first sight the method seems like a straight-forward and practical tool for assessment of GITC. Nevertheless, its applicability has not been quantified. For instance, it is not known whether various physicians will define the same region of interest. Furthermore, it has not

been researched what the effects of a factors, such as the defined region of interest, will have on the MVC.
It is also well recognized that the model is a simplification, focussed on cross-sectional area reconstruction. During this simplification some processes have not been accounted for and were deliberately ignored. These processes included:

1. The effect of trachea deviation

   Critical values for Reynolds number are higher for tubes with higher curvature [84, 85]. This effect can delay the occurrence of turbulent flow, which yields a higher airway resistance than laminar flow. However centrifugal forces generated by curvature can increase Δp which can negate the effect of increased critical numbers. Measurements on infants show that tracheal deviation causes an increase in Δp [84]. Whether these effects are significant in adults with tracheal compression in unknown.

2. Approximation of geometry

   Obtaining trachea dimensions from the patients physical anatomy in done through several steps. Firstly DICOM-files are generated through a CT-scan. These gray-scale images are converted to binary images which are thereafter transformed into an isosurface. Through this isosurface a CCL is constructed, and perpendicular to this CCL the diameter of the isosurface is calculated, which then is believed to represent the true diameter of the original physical anatomy of the patient. All these steps might impose a margin of error in the calculation of the dimensions. Comparing the current algorithm to validated third party software can help to quantify these errors. Another options would be to analyse a phantom with known dimensions using the current method. This phantom should include realistic trachea geometry and radio-density to ensure that obtained results representative for the assessment of actual trachea's.

3. Collapse of trachea

   Generally goitre is more likely to hinder inspiratory flow, rather than expiratory flow. This might be because during expiration the lungs generate a positive pressure difference, this enables them to generate flow through a mild stenosis because the stenosis is kept open by the positive pressure. However, during inspiration the lungs will generate a negative pressure gradient, which can constrict the trachea even further, and thereby hindering airflow[86]. Therefore it is important that CT-images images are obtained during inspiration, as is instructed in clinical practice. Nevertheless, to what extent these instructions are actually followed is quite questionable.

4. Irregular dimensions

   Sudden variations in cross sectional area and irregularities in shape will increase turbulence and therefore airway resistance [87]

5. Ventilation to flow

   In section 2.3.1 it is motivated that ventilation is proportional to flow if $\frac{I}{E}$ and VR are not altered over time. However, it is well-known that GITC mainly hinders inspiration, thereby increasing $\frac{I}{E}$. Even if expiration is not impaired, the total time needed for a single breath will increase, thereby lowering the VR.

6. Validity of d_{ref}

   As motivated in section 2.2.5, the vocal chords will usually impose a constriction of 40% on the trachea. The maximal dilation of the trachea after the constriction is therefore used as a reference for further trachea portions. This dilation after the vocal chords was observed in all 10 patients, however whether severe goitres will also affect d_{ref} was not assessed. Therefore it is not know whether d_{ref} is actually a reference value for healthy trachea dimensions.

2.6 RECOMMENDATIONS

The extent to which the processes mentioned in section 2.5.1 influence the accuracy of the model is unknown. Individually assessing all of these factors will most likely be a very costly project. Nevertheless, it is possible that the error imposed by these processes described are insignificant when it comes to devising a meaningful parameter for the assessment of GITC. Ideally this hypothesis would be tested by comparing the MVC with a valid golden standard. Unfortunately, the performance of the current golden standard is suboptimal, mostly because it is only able to identify very severe cases of GITC. Nevertheless, if the MVC shows poor agreement with the golden standard in even these severe cases of GITC, this would imply that the MVC is not a valid measure for GITC.
Due to the lack of a sensitive golden standard it is quite difficult to assess the relevance of the MVC for more subtle cases. Although other measures such as Self Reported Symptoms(SRS), Radiological Parameters (RP) and impulse oscillometry spirometry (IOS) also have their drawbacks (see chapter 1), agreements of these different measure on individual cases can shed light on the validity for the MVC on these more subtle cases. In current clinical practice only CT-data and SRS are assessed. Thus for assessment of a wide spectrum of measures, a prospective study is indicated with a more comprehen-

sive assessment of GITC. However, due to the limited supply of patients, acquiring a sample size suited for statistically sound analysis can be a protracted process. Therefore it is recommended to perform a prospective study to assess measures such as lung function tests and questionnaires (chapter 4), and use readily available CT-data for retrospective analysis of the MVC and RP (chapter 3). Additionally this retrospective data can be used to determine healthy anatomical variations for the prediction of healthy trachea dimension based on pre-surgery data.

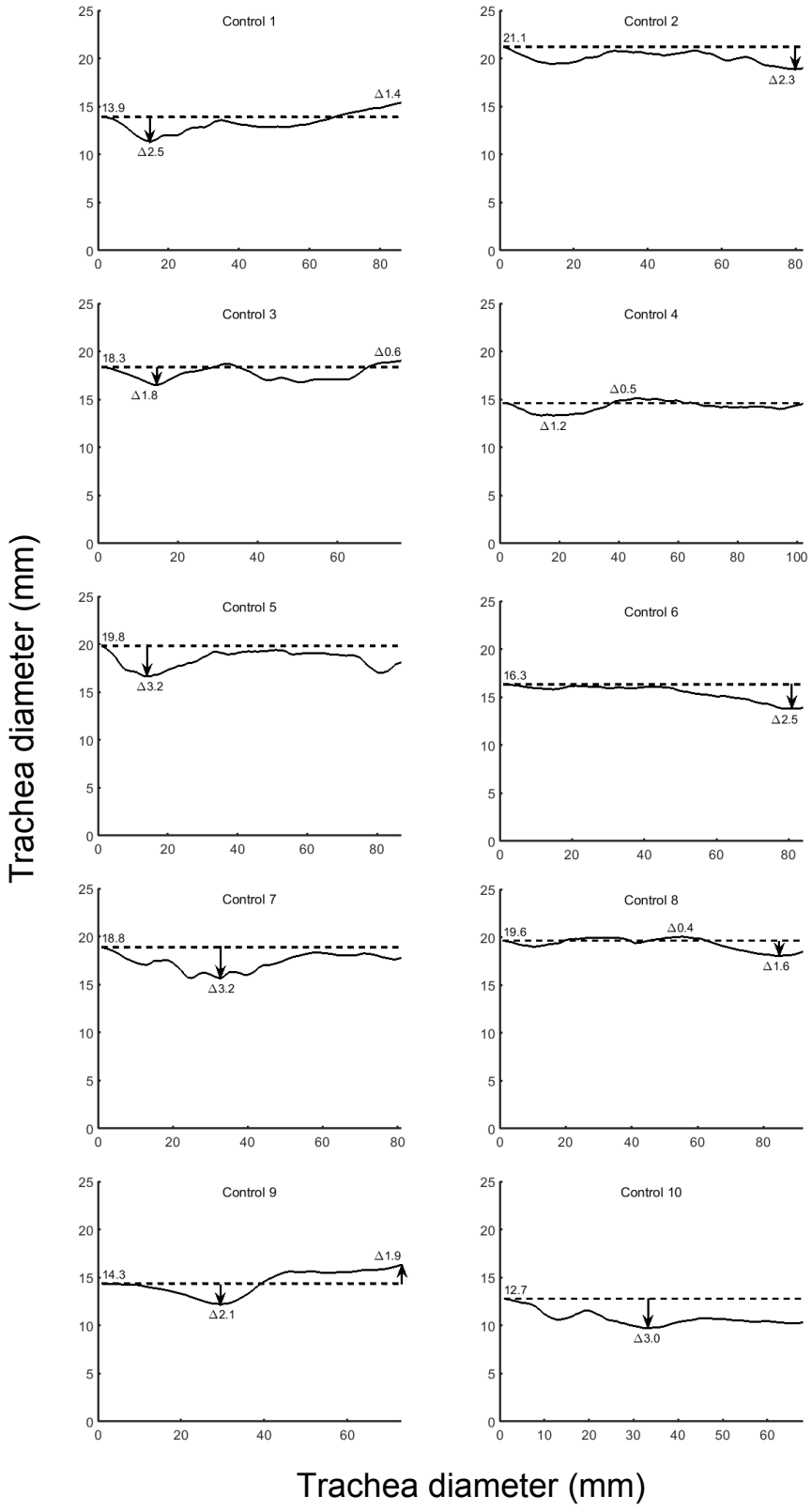


Figure 11: Trachea diameter (vertical axis (mm)) plotted against trachea length (horizontal axis (mm)) for all control subjects.

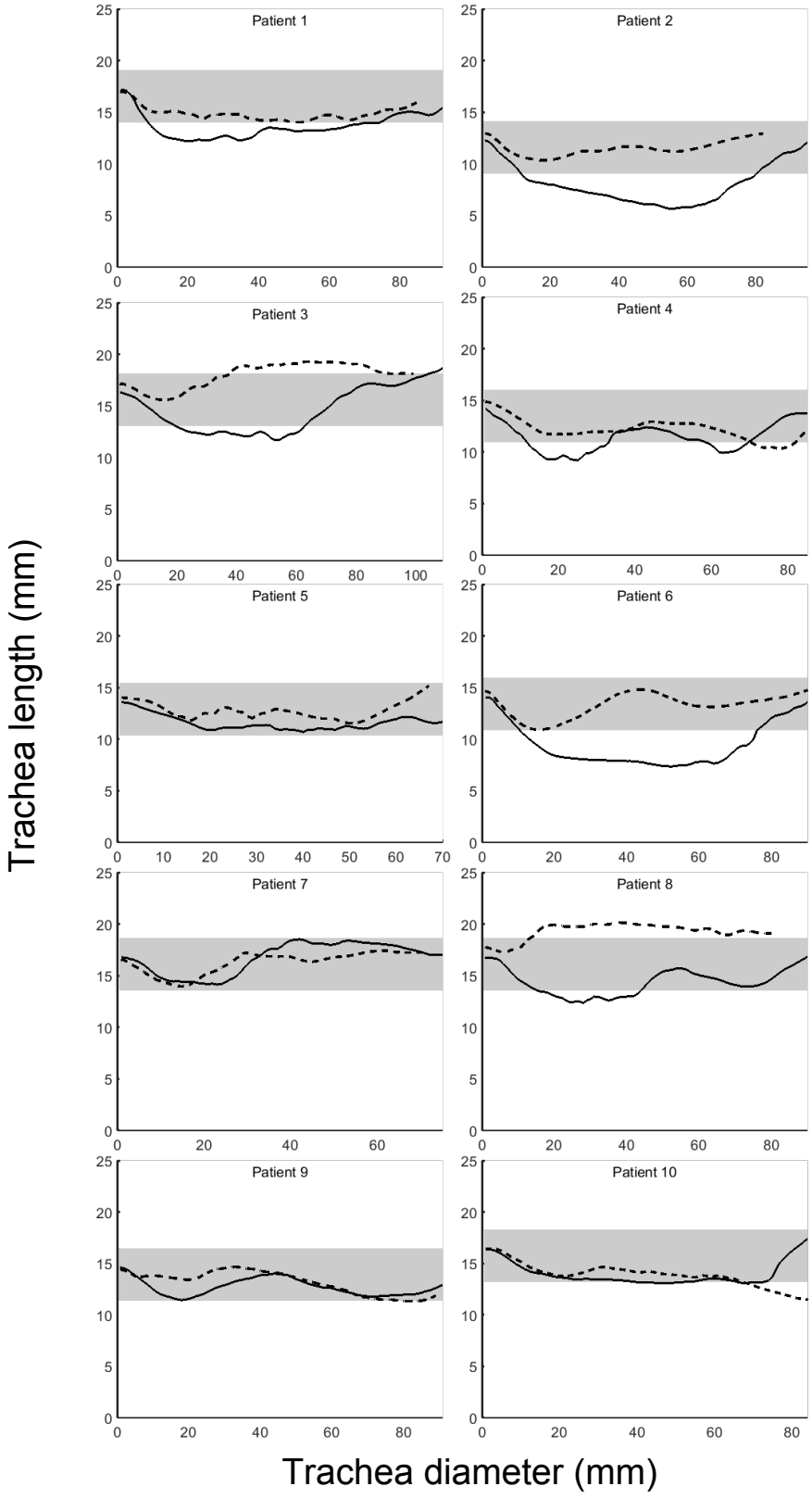


Figure 12: Pre- (solid-line) and post-surgery (interrupted line) trachea diameter (vertical axis (mm)) plotted against trachea length (horizontal axis (mm)), with 95%-confidence intervals (grey area) based on the anatomy of controls.

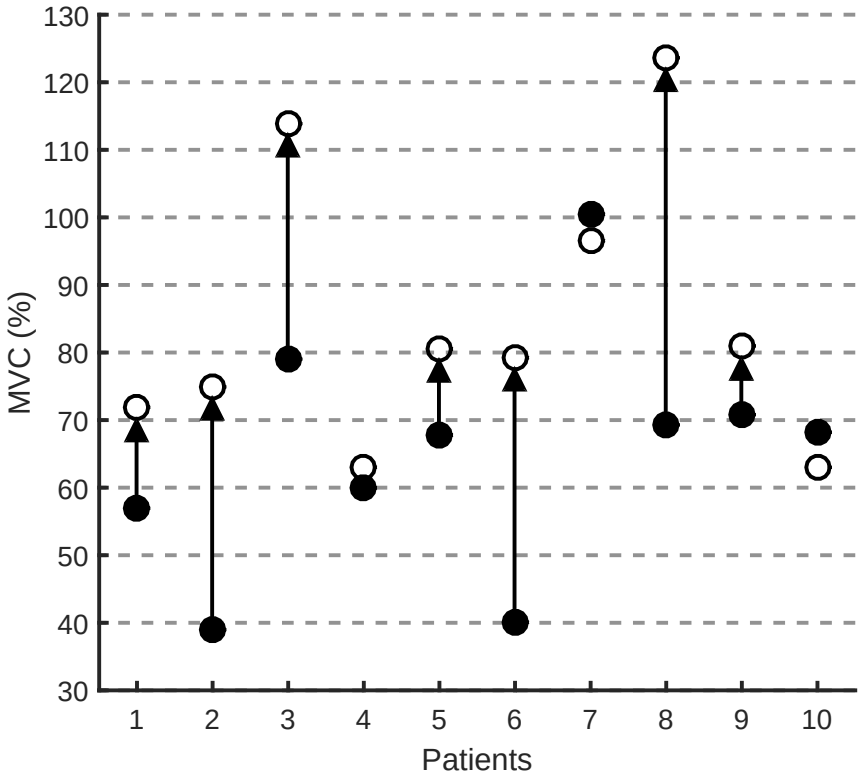


Figure 13: MVC's pre- (closed dot) and post-surgery(open dot).

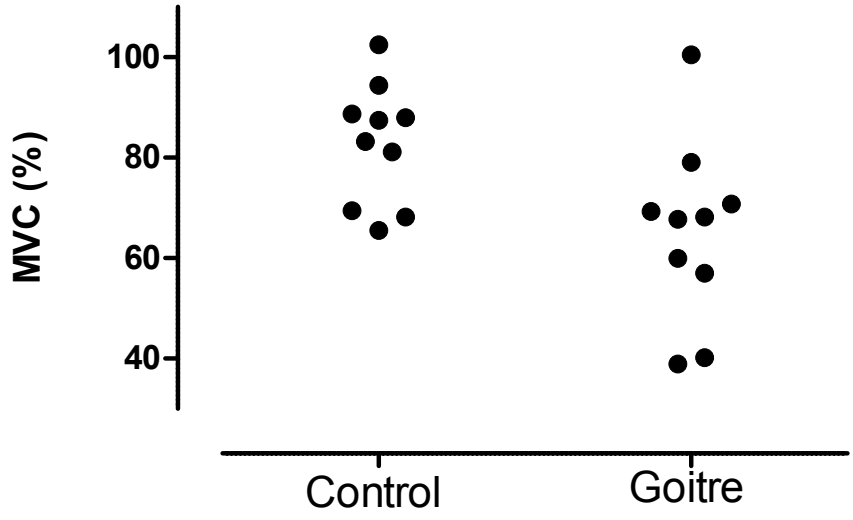


Figure 14: MVC's for goitre and control patients.

3 RETROSPECTIVE STUDY

3.1 BACKGROUND

Goitre is an palpable enlargement of the thyroid which may cause trachea malformation. This malformation can raise airway airway resistance, thereby causing ventilatory impairment [3]. Various modes of treatment with different efficacies are available for the treatment of Goitre Induced Trachea Compression (GITC) [4, 5, 6, 7, 15, 17, 18, 19]. To decide on the most suitable therapy it is important to assess the clinical impact of GITC accurately.
Due to the slow progressive nature of the disease, patients are often unaware to what extent GITC impairs their physical capabilities, which makes the assessment of Self Reported Symptoms (SRS) unreliable [61]. An alternative is the assessment of Radiological Parameters (RP), but these often lack a firm objective basis and functional consequences remain unknown [66, 67]. The current golden standard for assessment of GITC is whole-body plethysmography in combination with spirometry, also known as Pulmonary Function Tests (PFT) [28]. However, issues with standardisation, effort-dependence and functional interpretation make this a suboptimal measure for assessment of GITC [29, 35, 36, 68]. Moreover, sensitivity of PFT is low, and therefore only able to point out very severe cases of GITC [34, 68]. Physicians strive to intervene prior to severe escalation of GITC in order to keep patient discomfort to a minimum, therefore they resort to more sensitive but less reliable measures such as SRS and RP for assessment of GITC in clinical practice. A flowchart illustrating the decision-making process in current clinical practice is given in figure 15[1].

To provide insights in the shortcomings of commonly assessed methods and propose possible improvements, a literature-study was initiated.

3.2 LITERATURE OVERVIEW

Two prominent studies researched the correlation between RP and SRS.
In 2012 Stang *et al.* assessed 1081 patients suffering from goitre and reported a strong correlation between Positional Dyspnea (PD) and Trachea Constriction (TC) [24]. Another recent study by Shin *et al.*

1 As observed in the Rijnstate Hospital Group

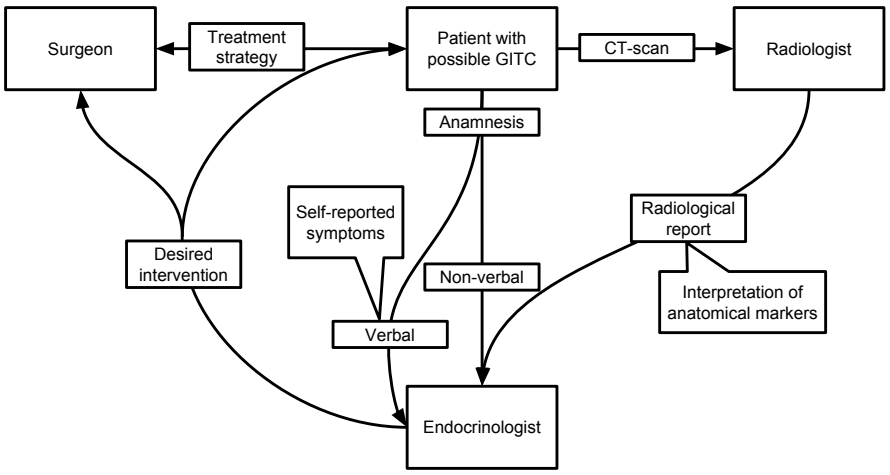


Figure 15: GITC-assessment in current clinical practice.

from 2011 also reported a significant correlation between a sensation of dyspnea and TC in 200 patients suffering from goitre ($p < 0.001$) [21].

Several other studies compared RP and/or SRS to results of PFT. In 2010 Albareda *et al.* compared the results of Flow-Volume Loops (FVL) to RP and SRS in 50 patients with endothoracic goitre [34]. They concluded that no clinical or radiological variable was related to Upper Airway Obstruction (UAO) diagnosed using FVL. Additional indices such as FEV_1 and FEV_1/PEF also showed poor correlation to UAO. However, FEV_1/PIF did show a significant correlation to UAO ($p = 0.003$). This is most probably since GITC mainly hinders inspiration, thereby making expiratory indices unfit to monitor GITC[2].

Albareda assessed TC by measurement of diameter and Cross Sectional Area (CSA) at a reference point 2 cm above the carina and at the site of maximal trachea constriction. They based this reference point on findings of Melissant et al., which stated that intra-individual dimensional differences are minor at the different levels of the trachea. The trachea diameter 2 cm above the carina may be 10% greater compared to the middle part, and this observation is highly reproducible among subjects [88]. However, this statement was formed based on observations in healthy subjects, whether this is also valid in patients suffering from TC was not verified.

The results of Albareda raised some questions. The maximum observed value for TC in terms of CSA reduction was 67%, whereby for most patients TC was well below this level (mean = 15%). In view of the findings of the earlier mentioned study by Brouns *et al.* (see chapter 1), which suggests that constriction below 50% are insignificant in terms of UAO, it is questionable whether these levels of TC were sufficient to elicit significant ventilatory impairment.

2 additional motivation provided in chapter 1

Another study by Gittoes *et al.* compared Flow-Volume Loops (FVL) to RP and SRS in 153 patients with goitre. They also reported poor correlation between FVL and these other measures. However, they used plain x-ray instead of CT, whereas it has been proven that CT is more accurate [31]. Moreover, they did not assess the absolute magnitude of TC, but verified if it exceeded a certain threshold. As Brouns *et al.* clearly demonstrated, the magnitude of TC is decisive for its consequences, therefore it is not remarkable that this study found a poor agreement between TC and UAO [68].
In 1994 Melissant *et al.* researched the correlation between RP and PFT [88]. Their methods for assessment of TC was similar to those in the study by Albareda *et al.*. Similarly to Albareda, they reported a poor correlation between TC and PFT. They put forth two explanations for these findings. Firstly trachea dimensions are known to change with alterations in position [89]. And secondly, trachea dimensions are known to changes during ventilation [90].
In 1999 Bonnema *et al.* researched the correlation between Minimal Cross-Sectional Area (CSA_m) and PFT in 23 patients suffering from goitre [91]. They found a strong relation between this CSA_m and FIF_{50} ($p < 0.001$), see figure 16.
The findings of Bonnema *et al.* are remarkable. To our knowledge it

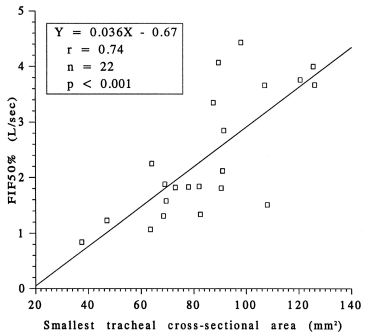


Figure 16: Minimal trachea CSA versus FIF_{50} [91]

is the only study that identified a correlation between trachea dimensions and PFT-indices. A presumable important difference between this study and others, is the quantification of absolute trachea dimensions (Minimal Cross-Sectional Area: CSA_m), rather than relative ones such Percentage Trachea Compression (%-TC).
One might expect that a relative index would be a better predictor for UAO than an absolute one, since persons who require more ventilation for certain physical activities will have larger trachea's. Data from the Consolidated Human Activity Database (CHAD) clearly indicate that an individual's ventilatory demands greatly depend on age, gender, bodyweight and height [92]. Nevertheless, literature indicated that gender greatly affects trachea dimensions, but bodyweight, age and stature do not [78]. Moreover, as the research of Brouns *et al.* indicates, only severe levels of TC (> 50%) may hinder ventilation in

humans. This suggests that in healthy individuals, the trachea has an overcapacity. This makes the theory that one's trachea dimensions are not affected by their build, very plausible. However, since ventilatory demands are influenced by one's physical size and age, i. e. required levels of airflow are dependent on an individuals physical size, it is also likely that a CSA_m that causes ventilatory impairment is not universal for all persons. Nevertheless, to our knowledge there has not been a study which compared both absolute (CSA_m) and relative indices (%-TC) to validated indices for the assessment of GITC. Thus it remains unknown whether this hypothesis is actually true.
Recently a new approach to assess the clinical impact of GITC, based on CT-imaging and simplified computational fluid dynamics, was introduced (chapter 2). With this approach CT-derived anatomical data are translated into functional data that express the Maximal Ventilation Capacity (MVC) of the trachea. The MVC of the trachea in a patient with GITC ($MVC_{patient}$) is then expressed as the percentage of this individual's predicted normal MVC. This individual specific normal values is based on trachea dimensions observed in unaffected trachea areas. One of the advantages of this method over traditional assessment of RP is its well-defined interpretation.
Ideally, this MVC would be compared to the golden standard (which are PFT). However, as mentioned in section 7.0.1, despite being the golden standard, PFT are not commonly assessed in clinical practice. Nonetheless, as figure 15 shows, SRS and RP are commonly assessed in current clinical practice. Therefore, fairly large amounts of these data are readily available for retrospective analysis.

3.3 AIM OF THE STUDY

This study aims to use readily available retrospective data, to evaluate %-TC and CSA_m and compare them to SRS. Additionally, this chapter will follow-up the pilot-study that introduced the MVC, in order to validate some of the assumptions described in chapter 2 as well as the establishment of reference values using control-data.

3.4 SUBJECTS

CT data of 55 control subjects (31 females and 21 males) and 74 patients with thyroid disease that warranted thyroid surgery were obtained. All subjects were older than 18 years. Control subjects had an indication for CT-scanning of the neck and mediastinum, not related to thyroid diseases, but as part of the work-up for parathyroid adenoma localization, lung cancer staging or analysis of vocal cord paralysis. Selection was based on the absence of abnormalities that might affect trachea dimensions, as judged by an experienced radiologist.

		Males	Females	Age (years)
Controls		24	31	58.5 ± 12.8
Patients	Pre	5	12	59.0 ± 10.4
	Post	0	2	60.0 ± 10.7
	Both	10	45	55.9 ± 13.4

Table 1: Subject characteristics for retrospective data.

All 74 patients with diagnosed goitre underwent a thyroidectomy or were awaiting one. The indications for surgery were symptomatic GITC, dysphagia, diagnostic hemithyroidectomy, cosmetics, or pre-emptive to GITC. For most goitre patients, pre- and post-surgery data were available, however for some only the pre- or post-surgery data were obtained. This was because some patients had not been operated yet, or one of their scans did not include the entire region of interest. Table 9 shows the characteristics for the included subjects.

3.5 METHODS

3.5.1 *Self reported symptoms*

All patients were screened for symptomatic GITC based on their medical records. These records contained patient's Self Reported Symptoms (SRS) documented by physicians during outpatient consultation. Additionally, all patients which underwent surgery filled out a questionnaire whereby they were asked whether they experienced shortness of breath during minor physical exertion. These two resources were used to quantify the reliability of SRS and were subsequently compared to Radiological Parameters (RP).

3.5.2 *Radiological parameters*

To ensure objective and reproducible results, an automated segmentation tool was used to obtain trachea dimensions, similar to the one described in chapter 2.
In brief, this tool segments the trachea lumen in transversal slices. Thereafter, these segmented slices are used to generated a 3d-model of the trachea. Next, a Centre Lumen Line (CLL) is constructed throughout this reconstructed trachea. This CLL is equidistant to the trachea wall, thereby allowing to assess trachea dimensions orthogonal to the mean direction of airflow, which is assumed to be more relevant in terms of fluid mechanics compared to basic transversal dimensions.
This tool is used to monitor trachea dimensions throughout its course. First, these obtained dimensions are used to define healthy anatomical variation. Next, obtained dimensions are used to quantify several

RP in patients and controls, these include the CSA_m, %-TC and MVC. This analysis serves two purposes, on the one hand the analysis of control subjects yields reference values for healthy subjects and on the other hand the effect of GITC on these parameters is assessed.

3.5.2.1 *Diameter*

It is well-known that trachea cross-sections are not perfect circles, but shaped like horse-shoes [93]. Due to this irregular shape, the magnitude of the diameter depends on the manner in which it is defined. In order to realise non-ambiguous and reproducible results, a circular equivalent diameter (d_e) was used, similar to the methods of chapter 2. This d_e is a function of cross-sectional area and perimeter, as defined by equation 21 [76]:

$$d_e = 1.55 \frac{CSA^{\frac{5}{8}}}{P^{\frac{1}{4}}} \tag{21}$$

Whereby CSA is the cross-sectional area in mm^2 and P the perimeter in mm. CSA and P are straight-forward indices, independent from the orientation of the trachea cross-section (just as long as it is orthogonal to the CLL). Note that from here and onwards the terms *circular equivalent diameter* and *diameter* are used interchangeably.

3.5.3 *The MVC*

Trachea diameters obtained using the automated segmentation were used to assess MVC values for all included subjects. Whereby the sum of diameters was used to calculate maximal achievable trachea airflow, i. e. the peak airflow. The MVC was defined as the calculated peak airflow in the diseased state divided by the calculated peak flow in the healthy state, as described previously in chapter 2. It reflects the subjects potential ventilation capacity of the trachea and is only affected by changes in trachea diameter. It is independent of the presence of cardiac or pulmonary disease, and this is the key to assess the limitations in daily life activities exclusively caused by trachea narrowing such as occurs with compressive goitres. Calculation of the MVC requires that a (reliable) prediction can be made of a subject's normal peak flow. Calculation of the predicted normal peak flow is based on estimated healthy trachea dimensions. To predict these healthy trachea dimensions, are reference point is used.
In healthy human subjects, the vocal cords will impose a trachea constriction (reduction in cross sectional lumen area) of approximately 40% [68]. Caudal from this constriction, the trachea will widen. The first peak in diameter caudal from the vocal cords is therefore used

as a reference (d_{ref}). The hyoid bone lies caudal from the vocal cords, this rigid structure encases the trachea and is assumed to prevent compression of the trachea due to GITC. Because this point is not affected by GITC, it can serve as a useful reference for healthy trachea dimensions, i. e. the dimensions for that specific individuals prior to the onset of GITC.
After defining d_{ref} the predicted normal peak flow is calculated on the assumption that the trachea is straight tube with a constant diameter equal to d_{ref}.

3.5.4 *Validation of the MVC model*

Figure 17 is an example of trachea diameter measurement in a control subject. It illustrates that the trachea is not a straight tube, there is always some degree of anatomical variation, such as imposed by an widened aortic arch, aberrant blood vessels and thickened oesophageal wall. The trachea diameters measured distal of the point of reference may be larger or smaller than d_{ref} and this could affect the accuracy of the MVC calculation procedure. If d_{max} represents the maximal diameter measured after d_{ref} and d_{min} represents the smallest diameter after d_{ref}, the normal variation around d_{ref} can be expressed by $\Delta_{max} = d_{max} - d_{ref}$ and $\Delta_{min} = d_{min} - d_{ref}$. Values of Δ_{max} and Δ_{min} were obtained in 55 controls to assess the 95% confidence interval of the variation around d_{ref}, i. e. Δd_{ref}. Subsequently the relation between Δd_{ref} and d_{ref} was examined. In addition to this normal variability around d_{ref}, it is conceivable that d_{ref} in the presence of a compressive goitre may not be the same as d_{ref} after removal of that goitre. Displacement of the trachea by a unilateral goitre may have a stretching effect that tends to reduce the trachea diameter, not only at the site of displacement but also in the area that is considered to be normal [94]. Thus, d_{ref} measured before surgery might be smaller than d_{ref} assessed after removal of the goitre, at least in theory, and this could affect the accuracy of the prediction of patients normal trachea dimensions. The magnitude of this potential error was assessed by comparing pre- and post-surgery d_{ref}.

Additionally, d_c will be obtained. d_c is defined as the trachea diameter 2 cm above the carina. As mentioned before, this point is commonly used in literature as a reference point, for healthy trachea dimensions (similar to d_{ref}). However, in this study, d_c was intentionally not chosen as a reference point, since it was assumed that extraordinary large goitres might compress trachea dimensions even near the carina. Furthermore, this d_c is not located near bony structures that might prevent trachea compression due to such goitres. Moreover, it is hypothesized that anatomical variation, such as widening

of the aortic branch, might cause compression of the trachea at these levels, thereby further distorting the prediction of healthy trachea dimensions.

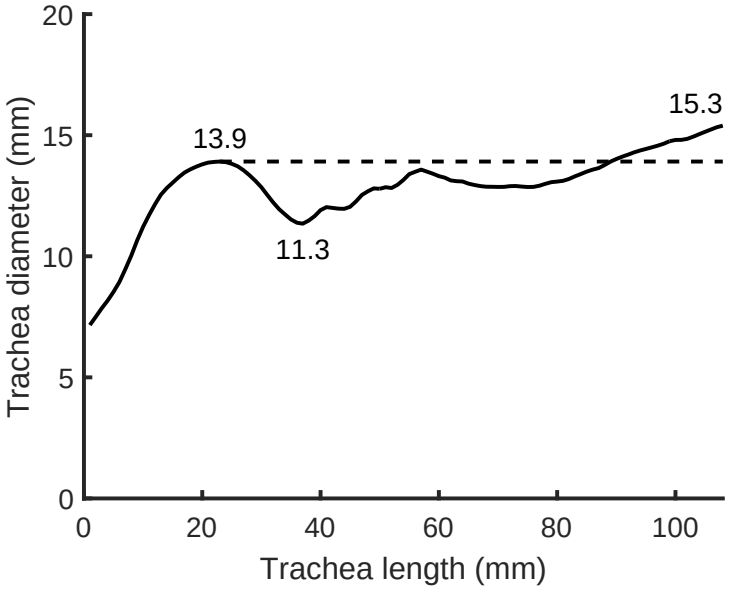


Figure 17: This figure illustrates the trachea diameter of a 55 year old female, throughout the course of her trachea (solid line). Analysis starts at the site of the hyoid bone and caudal from the vocal chords. The vocal chords usually impose a constriction of 40%, which explains the rise in diameter at the beginning of the graph. The first peak (d_{ref} = 13.9 mm) is used as a reference point and extrapolated over the course of the rest of the trachea (interrupted line). Whereby d_{min}(11.3 mm) is the site of maximal trachea constriction and d_{max}(15.3) the site of maximal trachea widening. Additionally d_c is observed, this the trachea diameter 2 cm above the carina. In this particular patient d_c is equal to d_{max}.

3.5.4.1 *Trachea compression and CSA_m*

The MVC is a novel measure which has not been analysed extensively yet. To date, studies have used other indices, such as relative trachea compression (%-TC) and Minimal Cross Sectional Area (CSA_m) to monitor the effects of GITC. To verify how these measures relate to the MVC and other measures such as SRS, they were also assessed for all included subjects.
Hereby, relative trachea constriction (%-TC) was defined as a percentage in cross-sectional diameter reduction, according to equation 22:

$$\%\text{-TC} = \frac{d_{ref} - d_{min}}{d_{ref}} \times 100\% \tag{22}$$

Whereby d_{min} represents the reference diameter and d_{min} the minimal diameter caudal from this reference point, similar to section 3.5.4.

3.5.5 *Statistics*

Results are shown as mean values ± standard error of the mean (SE), unless stated otherwise the values of the control data were tested for normality using the one-sample Kolmogorov-Smirnov test. Subsequently reference values were determined using $\mu \pm 2\sigma$ for normally distributed data or the middle 95%-percentile for non-normal data. Using these reference values a lower of limit of normal was established, subjects with MVCs below this limit were considered pathological.
In all instances a significance-level of 5% used.

3.6 RESULTS

3.6.1 *Self reported symptoms*

Data on reported symptoms during outpatient consultation were incomplete in 5 of 72 pre-surgery patients, for the pre-surgery questionnaire data was incomplete for 11 of 72 patients.
14 patients indicated a sensation of dyspnea in both instances, and 23 patients indicated an absence of dyspnea in both instances, for 14 patients dyspnea was indicated during consultation both not with the questionnaire and in 11 patients an absence of dyspnea was indicated during consultation but not with the questionnaire. This yielded a very weak correlation ($r = 0.18$, $p = 0.16$).

3.6.2 *Maximal Ventilation capacity*

MVC's assessed in patients before and after surgery are shown in figure 18, compared to controls. The mean MVC was 89.2 % in controls (normal range 58.1↔ 120.3%). It was similar for female and male control (89.6±15.0 versus 88.7±16.5, P = 0.8385). The overall mean pre-surgery MVC increased from 70.6 ± 22.7 to 93.3 ± 23.7 and was statistically different ($p = 1.31 \times 10^{-6}$).

3.6.3 *Trachea diameter*

CONTROLS

The control group had a mean d_{ref} of 15.5 ± 2.7 mm (range 9.7 ↔ 21.2). Multivariate analysis indicated that 62% of the variability in trachea diameter was explained by gender ($P = 2.00 \times 10^{-21}$). Stature and weight had no independent effects. Mean d_{ref} was 13.6 ± 1.5 mm in women (normal range 9.7 ↔ 16.8 mm) and 18.0 ± 1.6 mm in men, (normal range 15.3 ↔ 21.2 mm), and the difference between men and women was statistically significant ($p = 1.23 \times 10^{-14}$).
Δ_{min} and Δ_{max} of all 55 control subjects are illustrated in figure 19

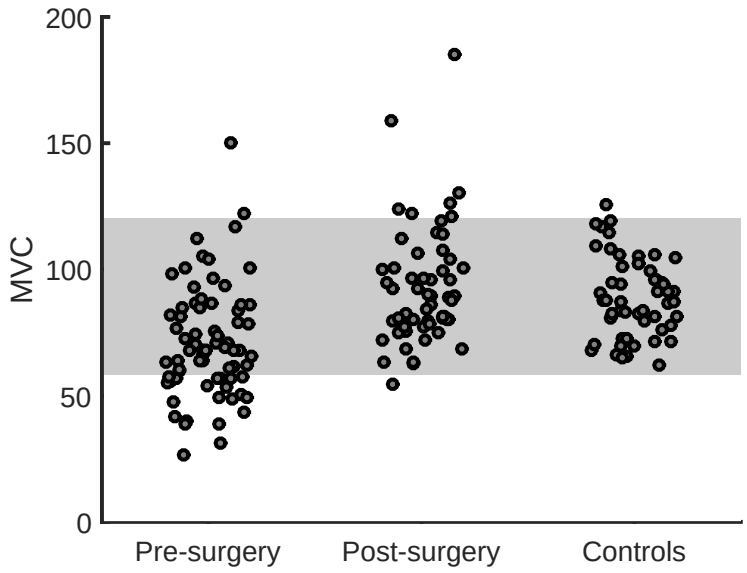


Figure 18: MVC-values for pre- and post-surgery and control subjects. The grey area represents reference values obtained using the control subjects.

Based on these data the 95%CI of the normal range of fluctuation around the reference line was +3.06 to -4.52.

PATIENTS

To examine the accuracy of the pre-surgery d_{ref} as predictor of the trachea diameter as it would have been if the patient had not developed a goitre, the pre-surgery d_{ref} was compared to the post-surgery d_{ref}. Valid reference points should not change after intervention, as well as between patients and controls.

The first condition was examined by comparing pre- and post-surgery d_{ref}. The correlation between pre- and post-surgery d_{ref} was very high ($r = 0.91$, $p = 8.01 \times 10^{-22}$). Nevertheless there was a significant difference between both parameters which was 0.3 mm on average (95% CI -1.2 $\leftrightarrow$ 1.8, p = 0.05). The difference between pre- and post-surgery d_{ref} can be simply corrected by adding 0.3 mm to obtain the post-surgery d_{ref}. It was decided not to perform this correction because it represented less than 5% of d_{ref} and thus considered to be not clinically relevant. Diameter variations of this magnitude affect the calculation of the MVC by 6%.

The second condition was examined by comparing the d_{ref} of patients and controls. However, the amount of women in the post-surgery group (82.5%) was significantly larger than in the control group (56.4%). To correct for this gender-inequality, a multivariate analysis was performed. Hereby the effect of the assigned group (post-surgery or control) as well as gender, on d_{ref} was assessed. This indicated that the

assigned group did not significantly affect d_ref ($p = 0.52525$). Additionally d_c was as for all patients. Mean d_c values were 13.33 ± 2.47 mm prior to surgery and increased to 13.95 ± 2.55 after surgery. Correlation between pre- and post-surgery values was moderate ($r = 0.62$, $p = 5.77 \times 10^{-22}$).

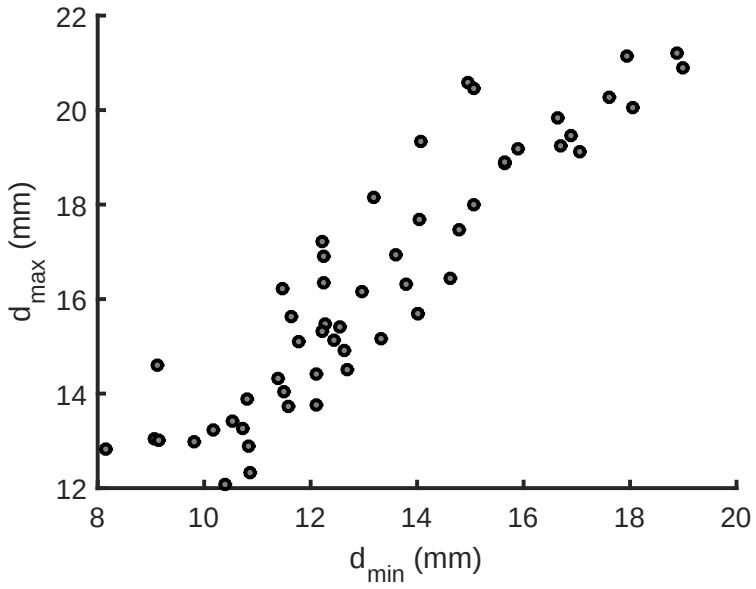


Figure 19: Δ_{min} and Δ_{max} for controls (n = 55).

3.6.4 *Trachea CSA*

CONTROLS

CSA_m assessed in patients before and after surgery are shown in figure 20b, compared to controls. Mean CSA_m was 147.00 ± 58.96 mm^2. Multivariate analysis revealed that gender greatly effected magnitude of CSA_m ($R^2 = 0.533$, $p = 2.49 \times 10^{-10}$), whereas parameters such as weight, stature and age did not.
In men, mean CSA_m was 195.49 ± 50.82 mm^2 (lower limit 93.85 mm^2) and in women this was 109.46 ± 30.66 mm^2 (lower limit 48.14 mm^2).

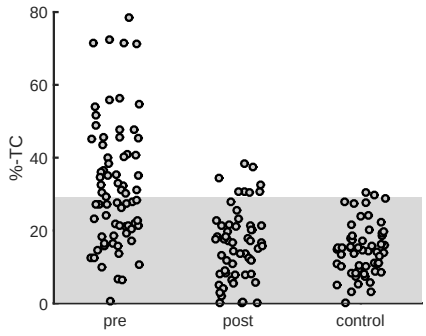
PATIENTS

Prior to surgery mean CSA_m was 78.87 ± 42.94 mm^2 (range $6.41 \leftrightarrow 243.12$), after surgery this increased to 123.07 ± 47.38 (range $42.53 \leftrightarrow 268.30$). Prior to surgery, gender did not affect CSA_m that much ($R^2 = 0.20$, $p = 0.000104$), but after surgery this increased slightly ($R^2 = 0.27$, $p = 2.91 \times 10^{-5}$.
Pre-surgery CSA_m-values were not a good predictor for post-surgery CSA_m-values due to a poor correlation ($r = 0.33$, $p = 0.015$).
Prior to surgery, mean CSA_m was 69.21 ± 31.43 mm^2 (range $6.41 \leftrightarrow 131.31$) in women and 115.57 ± 59.82 mm^2 (range $16.33 \leftrightarrow 243.1233$) in men, after surgery this increased to 111.73 ± 35.61 mm^2 (range

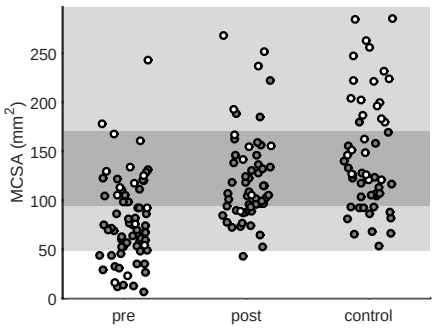
42.53 $\leftrightarrow$ 222.18) in women and 176.39 $\pm$ 60.48 mm^2 (range 89.21 $\leftrightarrow$ 268.30) and in men.
Gender-corrected post-surgery values were not statistically different from control values ($p = 0.66$).

3.6.5 *Trachea Constriction*

% trachea narrowing was obtained for all subjects, based on the smallest diameter detected anywhere distal of the reference diameter. It is an objective assessment of abnormalities in structure, not of function. TC-values assessed in patients before and after surgery are shown in figure 20a, compared to controls. Normal variation in trachea diameter relative to d_{ref} was associated with a mean percentage of constriction of 14.7 $\pm$ 7.2% and an upper limit of normal of 29.1%. Visual inspection of the data by an expert physician revealed that these observed constrictions had a wide variaty of causes such as widening of the aortic branch, thickening of the esophageal wall and other anatomical variations. Using 30 as cut-off, constriction was found to be excessive in 37 out of 72 patients (51%) scheduled for surgery.
In controls as well as goitre-patients, %-TC-values were not significantly effected by gender, stature, weight and age. The mean constriction percentage decreased after surgery from 32.6 to 16.0 % in patients but was still above the upper limit of normal for 8 patients.



(a) %-TC values for pre- and post-surgery and control subjects. The grey area represents reference values obtained using the control subjects.



(b) CSA_m values for pre- and post-surgery and control subjects. The upper grey area includes normal-values for men but not for women, the lower grey area includes normal-values for women but not for men and the middle (darker) grey area includes normal-values for both gender. White encoulered dots represent male and grey encoulered dots represents female data.

Figure 20: Pre- post and control data for TC and CSA_m

3.6.6 *Comparing indices*

Pre-surgery values for MVC, %-TC and CSA_m are illustrated in figure 21. Comparison of the MVC and %-TC in patients before surgery is illustrated in figure 21a and 21b. All patients that were below the lower limit of normal for the MVC were also above the upper limit of normal for %-TC (n = 22). Whilst this was not always the case the other way around, 15 patients fell in the normal range for MVC but not for %-TC.
A similar pattern can be observed when comparing female CSA_m and %-TC values. All these values that were pathological based on their reference values were also pathological based on their %-TC values (n = 13), however this was not necessarily the case the other way around, 18 females patients were pathological based on their %-TC but not their CSA_m values
Additionally, this pattern was less clear cut in CSA_m values for men. For both 5 measures, 5 males patients fell outside the normal ranges, however only 2 of these patients pathological according to both measures.
For %-TC, 8 post-surgery patients fell outside the normal range. These numbers were lower for MVC (n = 1)[3] and CSA_m (n = 2).

3.6.6.1 *Sensation of dyspnea*

Figure 22a and 22b show MVC and %-TC values for patients who did and did not experience dyspnea based on the outpatient consultation. Evidently, these figure do not show a distinct difference in patients with and without dyspnea. Similar patterns were observer during the assessment of CSA_m and %-TC, these figures are therefore not displayed.
Correlation between reported sensation of dyspnea during outpatient consultation and %-TC ($r = 0.22$, $p = 0.0849$), CSA_m ($r = -0.21$, $p = 0.086$) and MVC (r = -0.1018, p = 0.4196) were rather weak as well as reported sensation of dyspnea in the pre-surgery questionnaire and %-TC ($r = 0.28$, $p = 0.029$), CSA_m ($r = -0.17$, $p = 0.19$) and MVC ($r = -0.25$, $p = 0.048$).
Taking apart CSA_m-values for men and women did not significantly improve these correlations for reported sensation of dyspnea during outpatient consultation ($r = -0.43$, $p = 0.13$ for men and $r = -0.22$, $p = 0.13$ for women) as well as pre-surgery questionnaire ($r = -0.19$, $p = 0.54$ for men and $r = -0.28$, $p = 0.053$ for women).
An additional analysis was performed whereby all patients whereby both instances of SRS did not agree were disregarded. This yielded two groups, one whereby both the out-patient consultation as well as the pre-surgery questionnaire indicated a sensation of dyspnea (n = 14) and another whereby both these instances reported an absence of

3 Disregarding patients above the upper limit of normal

dyspea (n = 23). Both these groups were analysed using the CSA_m, %-TC and MVC. Differences between these groups were not significant for CSA_m ($r = -0.32$,$p = 0.057$) and were significant for %-TC, MVC ($r = 0.50$,$p = 0.0019$ and $r = -0.38$, $p = 0.021$ respectively). Analysing male and female values for CSA_m separately did not change these results.

3.7 DISCUSSION

3.7.1 *Prediction of healthy dimensions*

GITC does reduce d_{ref} values. Nevertheless, the strong correlation between pre- and post-surgery d_{ref} indicates that post-surgery dimensions can accurately be predicted using pre-surgery values. These post-surgery values for d_{ref} are most probably similar to healthy trachea dimensions. Therefore the correlation between pre- and post-surgery data can be used to more accurately predict healthy trachea dimensions.
Correcting for this effect yields slightly different MVC-values. Using these values only 20 instead of 22 pre-surgery patients fell outside of the normal range. This could mean that correcting the MVC has a positive effect on sensitivity, i. e. , the percentage of healthy individuals who are correctly identified as not having GITC has been increased. However, this difference could also be explained by a decrease in specificity, i. e. , the percentage of people with GITC who are correctly identified with GITC has decreased.. Currently available data are not sufficient to yield a conclusive explanation for this phenomenon, therefore MVC will not be revised based on these findings just yet.

3.7.2 *Radiological parameters*

The MVC and %-TC show a strong negative correlation ($r = -0.76$ and $\rho = 4.46 \times 10^{-36}$). Nevertheless, using the %-TC 37 of the 72 pre-surgery patients could be identified with GITC, whilst this was only 22 for MVC. This could mean that sensitivity of MVC is inferior to %-TC, or that specificity of MVC is superior to %-TC.
Figure 21 illustrates the strong correlations between the assessed radiological parameters (CSA_m, MVC and %-TC). Whereas figure 21a implies that the correlation between MVC and TC is non-linear, but rather exponential. Plotting the same data, with MVC on a logarithmic axis supports this theory (figure 21b). This findings can be explained using the mathematical theory behind the MVC. As section

2.3 mentions, the relationship between flow (Φ) and dimensions (d) can be described by equation 23:

$$\Phi \propto \sqrt{R} \times d^{\frac{5}{2}} \tag{23}$$

Which implies an exponential relationship between Φ and d at equal airway resistance (R). Explaining the curved trend between MVC and %-TC values at linear axis.
Figure 21a and 21b illustrate another phenomenon. The lower contour of the data cloud of these figure is sharply defined, whereas the upper portion shows more dispersion. Additionally, this upper portion of the data contains some outliers (patients with a high %-TC as well as a high MVC). Trachea dimensions of one of these outliers have been depicted in figure 23. The patient depicted in this figure shows some significant widening at the caudal portion of the trachea. The diameter of this caudal portion exceeds the reference line, thereby masking in the effect of the stenosis at the cranial portion of the trachea, on the MVC. In 3.2 it has been suggested that the dimensions of the stenotic portion of the trachea in patients with GITC are determinative in the onset of UAO. If this hypothesis is correct, the MVC will fail to recognise relevant trachea constrictions in a portion of assessed patients. This might explain why the MVC was only able to identify 22 patients with GTIC, whereas %-TC was able to diagnose 37 patients.
Additionally figure 21c, 21d, 21e and 21f show that only 18 patients fell below the lower limit of normal for CSA_m. Ostensibly, this might contravene with the hypothesis that absolute RP are more relevant for the prediction of UAO than relative ones. However, this is not the case per se, as intra-individual correlation for anatomical are quite strong (e. g. the correlation between d_{min} and d_{max} $r = 0.91$, $p = 2.66 \times 10^{-22}$), indices such as d_{ref} can serve as a useful reference to determine whether other indices such as d_{min} fall outside of the realm of healthy anatomical variation. However, this does not necessarily mean that such cases will actually experience ventilatory impairment.
For example, consider an individual named *patient A* and another named *patient B*. Patient A has a d_{ref} of 18 mm and his d_{min} is 10 mm. This corresponds to a %-TC of $\frac{18-10}{18} = 44.4$ %, which is above the upper limit of normal (29.1%). Let's assume that based on patient A's ventilatory requirements, he will only experience ventilatory impairment when his d_{min} falls below 8 mm. Patient B ventilatory requirements are similar to those of patient A, however his d_{ref} is 12 mm. Similar to patient A, patient B develops a trachea stenosis with a TC of 44.4 %. Thus his d_{min} is 6.7 mm, and limits his physical capabilities.
None of the assessed parameters (age, stature, weight and gender) could be linked to changes d_{ref}, %-TC, MVC or CSA_m as a result of

surgery. This implies that the effectiveness of surgery is not affected by these parameters, and therefore does not have to be accounted for when planning intervention.
The MVC, %-TC and CSA_m generated significantly different results for diseased and control patients as well as pre- and post-surgery data, whereas post-surgery and control data do not differ. This indicates that trachea's affected by GITC are indeed different from healthy trachea's. Furthermore, post-surgery trachea's are not statistically different from healthy trachea's, however post-surgery %-TC-values were above the upper-limit of normal for 8 patients. It should be assessed whether these patients did not fully recover, or whether the aberrant values are caused by non-goitre related anatomical phenomena.
Additionally this could yield insights in the findings of the pilot study (chapter 2), this study reported that trachea dimensions did not improve with surgery for 2 out of 10 patients. However, comparing their MVCs to observed normal-range values indicates that it is plausible that these trachea's were not pathological to begin with.
As shown in section , GITC does reduce d_{ref} values. Nevertheless, the strong correlation between pre- and post-surgery d_{ref} indicates that post-surgery dimensions can accurately be predicted using pre-surgery values. These post-surgery values for d_{ref} are most probably similar to healthy trachea dimensions. Therefore the correlation between pre- and post-surgery data can be used to more accurately predict healthy trachea dimensions.
A big overlap between pre-surgery and controls, in terms of %-TC, CSA_m and MVC, was observed. This is most probably because the pre-surgery group also contained subjects whom did not suffer from GITC (or non relevant levels of GITC) and were operated for cosmetic reasons, pre-emptive to GITC or symptoms such as dysphagia and globus pharyngis. Despite this overlap in pre-surgery and control patients, still some relevant differences can be observed. Reference values for CSA_m, %-TC and the MVC have been defined. Several pre-surgery patient fell outside of this normal range, indicating that at least a portion of patients suffering from GITC can be identified using these RP.

3.7.3 *Self-Reported Symptoms*

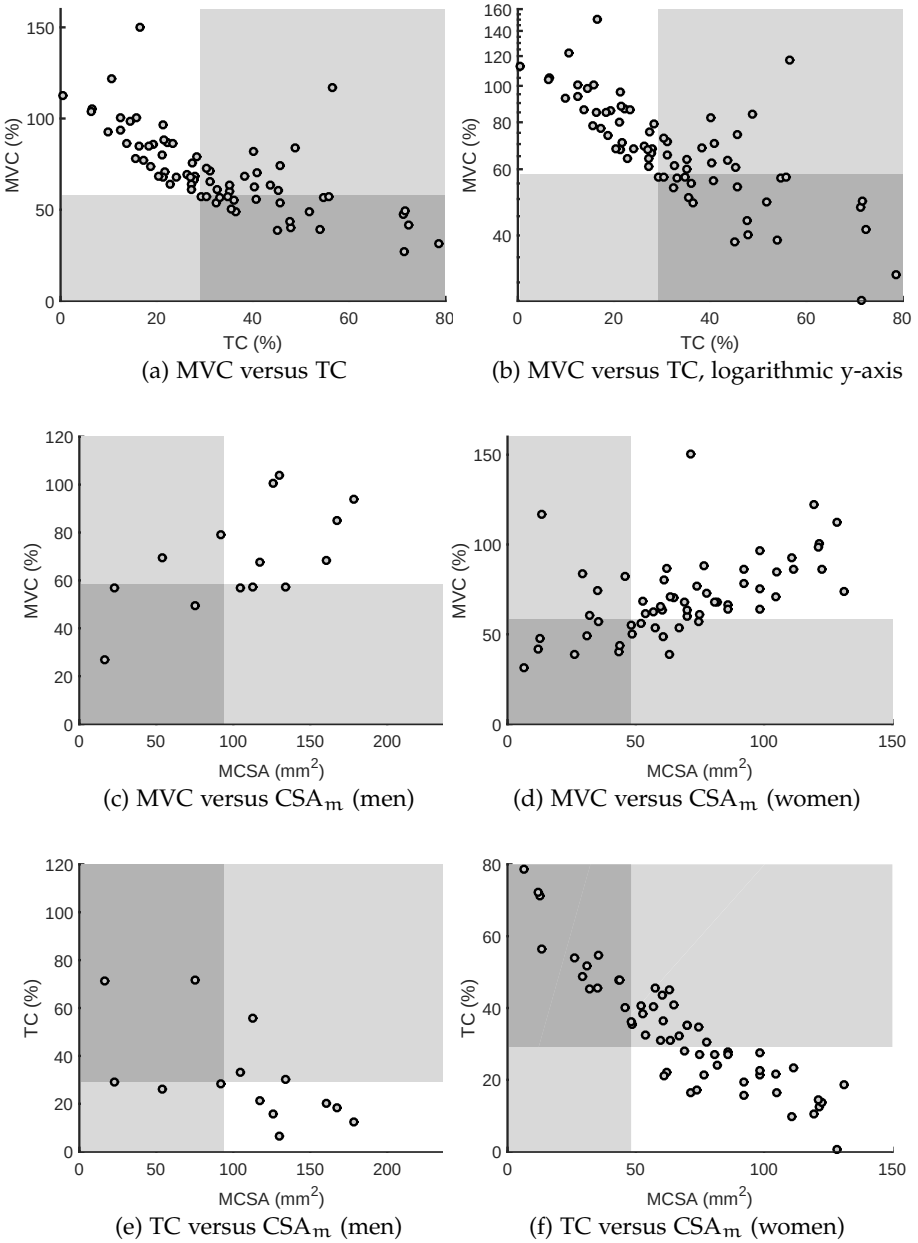
Documented symptoms from the questionnaire and consultation showed poor correlation. Due to the retrospective nature of this study, it is difficult to verify in which manner these data were collected. Moreover, definitions of dyspnea were vague. During outpatient consultation all mentions of shortness of breath were documented as dyspnea, whereas the pre-surgery questionnaire only inquired about a short-

ness of breath during minor physical exertion. Additionally, the timeframe wherein both databases were consulted differed vastly among patients. And finally, concomitant diseases were not monitored. This makes the reliability of the data questionable.
In addition to this, both sources of Self Reported Symptoms (SRS) showed poor correlation with other RP. These findings are contradicted by literature. The earlier mentioned study by Stang *et al.* assessed the relation between %-TC and Positional Dypnea (PD) in 1081 patients and found a strong correlation [24]. Hereby, PD was defined by trouble breathing that could be reduced by an alteration of posture. It is possible that this criterion weeds out non-GITC related causes of dyspnea and therefore is a better predictor for GITC. However, a study by Shin *et al.* in 200 patients suffering from goitre reported good correlation between a general sensation of dyspnea and %-TC ($p < 0.001$) [21].
These observations further support the theory that quality of documentation of the SRS were suboptimal and could be improved upon. Nevertheless, still some statements about the relation between dyspnea and RP could be made.
Disregarding patients whereby reported sensation of dyspnea did not agree for both databases (out-patient consultation and pre-surgery questionnaire), improved correlation between SRS and RP. Hereby, it is assumed that disregarding patients who were not consistent in their SRS weeds out the less obvious (i. e. , more subtle) cases of dyspnea and improves reliability of the data.
In view of these findings, it seems that %-TC is the best predictor for ventilatory impairment due to GITC. Naturally this raises the question why only Bonnema *et al.* were able to find a significant relation between RP and pulmonary function. In view of the findings of this study it could be stated that the methods adapted in literature show room for improvement. For instance, d_{ref} seems a better reference point for healthy trachea dimensions than d_c (correlation between pre- and post-surgery values for d_{ref} were $r = 0.91$, $p = 8.01 \times 10^{-22}$ and $r = 0.62$, $p = 5.77 \times 10^{-22}$ for d_c). Additionally, quantifying the exact degree of trachea constriction rather than its mere presence will most probably further increase the significance of documented findings. Additionally, different RP (such as the CSA_m, MVC and %-TC) should be assessed simultaneously and compared to other measures in a consistent fashion. This further stresses the need for prospective study which implements these recommendations.

3.8 CONCLUSION

This retrospective study has been able to identify which values for CSA_m, %-TC and MVC fall in the realm of normal anatomical variation. It seems that %-TC might be the best predictor for ventilatory

impariment due to GITC.
However a prospective study is necessary to establish this conclusion more firmly. Recommendations on how to perform such a study have been included.



(a) MVC versus TC

(b) MVC versus TC, logarithmic y-axis

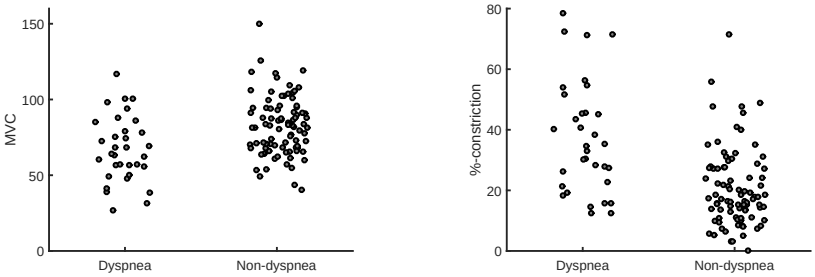
(c) MVC versus CSA_m (men)

(d) MVC versus CSA_m (women)

(e) TC versus CSA_m (men)

(f) TC versus CSA_m (women)

Figure 21: Comparative plots for TC, MVC and CSA_m for pre-surgery patients. Grey area's indicate pathological values (based on control data), in darker grey area patients lie outside the normal values for both indices.



(a) MVCs for patients with and without dyspnea assessed during out-patient consultation

(b) %-TC for patients with and without dyspnea assessed trough out-patient consultations

Figure 22

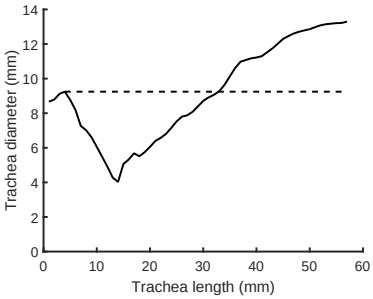


Figure 23: Trachea diameter of 71 year old female, MVC = 116.84 and %-TC = 56.41.

4

PROSPECTIVE STUDY: PATIENT X

Goitre is a palpable enlargement of the thyroid which might cause Upper Airway Obstruction (UAO) as a result of compression of the trachea. Several therapies for management of goitre are available, including surgery, I^{131}-therapy and radio frequency ablation. Each of these therapies have different efficacies [4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14]. To decide on the most suitable therapy, UAO should be assessed accurately.
Commonly used methods for the assessment of Goitre Induced Trachea Compression (GITC) include the assessment of Pulmonary Function Tests (PFT), Self Reported Symptoms (SRS) and Radiological Parameters (RP). The significance of these methods for the assessment of UAO has been assessed in several studies. Nevertheless, as described by chapter 3, methodologies for these study differ, and show room for improvement.
Literature points out that commonly used methods for the assessment of UAO, such as SRS, RI and PFT, often show poor correlation (with the exception of some isolated cases: see chapter 3). Although, some promising new methods are available, these have not been extensively researched. Furthermore, the application of commonly used methods could be improved upon. In view of these findings, the current study was initiated, which aims to research potential improvement of commonly used methods (PFT, RP and SRS) and evaluate the potential of novel measures.

4.1 BACKGROUND

In this section, a wide range of available measures is discussed. Results documented in previous chapters, as well as findings reported in literature are used to identify which of these measures are potentially useful for the assessment of GITC and how they should be interpreted.

4.1.1 *Self reported symptoms*

Generally SRS show poor correlation with other measures in terms of UAO [34, 30, 88]. Nevertheless, some studies, such as the one carried out by Stang *et al.* reported a strong relationship between self reported sensation of positional dyspnea and trachea constriction [24]. A possible explanation is that positional dyspnea might be more specific for

UAO than a general sensation of dyspnea.
In chapter 3, intra-patient-variability for SRS was found to be large. Because this was a retrospective study, it is difficult to verify the exact manner in which these data were documented. It is a well-known fact that the mode of administration can effect data quality, thus the mode of administration should be critically monitored and applied in a consequent manner [95]. This indicates the need for a well-organised and structured mode of symptom monitoring.
Usually SRS are assessed by either an interview or questionnaire. Generally a questionnaire will be more objective and structured, because the information gathered by an interview is physician-dependent. Harris *et al.* reviewed 19 questionnaire-interview comparison studies and found poor levels of agreement [96]. Both methods have different efficacies with respect to the type of data that have to be gathered. Generally questionnaires are more suited to explore a well defined area where most variables are known, whereas an interview is preferred to explore complex topics where several relevant variables have not been defined.
In view of the problems regarding the assessment of UAO, whereby possible symptoms are well-known and a structured approach is necessary, a questionnaire will be the most suitable method.
Extensive literature study showed that the Quality of Life Questionnaire for Patients with Thyroid Disease, also known as ThyPRO is probably the currently most relevant questionnaire regarding thyroid disease. ThyPRO has shown to have good responsiveness and ability to detect treatment effects [97]. Translated versions of ThyPRO have also been studied, and results showed only minor invariance compared to its original version. Therefore, this questionnaire is also recommended for use in international multi centre studies [98]. Nevertheless, extensive consultation with several endocrinologists, revealed that ThyPro contains a lot of irrelevant questions and does not cover all relevant symptoms linked to ventilation impairment in goitre. Although it covers a lot of common goitre symptoms, the assessment of dyspnea seems rather concise. Nevertheless, there are questionnaires which offer a more extensive investigation of dyspnea, such as the St. George Respiratory Questionnaire (SGRQ) which has proven to be viable for assessment of dyspnea symptoms due to several diseases[99, 100, 101]. Although these questionnaires give a detailed overview of the magnitude of shortness of breath and its effect on the patients quality of life, they lack questions relevant to goitre symptoms, which makes them not suitable to differentiate between ventilation impairment due to goitre or other pathologies such as chronic obstructive pulmonary disease. Devising a new, more relevant questionnaire based on the ThyPRO and SGRQ seems therefore a better solution.

4.1.2 *Pulmonary function tests*

PFT are the current golden standard for assessment of ventilaotry problems due to UAO [28, 102]. PFT include whole-body-plethysmography and spirometry. Spirometry is able to record Flow-Volume Loops (FVL) as well as Pulmonary Function Values (PFV). These PFV are predefined indices obtained from FVL. The plethysmograph is used to record airway resistance of the entire airways. As a result of this, plethysmography is not specific for UAO, but may be confounded by distal airway obstruction.
An alternative to plethysmography is Impulse Oscillometry Spirometry (IOS). IOS is a Forced Oscillation Technique (FOT). These techniques determine airway mechanics by imposing pressure fluctuations at varying frequencies on the airways during a subjects normal, tidal breathing [103, 104]. The absorption-coefficient of these waves determine how far they are able to penetrate the respiratory. Due to the frequency-dependence of this absorption-coefficient, IOS can be used to monitor the respiratory impedance at different levels of the airways.

4.1.2.1 *Flow-volume loops*

Spirometry is the simplest method for diagnosis of UAO [30, 31, 32]. The evaluation of Flow-Volume Loops (FVL) by an expert has been shown to be the most sensitive and specific test in PFT in determining UAO [31].
Usually upper airway obstructions such as GITC will hinder inspiratory flow prior to significant changes in expiratory flow, because inspiration is completely effort-dependent whereas expiration is only partially effort-dependent [37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47]. Additionally the negative pressure gradient generated during inspiration can cause the trachea to compress even further, whereas the positive pressure gradient generated by expiration prevents this collapse. As a result of this, UAO will cause inspiratory slowing, therefore it can be identified by a plateau on the inspiratory limb of the flow volume loop, see figure 24 [33].
However, the quality of such loops are dependent on the patient's effort and cooperation, whereby evaluation and interpretation of these loops are based on visual inspection thereby making it a rather subjective method [33, 34, 35, 36].
Apart from FVL, PFT is able to identify a wide variety of objective measures, such as Peak Expiraotry Volume (PEF), Forced Vital Capacity (FVC) and Peak Inspiratory Flow (PIF). Since UAO mainly hinders inspiratory flow, inspiraotry measures are expected to be a better predictor for UAO.
The magnitude of inspiratory flow in PFT is usually described by the

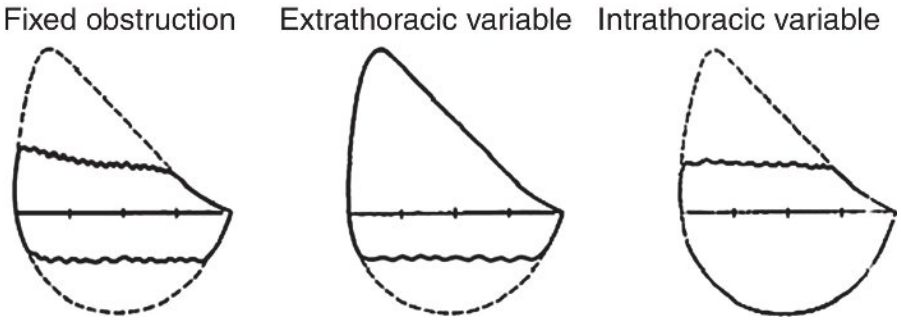


Figure 24: Interpretation of flow-volume curves [33]

PIF, Forced Inspitatory Flow at 50% of the vital capacity ($FIF_{50\%}$) and the Forced Inspiratory Volume in 1 second (FIV_1). Ideally, the measures would be interpreted using reference values. Unfortunately, no universal applicable model for the prediction of these values is available, apart from some incidental cases on small populations [45, 48, 49].

4.1.2.2 *Impulse Oscillometry Spirometry*

Impulse Oscillometry Spirometry (IOS) is an alternative to traditional spirometry. Its advantage over traditional spirometry is the absence of effort dependency contrary to traditional spirometry [50] Additionally, an enhanced sensitivity over spirometry for detecting airflow obstruction has been observed [51]. Although IOS does not require patient cooperation, its results still depends greatly on the investigator's level of familiarity with the device [52, 53, 54, 55, 56, 57, 58, 59, 60]. Generally, assessment of airway resistance using IOS shows similar repeatability compared to traditional assessment of airway resistance using plethysmography. However, correlation with spirometry varies greatly [51].
IOS uses pressure waves generated at different frequencies to measure respiratory impedance from which reactance (X), resistance (R) and fundamental frequency (f_0) can be derived [51].

REACTANCE

X includes the mass-inertive forces of the moving air column in the conducting airways and the elastic properties of lung periphery [103]. Because X characterises the lung periphery it is probably not relevant for diagnosis of UAO, this is supported by findings in literature [105]. Therefore, X will not be assessed in this study.

FUNDAMENTAL FREQUENCY

In 2001 Horan *et al.* assessed the correlation between stenotic trachea diameter and f_0 in 43 patients suffering from trachea stenosis and

found a strong correlation (r = 0.80 ↔ 0.91), see figure 25. This implies that f_0 might be a good index to monitor GITC, however f_0 also increases in restrictive airway diseases and is thus not specific for UAO [106, 107, 108].

The upper limit of normal for f_0 can be obtained using equation 24

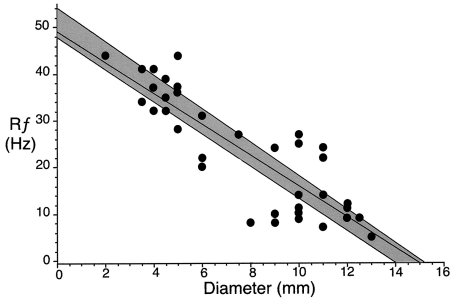


Figure 25: Relation between f_0 and trachea diameter [105]

[109]:

$$f_0 = 10.1428 + 0.1772A - 0.01201H \tag{24}$$

Whereby f_0 is the resonance frequency in Hz, A is the patients age in years and H is the patients height in cm.

RESISTANCE

R measured at lower frequencies (5 Hz) gives an indication of total airway resistance, whereas higher frequencies (20 Hz) indicates the resistance of the larger airways [110]. Thus in individuals with central airway obstruction, R will increase for lower (5 Hz) as well as higher frequencies (20 Hz), whilst distal airway obstruction will only raise R at lower frequencies [32, 111, 53, 112, 113]. In obstructions caused by trachea stenosis, R increases at these lower frequencies with decreasing tracheal luminal area, regardless of shape and length of the stenotic area[32].

Since R_{20} represents the resistance for the large airway, this parameter is regarded relevant in for the assessment of GITC. The upper limit of normal for R_{20} can be obtained using equation 25 [109]:

$$R_{20} = 1.1824 - 0.00036A - 0.00486H \tag{25}$$

Whereby R_{20} is the resistance in kPa s L^{-1} measured at 20 Hz, A is the patients age in years and H is the patients H in cm.

However, resistance corrected flow ($\frac{\Delta R}{\Delta \dot{V}}$) may be a more promising parameter than traditional R [32]. In 2009, Verbanck *et al.* carried out a study comparing 10 healthy individuals and 10 patients with tracheal stenosis. During the assessment of R, patients fell within the range of normal values, whereas this was not the case for obtained $\frac{\Delta R}{\Delta \dot{V}}$-values. Moreover, after treatment of the stenotic areas $\frac{\Delta R}{\Delta \dot{V}}$ values

returned to normal (upper limit of normal was 0.2 kPa $L^{-2}s^2$). Additionally Verbanck concluded that $\frac{\Delta R}{\Delta V}$ correlated better to CSA_m than any other spirometric UAO index.

4.1.3 *Radiological Parameters*

Various Radiological Parameters (RP) are available to define the extent of GITC. These RP either define Trachea Constriction (TC) or Trachea Deviation (TD). However, several studies have indicated that TC and not TD is most probably determinative in the onset of UAO (see chapter 1 for a more elaborate motivation of this statement) [36, 21]. Chapter 3 gave an overview of commonly assessed RP for quantification of TC. Thereafter, a selection of these RP were used to assess trachea dimensions in patients suffering from goitre as well as controls, whereby trachea dimensions were obtained using an automated segmentation procedure which was introduced in chapter 2. This procedure uses a radiodensity threshold to segment the trachea lumen in transversal CT-slices. Thereafter, these segmented slices are used to reconstruct a 3d-model of the trachea. Next, a Centre Lumen Line (CLL) is calculated throughout the course of the trachea. This CLL represents the morphological skeleton of the trachea and is assumed to be parallel to the direction of mean airflow. Calculating trachea dimensions perpendicular to this CLL (CLL-method) will yield different dimensions from assessment in the transversal-plane (transversal-method) as a result of the parallax-effect. This parallax-effect describes the overestimation of cross-sectional dimensions by cutting structures obliquely to their longitudinal axis (as in the transversal-method), instead of orthogonal to this axis (CLL-method), an illustration of this phenomenon is provided in figure 26.

Commonly assessed RP include the Minimal Cross Sectional Area (CSA_m) and Percentage Trachea Constriction (%-TC). Hereby the CSA_m is defined as the minimal observed Cross Sectional Area (CSA) of the trachea lumen caudal from constriction imposed by the vocal cords. %-TC describes the percentage of diameter reduction by comparing the minimal diameter (d_{min}) to a patient-specific reference point. In healthy individuals, the vocal cords will impose a prediction of approximately 40%[68]. Caudal from this constriction the trachea will usually widen. The vocal cords are located near the hyoid bone, which is a rigid bony structure encasing the trachea. It is assumed that this hyoid bone will prevent trachea compression in the case of externally applied forces such as GITC. Therefore it is assumed that this initial widening of the trachea caudal from the vocal cords will not be effected by GITC and thus can serve as a reliable reference point for healthy trachea dimensions based on the strong intra-individual correlation of trachea dimensions. The diameter at this reference point

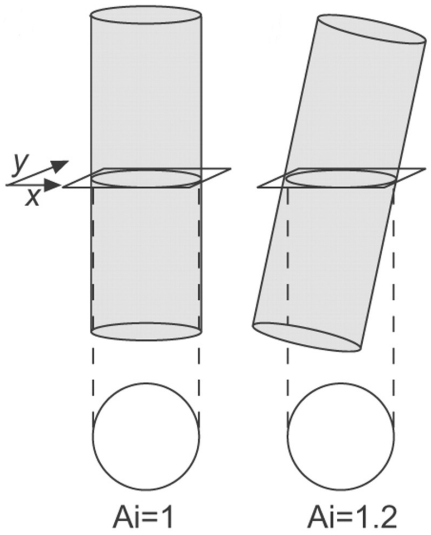


Figure 26: Illustration of the parallax-effect

is defined as d_{ref}. The ratio between d_{min} and d_{ref} is used to define %-TC according to equation 26:

$$\%\text{-TC} = \frac{d_{ref} - d_{min}}{d_{ref}} \times 100\% \tag{26}$$

Although these RP generally are highly reproducible and non-ambiguous, they are difficult to interpret, e. g. what are the practical consequences of a %-TC of 60%? A novel measures which aims to solve this issue is the Maximal Ventilation Capacity (MVC). With this MVC, CT-derived anatomical data are translated into functional data that express the Maximal Ventilation Capacity (MVC) of the trachea. The MVC of the trachea in a patient with GITC ($MVC_{patient}$) is then expressed as the percentage of this individual's predicted normal MVC. This individual specific normal values is based on trachea dimensions observed in unaffected trachea areas. Thus a MVC of 40% indicates that a patients current ventilation capacity is 40% of his potential healthy ventilation capacity, thereby allowing physicians and patients to interpret the imposed consequences of GITC (see chapter 2 for a more elaborate motivation of this MVC).
The MVC was introduced in a pilot study which evaluated 10 patients suffering from goitre and 10 controls (chapter 2). Thereafter, a follow-up study was performed whereby the MVC was evaluated in a larger sample size (chapter 3). This study accomplished to define the range for MVC-values in healthy subjects, nevertheless a large overlap between control and goitre patients was observed. This was most probably caused by the inclusion of goitre patients with insignificant levels of GITC. However, due to the retrospective nature of the study, it was difficult to assess to what degree these patients actually suffered from UAO. These results indicated the need for a prospec-

tive study, to determine the significance of the MVC for GITC more firmly.
All of these three RP (CSA_m, %-TC and MVC) are potentially useful for the assessment of GITC. To our knowledge, the MVC is the only RP with defined practical consequences . Whereas, the CSA_m is currently the only RP which has been successfully linked to a PFV. This correlation was found in 1999 by Bonnema *et al.* as is described by equation 27:

$$FIF_{50\%} = 0.036\, CSA_m - 0.67 \tag{27}$$

Hereby the $FIF_{50\%}$ is the patients forced inspiratory flow at 50% of his vital capacity in L s^{-1} and CSA_m is provided in mm^2.
Additionally, %-TC was the only RP assessed in chapter 3 that could succesfully be linked to reported sensations of dyspnea in chapter 3. Additionally, the %-TC was able to identify a larger portion of patients with GITC than the MVC and CSA_m.

4.2 IMPLEMENTATION

As explained in section 4.1, several methods are readily available to assess GITC. However, some of these can be discarded in advance due to the lack of potential added value for the assessment of GITC. Nevertheless, a substantial list of potential useful indices remains. This study aims to test the applicability and relevance of these indices, thereby providing recommendations for the methods of additional studies.
Because this study solely aims to explore the potential applicability and relevance of several methods for the assessment of GITC, only one patient is included. This patient will be analysed in detail and provide recommendations for the analysis of additional subjects.

4.3 METHODS AND SUBJECTS

The study protocol was approved by the Nijmegen-Arnhem medical-ethical committee.

4.3.1 *Patient characteristics*

Our patient, hereafter referred to as patient X, is a 69 year old male with a height of 179 cm, weighing 82 kg, with a large retrosternal goitre. He underwent a total thyroidectomy. Surgery was successful without any complications. The surgeon reported a goitre weight of 355 g. Three months before and three months after the surgery a non-contrast CT-scan of the neck region was obtained using a Philips Brilliance 40 slice CT scanner (Best, The Netherlands).

Substance	Pre-surgery	Post-surgery	Reference values	Units
FT_4	14	15	11.5 ↔ 22.7	pmol L^{-1}
TSH	0.59	2.3	0.35 ↔ 5.5	mU L^{-1}

Table 2: Laboratory results for patient X and reference values [116, 117]

4.3.2 *Concomitant pathologies*

It is well-known that goitre is linked to thyroid dysfunctions such as hypo- and hyperthyroidism [114]. In clinical practice, thyroid function tests are used to identify possible hypo- or hyperthyroidisms. Here, hyperthyroidism is indicated by suppressed Thyroid Stimiluating Hormone (TSH) levels and high or high normal thyroxine (FT_4) [115]. Literature indicates that hyperthyroidism is linked to cardiovascular and respiratory diseases such as heart failure, pulmonary hypertension and ventilatory muscle weakness, thereby causing dyspnea [62, 63, 64, 65]. As shown in table 2, both TSH and FT_4 fall within normal ranges for both hormones. This makes it unlikely that potential sensations of dyspnea are caused by hyperthyroidism induced co-morbidities.
The pre-surgery CT-scan was evaluated by an experienced radiologist. No pathologies besides his goitre, which could affect his trachea dimensions were identified. Furthermore, the patients medical history did not reveal any diseases that could significantly reduce his physical capabilities.

4.3.2.1 *Self Reported Symptoms*

As mentioned in section 4.1.1, the most suitable method for an effective assessment of SRS is most probably a questionnaire. Whereby a combination of ThyPRO and SGRQ seems the most sensible option.
In view of these findings a new questionnaire was devised. In consultation with an expert physicians, relevant passages of the ThyPRO and SGRQ were extracted and used to construct a new questionnaire. Additionally some potential relevant questions for the assessment of GITC, which were not mentioned in ThyPRO and SGRQ were added. An example of this new questionnaire has been included in the appendix (page 133). For completeness also an English version has been included (page i), nevertheless, for this research only the Dutch version was used.
The patient completed the questionnaire with the assistance of two researchers. Whereby the researchers orally presented the questions to the patient, and marked his answers. Both these researchers assisted in the establishment of the questionnaire, and were thus aware of its

contents, whereas the patients had no prior knowledge of the questionnaire.
The questionnaire contained 14 items in total. Two of these items focussed on general characteristics, namely gender and age. Five items focussed on dyspnea in general. Another five were aimed at specific symptoms such as dysphagia, hoarseness, coughing, wheezing and globus pharyngis. Although these five symptoms are not exclusively linked to goitre, they might give an indication of severity of GITC apart from dyspnea. These 14 items were used to score the goitre on four different aspects on a scale of 0-100%:

A. Functional consequences

B. Emotional impact

C. Symptoms

D. Dysnpea sensation

4.3.2.2 *Functional consequences*

The functional consequences of the goitre were used to define to what extent the goitre limits the patients physical capabilities. The patient was asked to indicate whether he has experienced eight different situations with varying degrees of intensity. Such as 'Because of my breathing problems I take longer to wash myself or to get dressed' up unto 'Because of my breathing problems I experience troubles when performing heavy psychical labour, running, riding a bike, swimming at a high pace or performing intensive sports'. Each of these situations yielded a additional score of 12.5%.

4.3.2.3 *Emotional impact*

The emotional impact of the goitre was used to verify the manner in which the GITC affects the patients' quality of life. This emotional impact was assessed in a manner quite similar to the functional consequences. The patient was presented with eight situations from which he indicated which applied to him. These situations covered several possible emotional consequences such as 'I am ashamed of my breathing' or 'My breathing makes me weak and helpless'. Each of these situations yielded a additional score of 12.5%.

4.3.2.4 *Symptoms*

Five different symptoms were assessed, namely dysphagia, hoarseness, coughing, wheezing and globus pharyngis. For each of these symptoms the patient indicated to which effect he experienced these symptoms. For instance:
I experience wheezing:

- several times a day (100%)
- several times a week (67%)
- several times a month (33%)
- never (0%)

This yielded a sub-score for all symptoms. The score for this aspect was obtained by averaging all sub-scores.

4.3.2.5 *Dysnpea sensation*

The dyspnea sensation was scored using the BORG-score. Here the patient was asked to rate his dyspnea on a scale from 0-10, with 0 representing no sense of dyspnea at all and 10 is the worst level of dyspnea the patient can imagine.

4.3.3 *Pulmonary Function Tests*

Pulmonary Function Tests (PFT) included the application of traditional spirometry as well as Impulse Oscillatory Spirometry (IOS). Spirometry was performed on a MasterScreen Body Plethysmograph (Jaeger, Carefusion Australia) and IOS measurements were done using IOS Impulse Spirometry (Jaeger, Carefusion Australia). Data were analysed using Labmanager v5.31 (Carefusion Australia).
Using the spirometry, Flow-Volume Loops (FVL) as well as Pulmonary Function Values (PFV) were obtained. IOS was used to monitor R_5, R_{20}, f_0 and $\frac{\Delta R}{\Delta \dot{V}}$.

4.3.4 *Radiological Assessment*

A CT-scan of the upper thorax was obtained using a Philips Brilliance 40 slice CT scanner, hereby a radiation exposure of 2 mG was administered. This scan consisted of 2 mm transversal slices. The earlier automated segmentation method introduced in chapter 2 was used to obtain trachea dimensions.
In brief, first the trachea lumen was segmented in transversal CT-slices based on a radiodensity threshold. Thereafter, these slices were used to reconstruct a 3d-model of the trachea. Next a Centre Lumen Line (CLL) was constructed throughout the trachea lumen. This CLL was basically the morphological skeleton of the lumen and equidistant to the trachea wall. It was assumed that obtaining trachea dimensions orthogonal to this line yields more relevant results compared to traditional assessment in the transversal plane. This is because the CLL follows the mean direction of airflow, dimensions orthogonal to this air-flow are assumed to be more relevant in terms of fluid

mechanics. Additionally, this method circumvents the parallax-effect, which may distort the assessment of trachea dimensions (figure 26). The region of interest for segmentation was defined as the first slice caudal from the vocal chords up to the first slice prior to the trachea bifurcation.
After these trachea dimensions have been quantified, the three RP discussed in section 4.1 (%-TC, CSA_m and MVC) were assessed. Their methodologies are also described in section 4.1.

4.4 RESULTS

4.4.1 *Self reported symptoms*

4.4.1.1 *Pre-surgery*

GENERAL

The patient is a 68 year old male. He rates his general health as good (4 on a 5 point scale, whereby a higher score indicates a better level of health). He is not familiar with any cardiovascular or pulmonary diseases.

DYSPNEA

The patients reported that his dyspnea started 2 years ago, whereas other goitre symptoms started 8 years ago. He reports that the severity of the dyspnea is increasing. He experiences dyspnea during washing and getting dressed, when walking, climbing up a flight of starts and more intensive forms of exercise such as tennis swimming, bowling and golf.
The emotional consequences of his dyspnea are limited to panic during episodes of dyspnea and a lack of control during breathing problems. The patients indicates a BORG-score of 8 on a scale from 0-10. Whereby 0 represents a complete lack of dyspnea and 10 the most severe form imaginable.

SPECIFIC SYMPTOMS

The patients reports coughing 0-5 times a day, he never has trouble swallowing, a feeling of tightness in the throat is always present, since the last 6 months he experiences a constant form of wheezing and he never experiences hoarseness.

4.4.1.2 *Post-surgery*

GENERAL

The patient indicated that his general health has not improved after surgery (4 on a 5 point scale). He was still not familiar with any cardiovascular or pulmonary diseases.

DYSPNEA

The patient reported that his dyspnea is completely gone, even during intense physical exertion. The patient indicates a BORG-score of 0 on a scale from 0-10.

SPECIFIC SYMPTOMS

The patients reports coughing 0-5 times a day, he never has trouble swallowing and he never experiences tightness in the throat, wheezing or hoarseness.

Figure 27 show the scoring of patient X in the four aspects proposed in section 4.1.1.

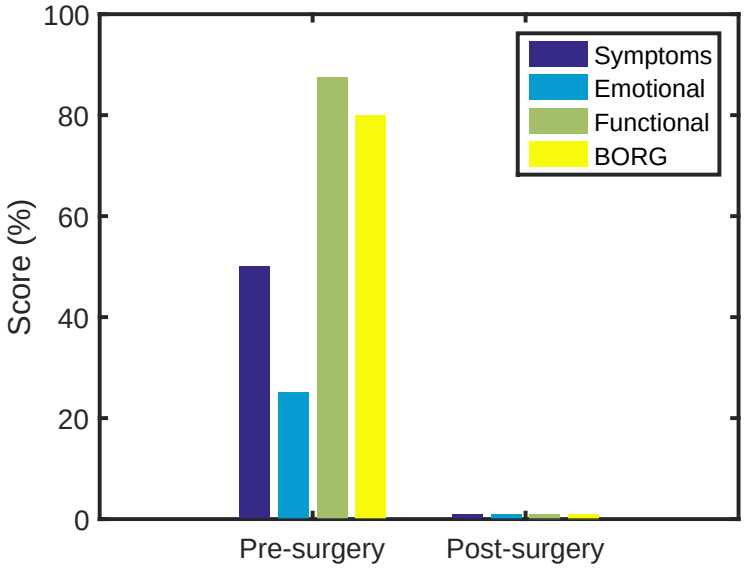


Figure 27: Questionnaire results for patient X.

4.4.2 *Pulmonary function testing*

4.4.2.1 *Traditional Spirometry*

FLOW VOLUME LOOPS

The report of the FVL indicates a obviously flattened inspiratory curve for the pre-surgery situation of patient X, whereas this flattened effect is absent in the post-surgery FVL (see figure 28a and 28b respectively).

PULMONARY FUNCTION VALUES

As table 3 shows, none of the expiratory PFV fell below their lower limit of normal, except for PEF. After surgery, all expiratory PFV were within their normal range. Table 14 of the appendix shows that mea-

		Pre-surgery		Post-surgery	
Variable	LLN	Measure	%-pred	Measure	%-pred
VC (L)	3.17	4.77	116.5	5.68	138.9
FVC_{in} (L)	3.17	4.05	99	4.85	118.5
FVC_{out} (L)	2.94	3.91	99.1	5.49	139.3
PEF (L s^{-1})	5.96	**3.86**	48.7	8.33	105
FEV_1 (L)	2.21	2.93	96.3	3.93	129.1
$\frac{FEV_1}{FVC}$	0.63	0.75	NA	0.72	NA

Table 3: Expiratory pulmonary function values for patient X. %-pred means percentage of predicted results based on weight, height, age and gender. Fields marked with NA indicate that reference values were not available.

Measure	Pre	Post	Pre/Post
PIF (L s^{-1})	2.78	7.33	0.38
FIV_1 (L)	2.58	4.85	0.53
$\frac{FIV_1}{FVC_{in}}$	0.64	1.00	0.64

Table 4: Inspiratory pulmonary function values for patient X.

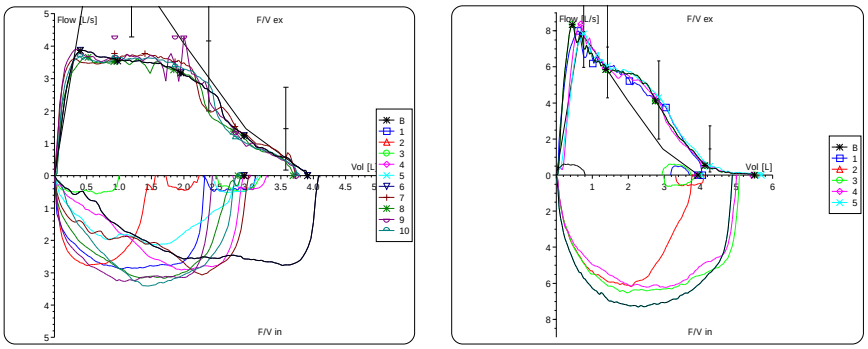
sured PEF values before surgery differed considerably ($\sigma = 0.27$ L s^{-1}), however, PEF was below this lower limit of normal consistently during all iterations.
Reference data were absent for inspiratory indices, nevertheless table 4 indicates that these PFV show a notable increase after surgery, whereby the increase in PIF was proportionally the largest.

4.4.2.2 *Impulse Oscillometry Spirometry*

Table 5 indicates that prior to surgery values for $\frac{\Delta R}{\Delta \dot{v}}$, R_5, $R_5 - R_{20}$ fell outside the realm of normal values, whereas R_{20} and f_0 did not. After surgery all IOS-indices returned to normal.

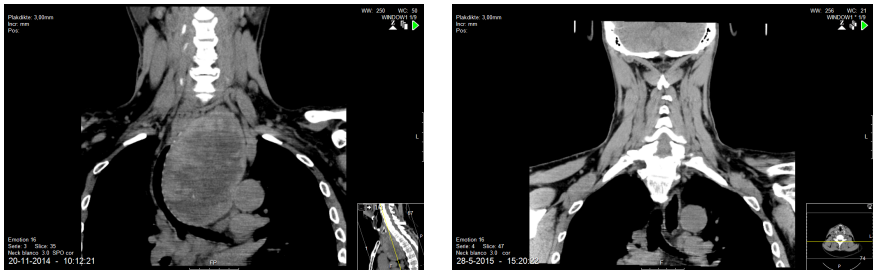
4.4.3 *Assessment of CT-data*

Figure 31 shows the diameter throughout the course of the trachea pre- and post-surgery for patient X. And figure 32a shows a 3d-reconstruction of the patients trachea anatomy. Table 6 shows the several radiological indices obtained during automated segmentation. All indices prior to surgery except for CSA_r were outside their respec-



(a) Pre-surgery (b) Post-surgery

Figure 28: Flow-volume loops for patient X



(a) Pre-surgery (b) Post-surgery

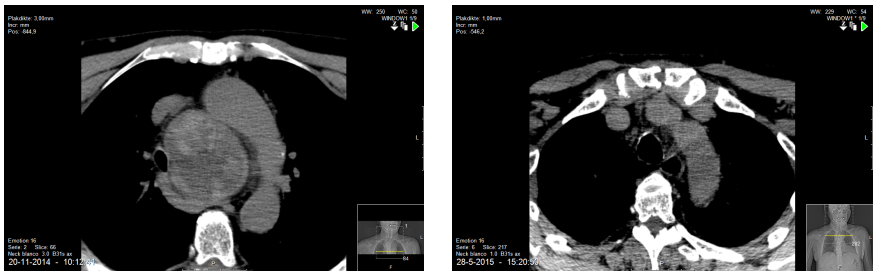
Figure 29: Coronal cross-sections of patient X

tive normal range. After surgery all values returned to normal, except %-TC, which was still slightly higher than the upper limit of normal.

4.5 DISCUSSION

4.5.1 *Self reported symptoms*

It is evident that patient X experienced a major relief in symptomatic, emotional and functional consequences of his goitre after interven-



(a) Pre-surgery (b) Post-surgery

Figure 30: Axial cross-sections of patient X

Variable	Pre	Post	ULN
$\Delta R/\Delta \dot{V}$ I	0.30	0.13	0.20
$\Delta R/\Delta \dot{V}$ E	-0.17	-0.12	NA
R_5 S ($\frac{kPa}{Ls^{-1}}$)	0.43	0.28	0.34
R_5 L ($\frac{kPa}{Ls^{-1}}$)	0.46	0.31	NA
R_{20} S ($\frac{kPa}{Ls^{-1}}$)	0.28	0.24	0.28
R_{20} L ($\frac{kPa}{Ls^{-1}}$)	0.42	0.25	NA
$R_5 - R_{20}$ S ($\frac{kPa}{Ls^{-1}}$)	0.15	0.04	0.13
$R_5 - R_{20}$ L ($\frac{kPa}{Ls^{-1}}$)	0.04	0.06	NA
f_0 S (Hz)	16.73	9.27	20.20
f_0 L (Hz)	9.37	11.07	NA

Table 5: IOS results for patient X. ULN indicates upper limit of normal based on equations of Schulz *et al.*, with the exception of ULN for $\frac{\Delta R}{\Delta V}$ which was defined by Verbanck *et al*[109, 32]. Fields marked with NA indicate that reference values were not available. I refers to inspiration, whereas E indicates expiration.

Parameter	Normal limit		pre-surgery	post-surgery
CSA_m (mm^2)	93.86	$\downarrow$	16.3	89.2
%-TC	29.1	$\uparrow$	71.4	34.4
MVC	58.1	$\downarrow$	26.8	81.0
$FIF^{pred}_{50\%}$ ($L\ s^{-1}$)	2.71	$\downarrow$	-0.082	2.5

Table 6: Radiological parameters for patient X. $\downarrow$ indicates a lower limit of normal (LLN), whereas $\uparrow$ indicates the upper limit of normal. The LNN of $FIF^{pred}_{50\%}$ is calculates using the LLN of CSA_m and equation 27.

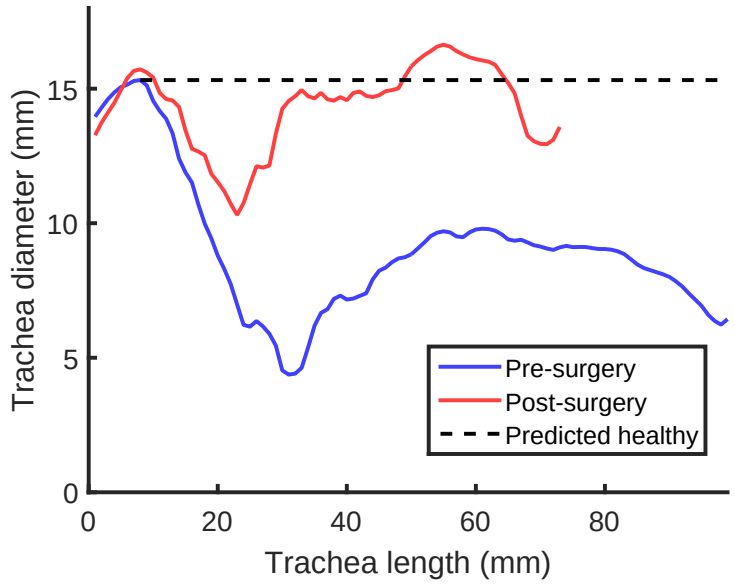
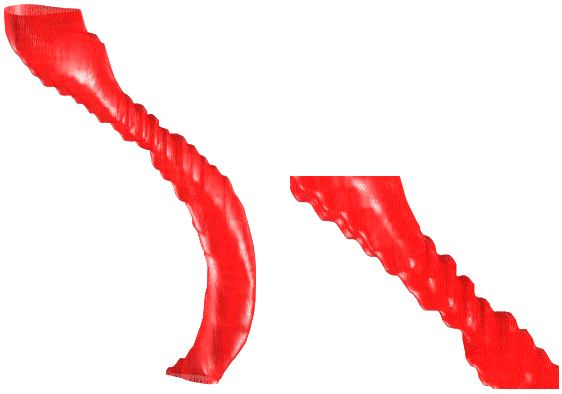


Figure 31: This figure illustrates the pre- and post-surgery trachea diameter of patient X throughout the course of his trachea (solid line). Analysis starts at the site of the hyoid bone and caudal from the vocal chords. The vocal chords usually impose a constriction of 40%, which explains the rise in diameter at the beginning of the graph. The first peak (d_{ref} = 15.3 mm) is used as a reference point and extrapolated over the course of the rest of the trachea (interrupted line). Whereby d_{min}(4.4 mm) is the site of maximal trachea constriction and d_{max}(15.3) the site of maximal trachea widening.

tion. This is the first time this questionnaire has been put to practice, which makes its interpretation difficult. Nevertheless, it was inspired by various other questionnaires which have been researched extensively and have proven to be valid tools for use in clinical practice. Nevertheless, results of multiple subjects and comparison with other validated measures are necessary to firmly assess its significance in assessment of GITC.

It is remarkable that the patient reported a notable functional limitation as a result of his disease (score = 87.5%) whilst the impact of his symptoms is considerably milder (score = 50%). As explained earlier, the interpretation of symptoms can be difficult for a patient. Moreover, the way in which the disease expresses itself differs per patient. This further supports the theory that the assessment SRS may provide a distorted image of the magnitude of the UAO.

Furthermore, the emotional consequences experienced by the patient seem fairly mild. The patient also indicated that he waited fairly long before he sought medical attention for his goitre. This illustrates that the impact on the quality of life is not solely dependent on its functional consequences, but also depends on the patients individual ambitions.



(a) A 3d-reconstruction of the trachea anatomy for patient X

(b) Enlarged view of the 3d-reconstruction of the patients trachea at the site of maximal constriction.

Figure 32

4.5.2 *Pulmonary function testing*

4.5.2.1 *Flow-Volume Loop*

The pre-surgery FVL clearly show a flattened inspiratory curve whereas the post-surgery FVL does not, this was also confirmed by an expert physician.
The 163.67% increase in PIF (see table 3) after surgery supports the idea that the patients inspiratory flow is hindered before intervention. Additionally, the expert physician indicated that the patients expiratory curve was also aberrant, suggesting that the patients ventilatory impairment is quite severe.

4.5.2.2 *Pulmonary Function Values*

As section 2.5.1 explains, goitre mainly hinders inspiratory flow. This might explain why the patients is still able to generate a fairly good FEV_1 at 96.3% of its predicted value. However, his PEF at 48.7 of its predicted value indicates that the patients expiratory flow is limited nonetheless.
Unfortunately, the literature regarding pulmonary tests lacks reference values for inspiratory indices. This hiatus in data severely limits the assessment of GITC through PFT-indices.

4.5.2.3 *Impulse Oscillometry Spirometry*

It is remarkable that R_5 exceeded the upper limit of normal, whereas R_{20} did not. Generally, R_5 reflects the total airway resistance, whereas R_{20} reflects the resistance for the central airways [110]. Additionally,

$R_5 - R_{20}$ is believed to represent the resistance of the distal airways. This value was also raised above its upper limit of normal, this implies that patient X distal airways are also obstructed to some extent. Figure 29a shows that the patients goitre prior to surgery was enlarged even at caudal regions of the trachea near the carina, which support the theory of distal airway obstruction.
Nevertheless, the fact that R_{20} was not increased above its upper limit of normal indicates that patient X upper airways were not severely obstructed or that R_{20} is not a sensitive measure for UAO. Since the results from other measures (SRS and RP) indicate that UAO in patient X is quite likely, it is implied that R_{20} is indeed not a sensitive measure for UAO.
After surgery, f_0 was reduced to 55.4% of its pre-surgery value. This suggests that f_0 is indeed linked to airway obstruction. Nevertheless, pre-surgery values were still below the upper limit of normal. This implies that f_0 may be an effective tool to monitor the progression of GITC, but is not suitable for an initial diagnoses.
Prior to surgery, $\frac{\Delta R}{\Delta \dot{V}}$ was raised above its upper limit of normal, whereas after surgery these values returned to normal. However, it should be noted, that this upper limit of normal was established using only 10 subjects. Although preliminary results regarding the implication of this measure are promising, its significance should be established more firmly before use in clinical practice.

4.5.3 *Assessment of CT-data*

4.5.3.1 *Assessment of dimensions*

Figure 30a and 29a clearly show that the patients trachea is constricted and deviated, this was confirmed by an experienced radiologist. Tracheal deviation is known to stretch the trachea, thereby decreasing its diameter due to the accordion-effect [94]. This might also explain the slight increase in CSA_r after intervention.

CSA_m
Table 6 shows that the CSA_m is 16.33 mm^2 prior to surgery and increases to 89.21 mm^2. In chapter 3 CSA_m values were assessed in various control subjects. These analysis indicated that for male adults, the lower limit of normal for CSA_m is 93.86 mm^2. The increase in CSA_m after surgery, suggests that the tracheal dimensions prior to surgery were indeed diminished for patient X. Additionally, intervention helped to relieve a significant portion of this compression, nevertheless post-surgery values were still pathological.
The obtained pre-surgery CSA_m is remarkable in view of the findings of Bonnema *et al.* As shown in figure 33, $CSA_m < 20$ mm^2 in-

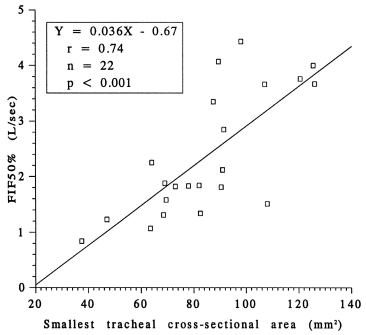


Figure 33: Minimal trachea CSA versus FIF_{50} [91]

dicates that FIF_{50} approaches 0 L s^{-1}. Since, the smallest CSA_m observed by Bonnema is 40 mm^2 is not sure whether this linear correlation also applies to CSA_m below this value. Nevertheless, even if this were not the case, one would not expect to find PIF levels of 2.78 L s^{-1}, which would correlate to a CSA_m of 94.44 mm^2 according to equation 29[91]:

$$FIF_{50} = 0.036 CSA_m - 0.67 \tag{28}$$

Measuring CSA_m in the transversal-plane can cause an overestimation of dimensions due to the parallax-effect (figure 26), especially in trachea's with high levels of TD, such as is the case with patient X [75]. Nevertheless, it is unlikely that this effect would contribute to a error with the magnitude of 5.7 times it's original value. A more plausible explanation can be obtained by zooming in on figure 32a, as is done in figure 32b. This figure illustrates that the `smooth3` (see chapter 2 section 2.2.3 for a more elaborate explanation on the implementation of this function) has some issues in generating a smooth surface in portions of the trachea with high levels of constriction and deviation, most probably due the large distances between the pixels of the trachea wall. The result of this is a wrinkled surface, masking the actual trachea dimensions.
The 3d-model of trachea (figure 32a) is constructed using a binary matrix. This matrix consists of subsequently segmentations of the trachea lumen in the transversal plane. Assessing trachea dimensions in these transversal planes (transversal-method), will prevent artefacts such as those displayed in figure 32b to distort the assessment of the tracheal geometry. Additionally, this transversal-method would be a good approximation fo the methods of Bonnema *et al.*
The results of this additional analysis are provided in figure 34. Comparing figure 31 and 34 yields several interesting observations. Firstly, the higher stenotic area in the transversal-method versus the CLL-method. In view of the before mentioned artefacts (figure 32b) and findings of Bonnema *et al.* (figure 33), it is likely that the CLL-method gave an underestimation of the actual trachea dimensions. The results

Variable	Magnitude
CSA_r (mm^2)	263.6
CSA_m (mm^2)	54.0
%-TC	79.5
MVC	27.9
$FIF_{50\%}^{pred}$ ($L\ s^{-1}$)	1.3

Table 7: Radiological parameters for patient X obtained using the transversal method.

of this additional transversal analysis are provided in table 7.

Entering the results of the transversal-method (CSA_m =) into equation 29, yields an $FIF_{50\%}$ of 1.3 L s^{-1}. Generally the $FIF_{50\%}$ is a good approximation of the PIF, whereby the PIF is usually slightly higher than the $FIF_{50\%}$ [118]. Additionally, in the post-surgery data a stenotic area of 175.48 mm^2 is found, which correlates to a $FIF_{50\%}$ of 6.99 L s^{-1} according to equation 29. Comparing these findings to those of the PFT (table 3) still reveals a notable gap between the findings of Bonnema *et al.* and the results of this study (1.51 L s^{-1}). However, visual inspection of figure 33, suggests that this discrepancy is not much greater than those of other cases in the research of Bonnema *et al* [91].

Additionally, the transversal-method shows a more than proportional case of trachea widening (figure 34), which is absent in the CLL-method (figure 31). Visual inspection of the post-surgery 3d-reconstruction does not indicate such widening. However, it shows

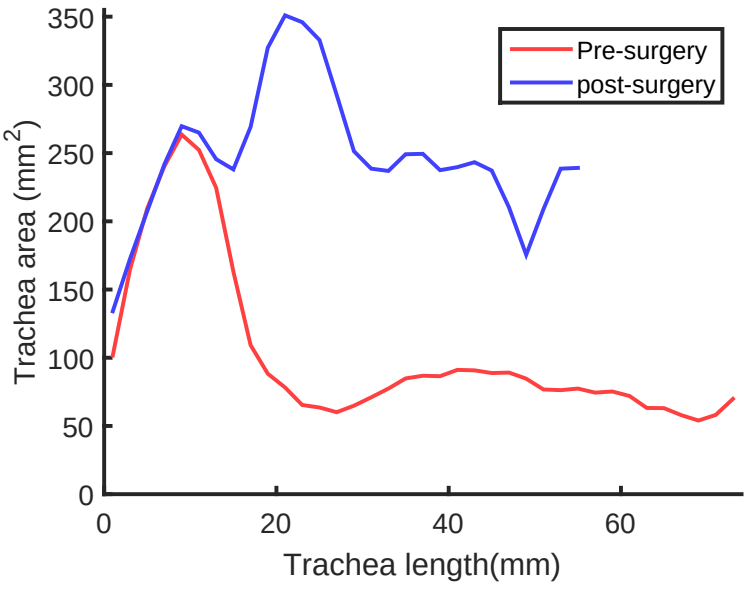


Figure 34

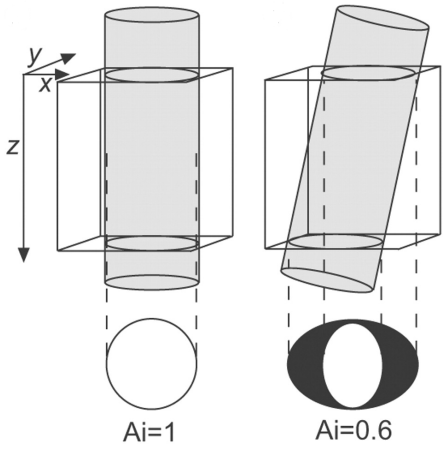


Figure 35: Illustration of the volume-averaging effect [75]

that the post-surgery trachea still contains a significant portion of curvature [1], this turtuosity might increase the error imposed by the parallax-effect (figure 26), which was the reason the CLL-method was developed in the first place. Additionally, both methods are affected by the volume-averaging effect (figure 35), which also imposes an error on the estimation of trachea dimensions. This volume average effect is caused by an off-set in the location of the trachea lumen in subsequent CT-images, this casts a shadow around the trachea lumen, causing the segmentation procedure to segment a smaller portion of the trachea lumen.
However, the phenomenon observed in figure 32a, hereafter referred to as the wrinkling-artefact, makes the accuracy of assessment of trachea dimensions using the CLL-method questionable. To yield more insights, regarding the reliability of this CLL-method, a comparative analysis between the CLL- and transversal-method was initiated and described in chapter 5.

4.5.3.2 *Comparative analysis*

Comparing obtained MVC values to those reported in chapter 3 indicates that pre-surgery MVC-values are pathological and post-surgery values are not, as is the case for %-TC. Nevertheless, interpretation of the MVC can be a bit misleading. As chapter 2 explains, a MVC of 81.04% predicts that the patients is able to reproduce 81.04% of the ventilation in a healthy state. However, as table 3 indicates, patient X is able to produce flow levels well above corresponding reference values. Additionally, chapter 2 and 3 both indicate that the MVC in

1 It is questionable whether this is pathological, visual inspection of 3d-reconstructions of the retrospective control data of chapter 3 revealed that healthy trachea's also show curvature to some expect, however the magnitude of these observations were not quantified.

healthy controls are often well below 100%.
Further, it is well known that in some case trachea widening might occur, which may generate MVC-values over 100% [119]. However, in view of the earlier mentioned overcapacity of the human trachea (section 7.0.1), it is likely that in a healthy individuals, trachea dimensions are not restrictive in generating maximal levels of ventilation[2]. Furthermore, in assessment of the MVC, widening in one portion of the trachea, might conceal the effects of constrictions elsewhere.
Additionally, the MVC is a relative geometrical measure, i. e. it uses the ratio between predicted healthy and observed diseased dimensions, to predict the consequences of the reduced airway dimensions. However, both literature (section 7.0.1) and findings of this study suggest that the absolute CSA_m is decisive in the generation of UAO. Whereas the MVC was able to differentiate between patients suffering from GITC and healthy controls for some cases in the retrospective analysis in chapter 3, since for most cases the MVC will correlate to the CSA_m. Nevertheless, it should be noted that these conclusions are solely based the findings of Bonnema *et al.* and the additional analysis of one patient (patient X). Therefore, the need for additional research explicitly stressed, to further investigate the significance of the MVC, %-TC and CSA_m for the detection of UAO.

4.6 PRELIMINARY CONCLUSIONS

Assessment of symptoms (coughing, dysphagia etc.) did not seem to be a good representation of functional consequences of GITC, but more relevant in assessment of patient discomfort.
Almost al parameters improved after surgery. Indicating that these measures are indeed effected by the presence of GITC. Nevertheless, prior to surgery some of these measures (R_{20} and f_0) still fell within the range of normal values, which makes them unfit for the management of GITC.
To further establish the relevance of other measures ($\frac{\Delta R}{\Delta \dot{V}}$, CSA_m, %-TC and MVC) for the management of GITC, a larger sample size should be evaluated.
Furthermore, it seems that assessment of trachea dimensions using the CLL-method might fail in complex geometrical structures. The effectiveness of the transversal- and CLL-method for the assessment of trachea dimensions should be researched further.

2 Thus surgical widening of the airways will most probably not benefit Olympic athletes

5 GEOMETRICAL ASSESSMENT OF CT-DATA

5.1 BACKGROUND

Goitre is a palpable enlargement of the thyroid. This enlargement may cause trachea compression, thereby increasing airway resistance and causing ventilatory impairment [3]. Assessment of Goitre Induced Trachea Compression (GITC) can be realised using various methods such as pulmonary function testing, anamnesis of self reported symptoms and Radiological Parameters (RP).
RP aim to quantify geometrical characteristics of the trachea anatomy. Generally, RP will either quantify Trachea Constriction (TC) or Trachea Deviation (TD). Nevertheless, several studies indicated that TD is probably not relevant in the onset of ventilation impairment due to goitre [21, 36].
Methodologies for the assessment of TC differ widely. Whereas some studies use a reduction of diameter, others choose to assess Cross-Sectional Area (CSA) instead. Furthermore, some studies will use the absolute magnitude of these indices, where as others chose relative ones (e. g. a percentage reduction in CSA). Additionally, some studies choose to manually these dimensions whilst other relied on automated methods.
Nevertheless, all studies documented in current literature choose to assess trachea dimensions in the transversal plane, hereafter referred to as the transversal-method. Nonetheless, it is well known that analysing geometrical complex structures such as the trachea, using the transversal-method will impose a margin of error due to the parallax-effect. This parallax basically entails the overestimation of trachea dimensions due to analysis of cross-sections obliquely cut to its morphological skeleton, see figure 39.
To eliminate such errors, a Centre Lumen Line (CLL) can be used. This CLL basically represent the morphological skeleton of the trachea lumen and is equidistant to the trachea wall. A method employing such a CLL has been introduced in chapter 2. In brief, first the trachea lumen is segmented in transversal CT-slices based on a radiodensity threshold. Thereafter these slices are used to reconstruct a 3d-model of the trachea. Next the coordinates of the CLL throughout the trachea lumen are calculated. Thereafter trachea dimensions are obtained by calculating where planes orthogonal to this CLL intersect with the trachea wall. It is assumed that obtaining trachea dimensions orthogonal to this line yields more relevant results compared to traditional assessment in the transversal plane. This is because the CLL

follows the mean direction of airflow, dimensions orthogonal to this air-flow are assumed to be more relevant in terms of fluid mechanics. However, application of this CLL-method in the analysis of a prospective patient (chapter 4) revealed a worrisome phenomenon. It seems that in patient with high levels of TD and TC, the CLL-method fails to accurately reconstruct trachea dimensions, as illustrated in figure 37b. It seems that the the `smooth3` which is used to generate a smooth 3d-model from the pixelated transversal trachea slices (see figure 38a). These artefacts caused by the smooth3 function cause the surface of the 3d-reconstruction of the trachea lumen to wrinkle at portion of the trachea with severe TD and TC. To verify whether the CLL-method is a reliable method for the assessment of trachea dimensions, it is important to verify whether the observation in chapter 4 is an isolated case, or that such artefacts occur more frequently.
The aim of this study is to assess the prevalence of the wrinkling-artefact and yield recommendations on accurate assessment of trachea dimensions.

5.2 METHOD

5.2.1 *Subjects*

CT-data of 72 patients (15 males and 57 females) suffering from goitre were collected (identical to those used in chapter 3). All patients underwent a thyroidectomy or were awaiting one. Mean age of subjects was 56.90 ± 12.84 (mean ± SE) years.

5.2.2 *Measurements*

Minimal Cross Sectional Area (CSA_m) was assessed in all subjects using the CLL- and transversal-method.
Usually the vocal cords will impose a reduction in cross sectional lumen area of approximately 40% [68]. Caudal from this constriction imposed by the vocal cords, the trachea widens. The CSA_m was defined at the site of minimal cross-sectional area caudal from this initial widening, as illustrated by figure 36.
These CSA-values were obtained using the CLL- and transversal-method. To illustrate relevant differences between these methods, they are both briefly described.
The segmentation-process preceding these analysis is similar for both methods. First a region of interest is defined from the first slice caudal from the hyoid cartilage up until the carina of the trachea using a graphical user interface. Thereafter slices within the region of interest are converted to binary images using a radiodensity threshold which is believed to differentiate between air-containing structures and tissue. Next these slices are cropped around the presumed tra-

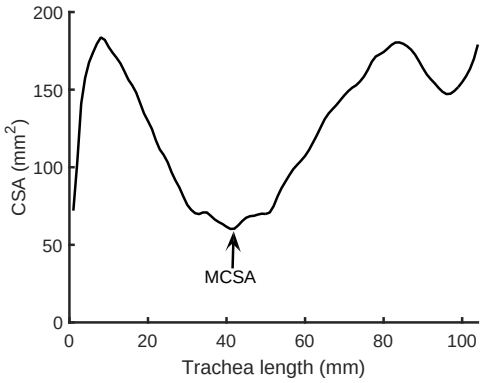
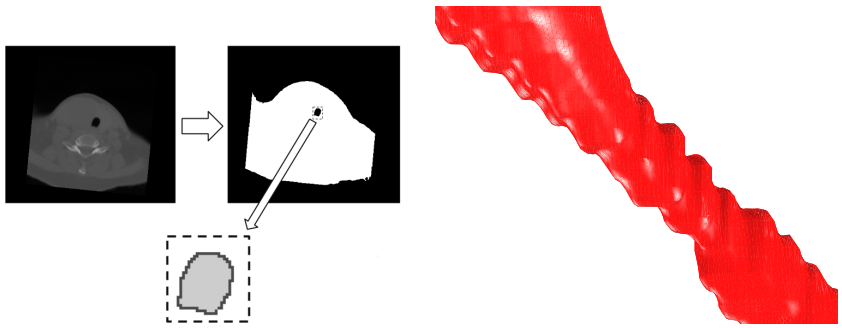


Figure 36: A 43 year old female suffering from goitre with a CSA_m of 60.24 mm^2.

chea cross-section as illustrated in figure 37a. Subsequently, this process is repeated for all slices.



(a) Segmentation process (b) Wrinkling artefact

Figure 37

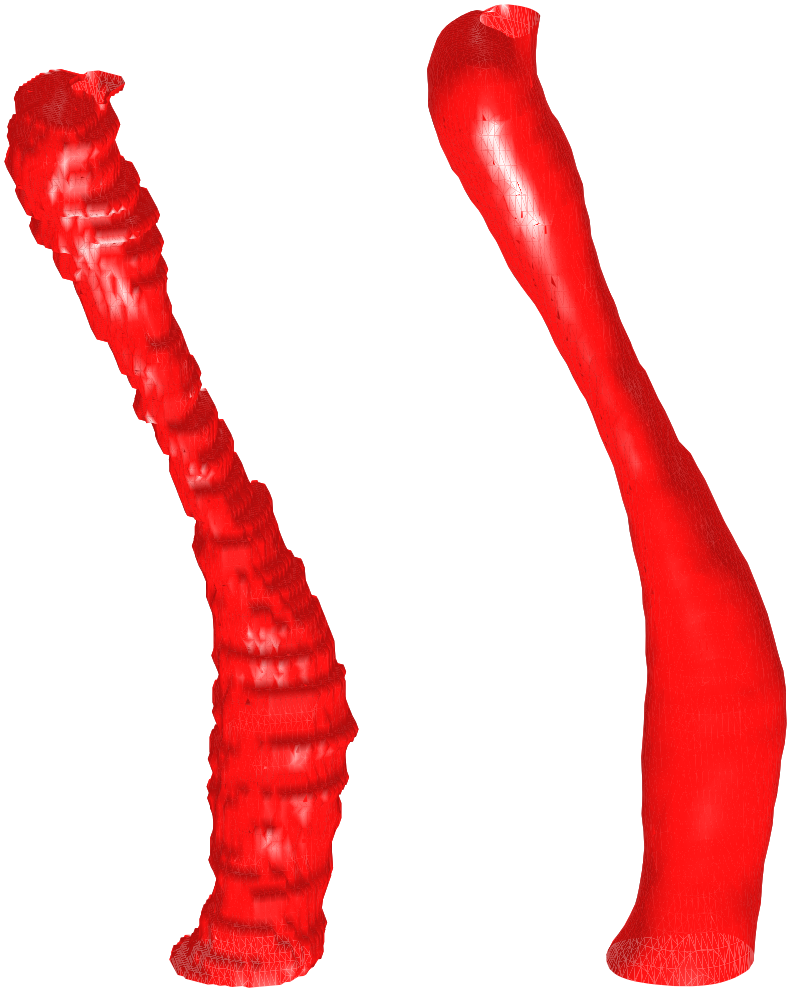
5.2.2.1 *Transversal-method*

CSA is obtained by counting the amount of segmented pixels per slice. This yields the CSA in $pixel^2$. Next the CSA in mm^2 is obtained by multiplying this CSA in $pixel^2$ with the $pixelspacing^2$ (obtained from the scan's dicom-tags). This pixelspacing indicates how pixels relate to actual dimensions, i. e. a pixelspacing of 0.5 mm $pixel^{-1}$ indicates that a pixel in a CT-image has a width and length of 0.5 mm[1].

5.2.2.2 *CLL-method*

Contrary to the transversal-method, the CLL-method uses the segmented slices to reconstruct a 3d-model of the trachea. This is done by generating an isosurface of the segmented slices in their spatial position relative to each other (figure 38a). Thereafter this isosurface is

1 The pixelspacing actually contains two arguments, defining the length and width of a pixel. In practice these are almost always identical.



(a) Isosurface of the tracheal wall before smoothing

(b) Isosurface of the tracheal wall after smoothing

Figure 38: Smoothing of the 3d-model

smoothed using matlab's `smooth3` function (figure 38b). Next, a CLL is generated and trachea dimensions are calculated orthogonal to this CLL. As figure 39 illustrates, calculating trachea dimensions using the CLL yields different results from assessment in the transversal plane.
This CLL is believed to follow the direction of the mean airflow, hereby it is assumed that dimensions orthogonal to this CLL are therefore more relevant in terms of fluid mechanics.

5.3 RESULTS

From here and on, CSA_m-values obtained using the transversal method are referred to as CSA_m^t whilst values for the CLL method are labelled CSA_m^c.

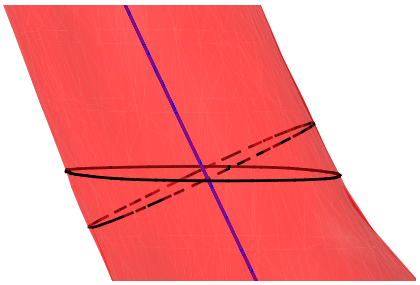


Figure 39: Obtaining trachea dimensions using the Centre Lumen Line (blue line). The interrupted black line shows where coordinates in the trachea wall intersect with a plane orthogonal to the CLL, whereas the solid black line displays the intersection in the transversal plane.

Mean CSA_m^t was 88.23 ± 42.15 mm^2 (mean $\pm$ SE) values ranged from 28.13 to 260.61 mm^2.
Mean CSA_m^c was 78.87 ± 42.94 mm^2 (range $6.41 \leftrightarrow 243.12$ mm^2.)
The mean difference between both measures (CSA_m^t - CSA_m^c) was 9.36 ± 14.91 mm^2 (range -20.02 $\leftrightarrow$ 60.91). This difference was not strongly correlated to the magnitude of CSA_m^t ($r = 0.12$, $p = 0.30$) or CSA_m^c ($r = -0.23$, $p = 0.056$).
Correlation between both methods was strong ($r = 0.94$, $p = 4.16 \times 10^{-34}$), see figure 40.
Visual inspection of the data revealed that the wrinkling-artefact was

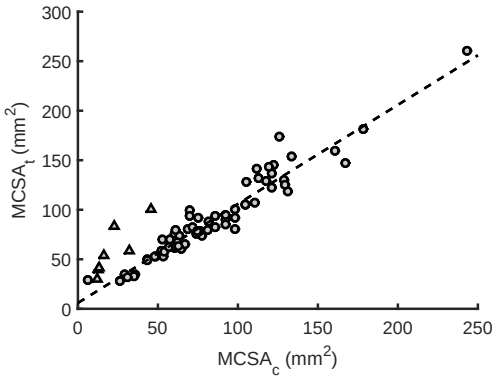


Figure 40: Relation between CSA_m^c and CSA_m^t, abberant values (indicated by triangles, n = 7) were not used in the construction of the trend-line (interrupted black line).

present in 7 patients (indicated by triangles in figure 40). Exclusion of these cases improved correlation between CSA_m^t and CSA_m^c even further ($r = 0.96$, $p = 5.12 \times 10^{-37}$).

5.4 DISCUSSION

Results indicate that the magnitude CSA_m^t and CSA_m^c are strongly correlated and that in most cases the difference between these two indices is only small (mean = 9.36 mm^2). However, in some cases this difference might even reach up to 60.91.
In 1999 Bonnema researched the relationship between CSA_m and $FIF_{50\%}$, and found a strong correlation ($p < 0.001$) [91]. $FIF_{50\%}$ is the forced inspiratory flow at 50% of the vital capacity and is strongly correlated to Peak Inspiratory Flow (PIF)[118, 120]. Based on their findings, Bonnema devised equation 29:

$$FIF_{50\%} = 0.036CSA_m^t - 0.67 \tag{29}$$

Hereby $FIF_{50\%}$ is defined in L s^{-1}. Assuming this relation between CSA_m^t and $FIF_{50\%}$, a difference of 60.91 mm^2 in CSA_m will cause a change of 2.19 L s^{-1} in $FIF_{50\%}$. Although no suitable reference values are readily available for $FIF_{50\%}$, it is not difficult to imagine that such errors will have significant consequences for a patients ventilatory capacity.
Additionally, considering the mean error between CSA_m^c and CSA_m^t of 9.36 mm^2 and equation 29, one might expect a difference of 0.34 L s^{-1} in $FIF_{50\%}$. However, due to the lack in reference values for inspiratory pulmonary indices, it uncertain whether such differences are clinically significant. Additionally, it is important to realise that these predictions are based on the observation of only one study. Therefore, it is sensible to establish the relationship between RP and pulmonary function more firmly.

Empirical screening revealed that in 7 of 72 subjects, the CLL-method failed to produce accurate results. Nevertheless, using the strong correlation between CSA_m^t and CSA_m^c such cases can easily be identified.

5.5 CONCLUSION

Generally CSA_m^t and CSA_m^c only differ slightly, however for some cases these difference were quite substantial. It is likely that such errors will have relevant practical consequences for the assessment of GITC. The current hypothesis is that CSA_m^c are more relevant in term of ventilatory consequences, however this has not yet been verified.
Furthermore, the current method for assessment of CSA_m^c fails to reproduce accurate results in rare cases (9.72% of assessed patients). Nevertheless, using the strong correlation between CSA_m^t and CSA_m^c, and additional visual inspection, such cases can easily be identified.

6

RADIOLOGICAL ASSESSMENT

Goitre is an enlargement of the thyroid which may induce trachea compression, thereby increasing in airway resistance, which may cause ventilatory impairment.
Assessment of Radiological Parameters (RP) is a commonly used method used for the evaluation of Goitre Induced Trachea Compression (GITC). However, documentation and interpretation of these parameters has not been standardized yet.
Radiologists generally describe their findings in rather subjective statements such as *'mild compression'* or *'severe deviation'*, whereas, the interpretation of these terms is unclear. Although some radiologist also quantify their observations (e. g. a constriction of 50%), still a lot of ambiguity remains concerning the interpretation of such statements, e. g. does this concern reduction in Cross-Sectional Area (CSA) or diameter, and how is this diameter defined?
Providing radiologists with a structured protocol for the assessment of RP, will most likely improve the reproducibility and objectivity of the assessment of RP. An alternative might be the application of an automated segmentation tool. The usage of automated methods will most probably further improve the reproducibility and objectivity of the assessment of RP. However, it is currently unknown whether these differences are actually significant for the assessment of GITC in clinical practice.
This study aims to address this issue by the implementation of a protocol for the assessment of RP. The reliability of this method will be tested by assessing inter- and intra-observer agreement between physicians using this protocol. Additionally, the results of this analysis will be compared to an automated segmentation tool, to yield insights in the added value of such a tool in clinical practice.

6.1 METHODS

A protocol was devised for the assessment of trachea dimensions in the transversal and coronal plane.
The assessment in the transversal plane aimed to quantify the trachea diameter, whereas the assessment in the coronal plane was used to quantify trachea deviation.
Automated assessment of RP was limited to the assessment of trachea diameters.

6.1.1 *Data*

Analysis were all performed on CT-images of the neck-region obtained from various patients. These CT-images were generated using a Philips Brilliance 40 slice CT scanner (Best, The Netherlands). This scanner generated transversal CT-images with a slice thickness of 2 mm each. Data in the coronal-plane were obtained by reconstruction of these transversal slices, as is customary in the assessment of RP.
In total 30 anonymous scans were scored. These 30 scans contained 10 scans from 9 control subjects (1 patient provided 2 scans), and 20 scans from 10 patients which underwent thyroidectomy (10 before and 10 after surgery scans). Control subjects included 2 males and 7 females, with an average age of 62.0 $\pm$ 14.3 years. They all had an indication for CT-scanning of the neck and mediastinum as part of the work-up for lung cancer staging, analysis of vocal cord paralysis or for parathyroid adenoma localization. Goitre-patients consisted of 2 males and 8 females with an average age of 58.3 $\pm$ 10.5 years.

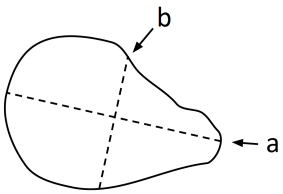
6.1.2 *Manual assessment*

RP were assessed in these scans by three radiologists using a novel protocol devised for this study. Prior to these measurements, radiologists were verbally briefed on this protocol. In addition to this, they received a written version of the protocol. During the briefing the radiologists were not informed about the characteristics of the study population, thus they were unaware about the amount of control and goitre-patients. Radiologists repeated measurements approximately 2 months later to allow the assessment of intra-observer variability. All measurements were performed individually and data were presented in random order during these measurements.

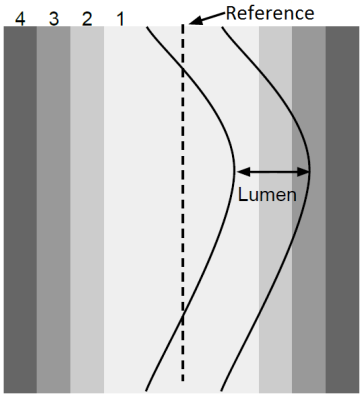
6.1.2.1 *Protocol*

TRACHEA DIAMETER

Trachea diameters were assessed in the transversal plane. The slice with maximal trachea constriction (i. e. largest reduction in cross-sectional diameter) was chosen as landmark for this assessment. Radiologist scrolled through the provided transversal slices and identified this Maximal-Constriction-Slice (MCS) by visual inspection.
After identification of the MCS, the radiologist measured and documented the trachea diameter in this slice. Hereby the diameter was read perpendicular midway to the longitudinal axis, as depicted in figure 41a. In order to provide insight in the radiologists decision making process, they also reported the slice-number of their presumed MCS for each scan.



(a) Assessment of trachea diameter. a represents the longitudinal axis and b is placed perpendicular midway to a.



(b) Optical illustration of assessment of trachea deviation in a trachea with deviation category 3.

Figure 41

TRACHEA DEVIATION

Trachea deviation was assessed in the coronal plane. Hereby the radiologists scrolled through the reconstructed coronal slices to identify the slice, which to their judgement was most relevant for the assessment trachea deviation.
After identification of this slice, the radiologists quantified the magnitude of trachea deviation by assigning it to one of four categories:

1. 0 to $\frac{1}{2}$ times the width of the trachea lumen
2. $\frac{1}{2}$ to 1 times the width of the trachea lumen
3. 1 to $1\frac{1}{2}$ times the width of the trachea lumen
4. $> 1\frac{1}{2}$ times the width of the trachea lumen

A visual reference of these categories has been included in figure 41b. The width of the trachea lumen was based on the trachea diameter 1 cm below the cricoid. This cricoid is a rigid bony structure which encircles the upper portion of the trachea. It is assumed that this cricoid prevent the compression of the trachea in the presence of pathologies such as goitre. Therefore is believed that this point is a valid reference for the general width of the trachea lumen.

6.1.3 *Automated method*

For a relevant comparison between automated and manual assessment of RP, it is important that this automated method uses the same guidelines as the manual method.
Therefore, a new segmentation tool was devised. This tool was an adaptation of the tool introduced in chapter 2. In brief, first a region

of interest was defined using a graphical user interface. Transversal slices within this region of interest were selected for further analysis. The user manually indicated the position of the trachea in the first selected slice. Next, all slices were converted to a binary image, based on a threshold which was believed to differentiate between tissue and air-containing structures. Using the presumed position of the trachea, provided by manual input, the trachea was segmented within the first slice. Coordinates of this segmented trachea slice were used as a reference for trachea position in subsequent slices. This allowed for an automated segmentation of the entire trachea. After these steps, the current method diverges from the method described in chapter 2.
Rather than constructing a 3d-model, slices are assessed individually in the transversal plane using the Matlabs regionprops function. This function was used to assess the Minor Axis Length (MAL) of the trachea. This MAL is always obtained orthogonal to the longitudinal axis, similar to the methods of section 6.1.2.1. Thereafter, the slice with the smallest MAL was chosen as the MCS.

6.2 RESULTS

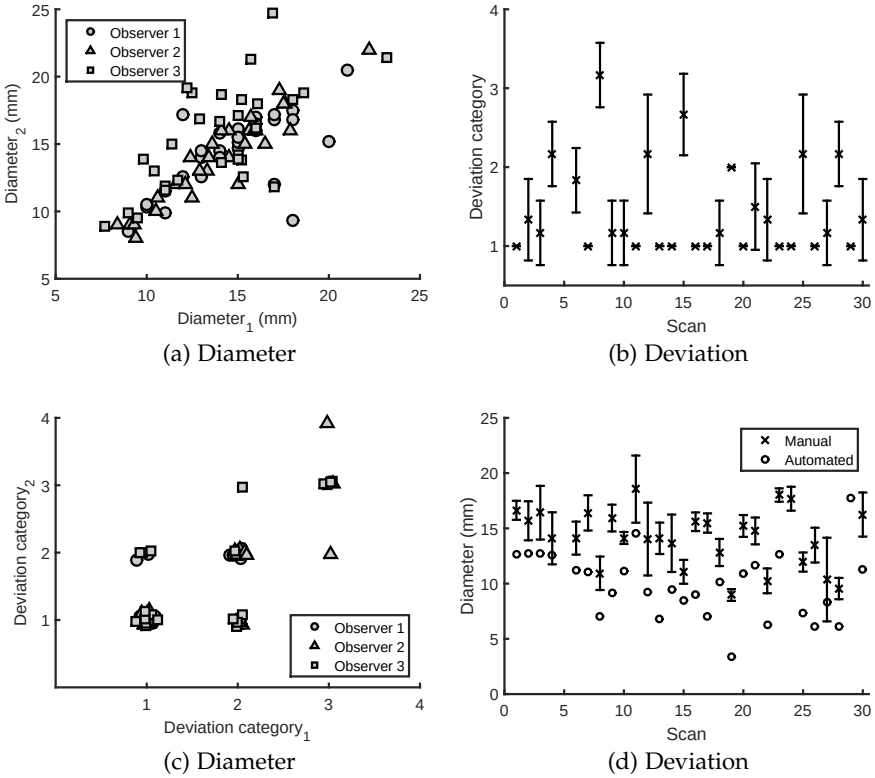
6.2.1 *Manual assessment*

6.2.1.1 *Diameter*

Each scan was assessed 6 times (twice by every observer), the mean documented diameter was 14.41 mm.
Figure 42a compares documented diameters in both measurements for all radiologists. Intra-observer agreement-values for the assessment of trachea diameter were 0.62, 0.95 and 0.65 (figure 43a).
Figure 42d shows the variance for documented diameters for each of the assessed CT-scans. The overall mean SE was 1.52 ± 0.85 mm. This SE was not correlated to the mean diameter ($r = -0.0104, p = 0.9573$).
Inter-observer agreement-values for the assessment of trachea diameter were obtained using these data and are provided in figure 43a. The mean inter-observer r was 0.75 ± 0.15.

SLICE SELECTION

Figure 43b shows that r for slice selection during diameter assessment varies greatly per observer and time of measurement (range $0.2470 \leftrightarrow 0.9920$). Correlations of the second observer during the second measurement are all fairly low (mean 0.2860 ± 0.0359) and not significant ($p > 0.05$).
Figure 44 shows that there is no obvious relationship between the agreement in diameter assessment and slice selection ($r = -0.4996, p = 0.0579$), exclusion of correlations of the second observer during the second measurement does not improve this correlation ($r = -0.4807$,



(a) Diameter (b) Deviation

(c) Diameter (d) Deviation

Figure 42: Documented diameters and magnitude of deviation for all scans.

$p = 0.1596$).

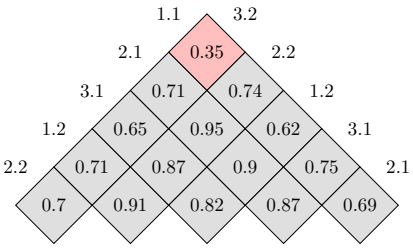
6.2.1.2 *Trachea Deviation*

Figure 42b compares documented categories in both measurements for all radiologists. Intra-observer agreement values for the assessment of deviation category were 0.76, 0.84 and 0.72 (figure 43c). Variance for documented categories for each of the assessed CT-scans are illustrated by figure 43c. The overall mean SE was 0.27 ± 0.26 category. Inter-observer agreement values for the assessment of trachea deviation were obtained using these data and are provided in figure 43d. The mean inter-observer r was 0.72 ± 0.09.

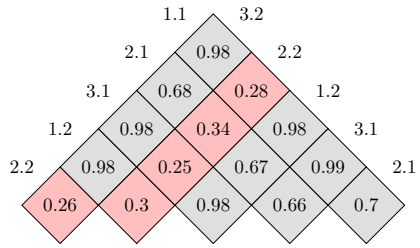
SLICE SELECTION

Observer 3 failed to document the selected slices for assessment of trachea deviation. Therefore only the results of observer 1 and 2 are included.

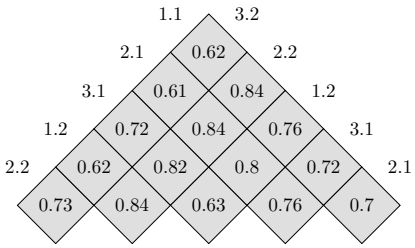
Intra-observer r were 0.97 and 0.93. Mean inter-observer r was 0.94 ± 0.05.



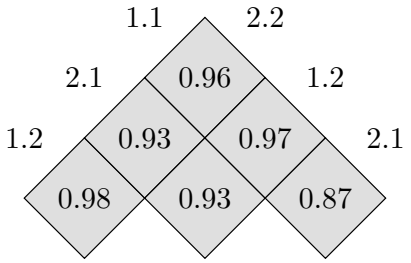
(a) Assessment of diameter



(b) Slice selection for diameter assessment



(c) Assessment of deviation



(d) Slice selection for diameter assessment

Figure 43: r-values for manual assessment of trachea dimensions. The number preceding the period indicates the observer, wheras the number after the period indicates the time of measurement, e. g. 1.2 indicates the second measurement by the first observer. Non-significant correlations ($p > 0.05$) are highlighted in red. The third observer did not document the slices for the deviation assessment, therefore these values are absent in figure 43d.

6.2.2 *Automated assessment*

Mean diameter was 9.88 ± 3.08 mm.
Table 8 and figure 42d show that correlation between the automated and manual assessment is fairly good. However, the manual method yields bigger diameters than the automatic method for all instances (correlation between average manual values and automated method: $r = 0.8397$, $p = 2.3196 \times 10^{-8}$). Nevertheless, the magnitude of this difference does not appear to be greatly affected by the diameter magnitude ($r = -0.3777$,$p = 0.475$, diameters obtained through automated assessment are used as reference).

6.3 DISCUSSION

6.3.1 *Manual assessment*

It is assumed that RP documented by a standardised protocol, such as the one described in this chapter, yields more objective and reliable

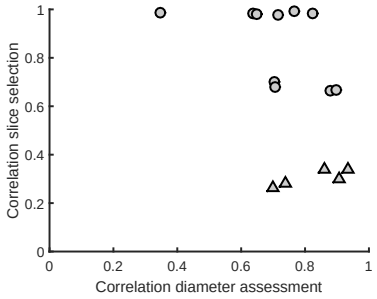


Figure 44: Relation between diameter assessment and slice selection, triangular markers indicate measurements by the second observer during the second measurement, the remainder are indicated by the round markers.

	1st	2nd
1	0.61	0.71
2	0.79	0.78
3	0.76	0.84

Table 8: Correlation between manual and automated assessment for the three radiologists during first and second measurement.

results than the methods currently used in clinical practice.
Nevertheless, there is still a lot of variance in documented results using this protocol. In the case of trachea diameters, this involves several millimetres. As the results indicate, these variances are not linked to magnitude of the diameters. This implies a absolute error rather than a relative one. As literature indicates, when a trachea is already severely constricted, any additional constriction will have major consequences in terms of airway resistance [68]. Thus such errors might have major consequences for the prediction of ventilation impairment using RP. However, it is likely that Cross-Sectional Area (CSA) is more relevant than trachea diameter in terms of ventilation impairment due to GITC [91]. Therefore it is important to also determine the accuracy of manual assessment of CSA.
Results show that a high correlation between selected slices is not a good predictor for a high agreement in documented trachea diameters. This implies that these observers interpreted and/or executed the protocol differently. However, from the documented results is difficult to determine the exact cause of these discrepancies.
Inter-observer agreement is slightly higher than intra-observer agreement for diameter assessment and vice versa for assessment of deviation. Nevertheless, these differences were only minor. Whether such differences are relevant in clinical practice remains unknown. Nevertheless, it seems that there is no added value in preferring the same radiologists when comparing RP for different intra-patient measurements according to the proposed protocol.

Differences in obtained diameters could not be linked to slice selection. Thus it seems that the implementation of the protocol differs per observer and is not (exclusively) linked to the visual identification of the slice with maximal trachea constriction.
More insight in the causes of the differences in obtained diameters can be generated by letting several radiologists score the same slice.

6.3.2 *Automated versus manual*

From figure 42d it is evident that automated and manual assessment show good correlation, however diameters obtained by manual assessment are consequently higher than those obtained by automated assessment. These findings could be explained in various manners.
First there is the radio-density threshold used for the segmentation of trachea slices. In chapter 2, it was motivated that this threshold was chosen lower than the values recommended in literature. This was done to prevent leakage, i. e. assigning lung-tissue to the trachea lumen, in CT-images were the distance between the trachea lumen and lung-tissue was fairly small. This lower threshold might cause the algorithm to disregard a portion of the pixels which contained noise due to their proximity to tissue with different radio-density values, such as the trachea wall. An option to circumvent this problem would be to use Magnetic Resonance Imaging (MRI), which generally has lower noise-signal ratio than than CT-imaging, allowing for the usage a less extreme radio-density threshold.
Another explanation would be a consequently different interpretation of the diameter-assessment protocol by the radiologists or an inability to correctly identify the slice with maximal trachea constriction. This theory can be verified by letting the radiologist as well as the automated method assess identical slices.
Another option would be to compare the results to a 'golden standard'. This can be done by scanning a phantom with known dimensions and assessing its dimensions using both methods or comparing its results to another method with proven accuracy.

6.4 CONCLUSION

Assessment of trachea dimensions using the various assessed methods show good correlation and reproducibility. The absolute magnitude of obtained RP differ per method. The functional consequences of these disparities in terms of UAO should be evaluated using additional validated methods to decide on the most suitable method for assessment of RP.

7

PROSPECTIVE STUDY: ADDITIONAL PATIENTS

Goitre Induced Trachea Compression (GITC) is a disease which compromises airway dimensions, this may raise airway resistance, thereby causing ventilatory impairment. As chapter 1 described, several therapies, each with different efficacies, are available for the management of GITC. To choose the best suited therapy, the consequences of GITC should be assessed accurately. However, as motivated in chapter 1, tools commonly used tools in current clinical practice are suboptimal for management of GITC. In chapter 2 a new and promising measure (the MVC) was introduced and tested on 10 patients suffering from goitre and 10 controls. Preliminary results suggested that the MVC may be useful in clinical practice but follow-up research is needed to establish the significance of this method more firmly.
In view of these findings a follow-up study was performed (chapter 3). This chapter evaluated the methods of various studies documented in literature regarding the assessment of GITC, and proposed several improvements on the methods of these studies. Thereafter a retrospective analysis was performed, which analysed 72 patients with goitre and 55 healthy controls. This analysis managed to firmly define the range for healthy MVC-values as well as other Radiological Parameters (RP) and evaluate some of the improvements proposed in the literature study. However, due to the retrospective design of this study, it was not possible to compare the results of this analysis to a validated alternative.
Therefore, a another follow-up study was initiated as described in chapter 4. This was a exploratory study which aimed to identify potentially useful measures for the evaluation of GITC, including some novel methods which have not been evaluated extensively yet. These measures may be useful an sich, but also allow to evaluate the MVC and other RP assessed in chapter 3.
This study aims to evaluate the potential of the measures identified by chapter 4 and verify the improvements proposed in other chapters. First interpretation of these measures will be described. Next, the improvement proposed in the preceding chapters will be discussed, and how they can be verified.

7.0.1 *Background*

In chapter 4 several methods were employed to assess GITC in a single patient (patient X). This included the assessment of Self Reported Symptoms (SRS), Pulmonary Function Tests (PFT) and assessment of

Radiological Parameters (RP). Based on findings in literature and the results of chapter 4, some of these methods could already be disregarded as potential useful tools for the management of GITC. The remainder of these methods, which do seem useful for the management of GITC, are included in the methods of the current study. This section will discuss the relevance and interpretation of these tools.
Additionally, this study aims to verify some of the assumptions made in preceding chapters.

7.0.1.1 *Pulmonary Function Tests*

SPIROMETRY

Pulmonary Function Tests include a wide variety of methods, including traditional spirometry, plethysmography and Impulse Oscillometry Spirometry (IOS). Traditional spirometry is used to generate Flow-Volume Loops (FVL) from which Pulmonary Function Values (PFV) are monitored. These PFV are objective, predefined variables obtained from FVL. However, as motivated in literature as well as chapter 4, expiratory PFV are not sensitive for GITC, contrary to inspiratory PFV. However, reference values for these PFV are not available. Nevertheless, a correlation between inspiratory PFV and RP has been documented in a study performed by Bonnema *et al.* [91]. Nevertheless, the methods for assessment of RP applied in this study, differ from those of Bonnema *et al.* Naturally, it is important to know how these RP relate to the findings of Bonnema *et al.* Therefore, PFT were performed whereby Flow-Volume Loops (FVL), Pulmonary Function Values (PFV) are monitored by traditional spirometry. These FVL were visually inspected by expert physicians, to ensure that the patient was able to successfully perform spirometry and that PFV derived from these FVL were reliable.

IMPULSE OSCILLOMETRY SPIROMETRY

Additionally, Impulse Oscillometry Spirometry (IOS) has been identified as an useful alternative to traditional spirometry. IOS can be used to monitor a wide variety of indices, including resistance, reactance and fundamental frequency. However, as illustrated by literature, these measures are either not specific (because they might be confounded by diseases, such as distal airway obstruction and/or airway restriction), or not sensitive for GITC (because intra-individual variation in healthy patients is too high). A possible exception is flow corrected resistance ($\frac{\Delta R}{\Delta \dot{V}}$). To our knowledge, usage of the $\frac{\Delta R}{\Delta \dot{V}}$ for the identification of tracheal stenosis was only performed by Verbanck *et al*[32]. Verbanck *et al.* measured this $\frac{\Delta R}{\Delta \dot{V}}$ at 5 Hz oscillations during quiet, tidal breathing. Their results suggested that $\frac{\Delta R}{\Delta \dot{V}}$ increases with decreasing CSA_m. However, these measurements were performed us-

ing tradional Forced Oscillation Technique(FOT), which is similar to IOS but does not allow to vary the frequency oscillations. It is well-known that Resistance measured at 5 Hz represents the total airway resistance (distal and central airways), whereas R measured at 20 Hz represents the central airways. This makes it likely that $\frac{\Delta R}{\Delta \dot{V}}$ measured at 20 Hz is a better indicator for GITC than $\frac{\Delta R}{\Delta \dot{V}}$ measured at 20 Hz. To verify this assumption, $\frac{\Delta R}{\Delta \dot{V}}$ will be measured at both frequencies and compared to the CSA_m.

7.0.1.2 *Radiological parameters*

Prior to the assessment of Radiological Parameters (RP), trachea dimensions have to be obtained. This can be done in various manners, for instance using a automatic (using a computational algorithm) or manual method (visual inspection). Chapter 6 assessed both methods and concluded that automated assessment and manual assessment show a good correlation. Additionally, the automated method has superior reproducibility and is less labour-intensive. However, although correlation between both methods was good, there still was a notable underestimation of dimensions using the automated method compared to the manual method. However, results of chapter 6 were not able to identify which of these methods gave the best approximation of actual trachea dimensions. In view of these findings, the automated method is regarded the best option, however additional research is strongly suggested, in order to verify this assumption.
Additionally, assessment of trachea diameters can be done in the original transversal CT-slices (transversal-method), or angled using a Centre Lumen Line in a reconstructed 3d-model of the trachea (CLL-method). The implementation of both these options has been evaluated in chapter 5. This chapter reported that both methods show good correlation for the assessment of Minimal Cross Sectional Area (CSA_m), however in some cases differences of 60 mm^2 have been observed. In view of the earlier mentioned linear correlation between CSA and PFV, it was found that such differences might cause errors up to 2.2 L s^{-1} in predicted inspiratory PFV [91]. Hereby it is assumed that the CLL-method yields more relevant results, since it calculated trachea dimensions orthogonal to the direction of the airflow. However, this assumption has not been verified yet. Additionally, it was reported that in a small portion of assessed scans ($\approx$ 10%), the CLL-method failed to successfully reconstruct the trachea geometry. However, these cases could easily be identified through visual inspection. In view of these findings, it is recommended to employ both methods for the assessment of CSA_m and compare them to $\frac{\Delta R}{\Delta \dot{V}}$ and PIF. The $\frac{\Delta R}{\Delta \dot{V}}$ and PIF were chosen since both these methods have previously been linked to CSA_m [32, 91].
Chapter 4 assessed three different RP, including the Minimal Cross Sectional Area (CSA_m), Maximal Ventilation Capacity (MVC) and

Percentage Trachea Constriction (%-TC). Assessment of the CSA_m is self-explanatory, it is the minimal observed cross-sectional area of the trachea lumen. %-TC describes the percentage of diameter reduction by comparing the minimal diameter of the trachea (d_{min}) to a patient-specific reference diameter (d_{min}). This reference CSA is assessed in a portion of the trachea which is assumed to be not be affected by GITC.
The MVC also uses this reference point. Hereby d_{ref} is extrapolated over the rest of the course of the trachea, to generate a prediction of healthy trachea dimensions, i. e. the presumed dimensions of the trachea if the patient would not have developed a goitre. Thereafter, the actual trachea dimensions are used to predict the magnitude of flow the patient is able to generate in this diseased state ($\Phi_{disease}$). Next, the predicted healthy trachea dimensions are used to predict the magnitude of flow the patient would be able to generate in the healthy state ($\Phi_{disease}$). The ratio between these levels of flow are assumed to reflect the magnitude of ventilation impairment imposed by the GITC and is defined as the Maximal Ventilation Capacity (MVC) according to equation 34:

$$MVC = \frac{\Phi_{disease}}{\Phi_{healthy}} \times 100\% \tag{30}$$

Assessment of MVC, %-TC and CSA_m in chapter 3 revealed that a larger portion of pre-surgery patients fell outside the range of normal values based on %-TC (51 % of patients) compared to CSA_m (21 % of patients) and MVC (31% of patients). This implies that %-TC might be a more sensitive measure for identification of GITC compared to CSA_m and MVC. This can be explained by the strong correlation of intra-individual trachea dimensions. Because %-TC defines the magnitude of the minimal trachea diameter (d_{min}) to a reference diameter for that specific patient (d_{ref}), it has a strong indication for the expected magnitude of d_{min} (at least 70.9% of d_{ref} for healthy patients). On the other hand, CSA_m shows a lot of intra-individual variation, even when corrected for gender, height, age and weight, thereby hampering the identification of diminished CSA_m-values as a result of GITC.
Although the MVC also uses reference values to assess the magnitude of GITC, the effect of trachea stenosis may be masked by portion of the trachea with more than proportional trachea widening. This is illustrated by figure 45. Basically, the MVC is a measure of the area which lies under the reference line and above the actual trachea dimensions. Portions of the graph above this reference line (caudal portion of the trachea in figure 45), might mask the effect of stenosis in other parts of the trachea.
However, out of these three measures (%-TC, CSA_m and MVC), only for CSA_m a significant correlation with pulmonary function has been

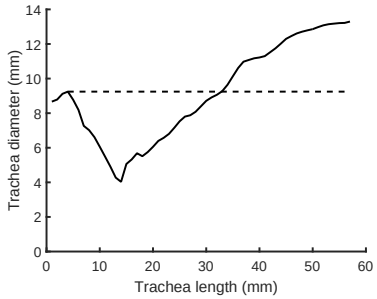


Figure 45: Trachea diameter of 71 year old female, MVC = 116.84 and %-TC = 56.41.

identified [91]. This earlier mentioned correlation was described by Bonnema *et al.*, and translated into equation 31:

$$FIF_{50\%} = 0.036\, CSA_m - 0.67 \tag{31}$$

Hereby the $FIF_{50\%}$ represents the forced inspiratory flow at 50% of the vital capacity in L s^{-1} and CSA_m provided in mm^2. However, the issue with radiological parameters such as CSA_m is that their interpretation in unclear. Since reliable reference values for inspiratory PFV such as the $FIF_{50\%}$ are absent, it is difficult to verify whether calculated levels of $FIF_{50\%}$ are pathological. Nevertheless, it is still highly relevant to reproduce this correlation, because this can help to provide insides in the physical processes underlying ventilation impairment due to GITC.
Additionally, in view of the documented correlation between CSA and PFV, it is interesting to know whether such correlations can also be observed when using a circular equivalent diameter (d_e), as defined by equation 32:

$$d_e = \frac{CSA^{\frac{5}{8}}}{P^{\frac{1}{4}}} \tag{32}$$

Whereby CSA is the cross sectional area in mm^2, P the perimeter in mm^2. In chapter 2 it was hypothesised that d_e is more relevant for an approximation of airway resistance than CSA. Since compressed trachea's often have irregularly teardrop shaped cross sections, ratio between the trachea wall-surface and trachea lumen volume may increase, thereby causing an increase in shear stress.

7.0.1.3 *SRS*

Self Reported Symptoms can be assessed in various ways, but most commonly this will be done by means of a questionnaire or an interview. Generally, a questionnaire will be most suited to monitor diseases with known consequences in a structured manner, whereas an interview will be more suitable to explore more complex diseases,

whereby relevant variables are unknown. Since GITC symptoms are already thoroughly assessed in literature and to ensure reproducible results a questionnaire was assumed most relevant. However, no questionnaire relevant for GITC could be obtained, therefore a novel questionnaire was devised, aimed specifically at GITC.
This questionnaire was used to monitor Functional Consequences (FC), Emotional Impact (EI), Symptomatic Consequences (SC), Borg Score (BS), General Health (GH) of the patient. Logically, some of these variables, such as SC, GH and EI, will not provide a accurate reflection of ventilation impairment due to GITC, but were included to monitor the effects of GITC on the patients quality of life. Therefore, only BS and SC were compared to PFT and RP.

7.1 METHODS AND SUBJECTS

This is a prospective study and inclusion of patients is still ongoing. This chapter describes the methodologies used in this study and presents its preliminary results.

7.1.1 *Patient characteristics*

12 patients (2 males and 10 females) suffering from goitre were included. All of these patients were awaiting a thyroidectomy. Average age was 56.54 $\pm$ 11.95 years (mean $\pm$ SE), average weight was 88.08 $\pm$ 25.29 kg and average height 165.38 $\pm$ 9.46 cm, one of the patients used tobacco and three had Possible Concomitant Diseases (PCD), including hypertension, renal failure, asthma and heart palpations. Table 9 provides a more detailed overview of patient characteristics.
This study aims to evaluate GITC in these patients prior to and after surgery. To ensure patients had sufficient time to recover from surgery, a 3-month hiatus between surgery and the post-surgery evaluation was indicated. At the current time, some patients already underwent surgery, however, the criterion of this 3-month hiatus has not yet been met, as a result of this, only pre-surgery results were included.

7.1.2 *Study design*

Section 7.0.1 identified several useful measures for the assessment of GITC, these measures will be assessed and discussed in this study. Additionally, the measures which are not evaluated but measured nonetheless, are provided in the appendix ....
First RP were assessed, these include CSA(measured using the transversal and CLL-method), MVC, d_{min} (minimal circular equivalent di-

Pt.	Gender	Age (yr.)	Weight (kg)	Stature (cm)	Smoking	PCD
1	F	71	76.5	163	Y	Y
2	F	55	76	168	N	N
3	F	56	71	154	N	N
4	F	69	61.5	159.5	N	Y
5	F	53	76	164.5	N	N
6	M	52	80	167	N	Y
7	F	34	80	155	N	N
8	F	53	72	157	N	Y
9	M	75	85.5	177	N	N
10	F	45	94.5	164.5	N	N
11	F	59	107	159	N	N
12	F	44	105	175	N	N

Table 9: Patient characteristics. PCD indicates Possible Concomitant Diseases.

ameter) and %-TC (percentage reduction in diameter). These indices have already been compared to each other in preceding chapters, therefore they are only evaluated to compare them to SRS and PFT later on.

Thereafter PFT were performed, whereby PIF and $\frac{\Delta R}{\Delta \dot{V}}$ were selected for further analysis. CSA_m has been linked to inspiratory flow in a study by Bonnema *et al.* Bonnema analysed CSA_m in the transversal plane (CSA_m^t), however in section 7.0.1 it has been motivated that CSA_m measured orthogonal to a Centre Lumen Line (CLL-method), might provide more relevant results in terms of ventilation impairment. To verify this assumption, both these measures were compared to the Peak Inspiratory Flow (PIF).

After the most relevant orientation has been defined (transversal or CLL), d_{min} and CSA_m will be assessed in this plane. Thereafter, both these measures will be compared to PIF.

$\frac{\Delta R}{\Delta \dot{V}}$ measured at 5 Hz has been linked to CSA_m by Verbanck *et al.*, however in section 7.0.1 it has been suggested that $\frac{\Delta R}{\Delta \dot{V}}$ measured at 20 Hz might be a more relevant measure. To verify this, both these indices were compared to CSA_m. As an additional check, $\frac{\Delta R}{\Delta \dot{V}}$ for both frequencies were also compared to PIF.

These analysis will indicate whether CSA_m^c or CSA_m^t, d_e or CSA_m and $\frac{\Delta R}{\Delta \dot{V}}$ measured at 5 Hz or $\frac{\Delta R}{\Delta \dot{V}}$ measured at 20 Hz is the most relevant in terms of ventilation impairment due to GITC. After excluding two of these measures, the MVC and %-TC were also be compared to the PIF and $\frac{\Delta R}{\Delta \dot{V}}$, to verify which RP is most relevant.

Next, SRS will be assessed, this includes the assessment of EI, FC, SC,

BS and GH. Hereby the correlation between SRS will be quantified. Not all these indices are expected to reflect ventilation impairment, such as EI, SC and GH, these measures are solely assess to identify the impact of GITC on a patients quality of life. Logically, FC and BS will be the most relevant parameters for assessment of ventilation impairment. Therefore, this both these indices will be compared to presumed most relevant RP and PFT, based on the results from this study.

7.2 RESULTS

7.2.1 *Radiological parameters*

Mean MVC was 69.55 $\pm$ 13.22 %, 2 measurements were below the lower limit of normal (57.1%). Mean %-TC was 29.94 $\pm$ 15.43 %, 5 measurements were above the upper limit of normal (29.1). Mean MCSA-values 88.67 $\pm$ 42.81 mm^2 were using the transversal method and 87.66 $\pm$ 43.40 mm^2 calculated using the CLL-method. 2 measurements were below the lower limit of normal(93.86 mm^2 for males and 48.15^2 for females).

7.2.2 *Pulmonary Function Tests*

Mean $\frac{\Delta R}{\Delta \dot{V}}$ at 5 Hz was 0.1613 $\pm$ 0.065 kPA L^{-2} s^2, 3 measurements were above the upper limit of normal (0.2 kPA L^{-2} s^2 [32]). Mean $\frac{\Delta R}{\Delta \dot{V}}$ at 20 Hz was 0.1229 $\pm$ 0.066 kPA L^{-2} s^2, no reference data were available for this measure.
Mean PIF was 4.17 $\pm$ 1.45 L s^{-1}, no reference data were available for this measure.
Correlation between $\frac{\Delta R}{\Delta \dot{V}}$ at 5 and 20 Hz were and PIF was assessed. These analysis showed that $\frac{\Delta R}{\Delta \dot{V}}$ measured at 20 Hz had better correlation with PIF ($r = 0.057$, $p = 0.066$) than measured at $\frac{\Delta R}{\Delta \dot{V}}$ 5 Hz ($r = -0.37$, $p = 0.27$), see figure 47d and 47e.

7.2.3 *Self Reported Symptoms*

Table 10 shows the questionnaire results for all patients. Correlations between the different assessed indices are provided in figure 46.
All correlations between SRS were significant, except for GH with EI, FC and SC.

Pt.	FC (%)	EI (%)	SC (%)	BS (0-10)	GH (1-5)
1	0	12.5	6.6	3.5	4
2	0	0	5	2	4
3	100	37.5	60	10	2
4	0	0	4	1	4
5	12.5	25	9	1	4
6	0	0	22.2	1	4
7	25	25	33.3	5	1
8	37.5	37.5	65	6	3
9	62.5	37.5	54	4	3
10	12.5	0	9	2	3
11	50	0	14	5	3
12	75	25	25.3	6	3

Table 10: Questionnaire results for all patients. Abbreviations: Patient(Pt.), Surgery Status (SS), Functional Consequences (FC), Emotional Impact (EI), Symptomatic Consequences (SC), Borg Score (BS) and General Health (GH).

7.2.4 *Comparative analysis*

To yield insights in the relevance of proposed measures, they were compared to PIF, which served as a golden standard.
First, CSA_m^t and CSA_m^c were assessed to yield insights in the relevance of orientation in which RP are assessed, see figure 47a and 47b. Hereby CSA_m^c ($p = 0.074$) had a better significance level than CSA_m^t ($p = 0.156$).
Thereafter the relevance of the circular equivalent diameter and cross-sectional area were assessed by comparing d_{min} to PIF, see figure 47c.

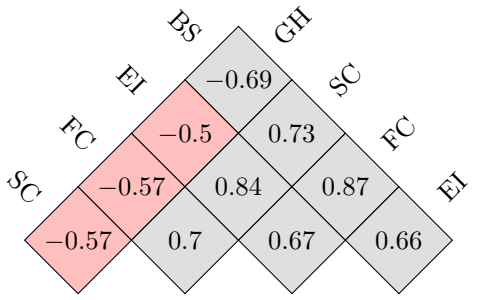


Figure 46: Correlation between self reported symptoms. Values in red cells were not significant ($p > 0.05$). Abbreviations: Functional Consequences (FC), Emotional Impact (EI), Symptomatic Consequences (SC), Borg Score (BS) and General Health (GH).

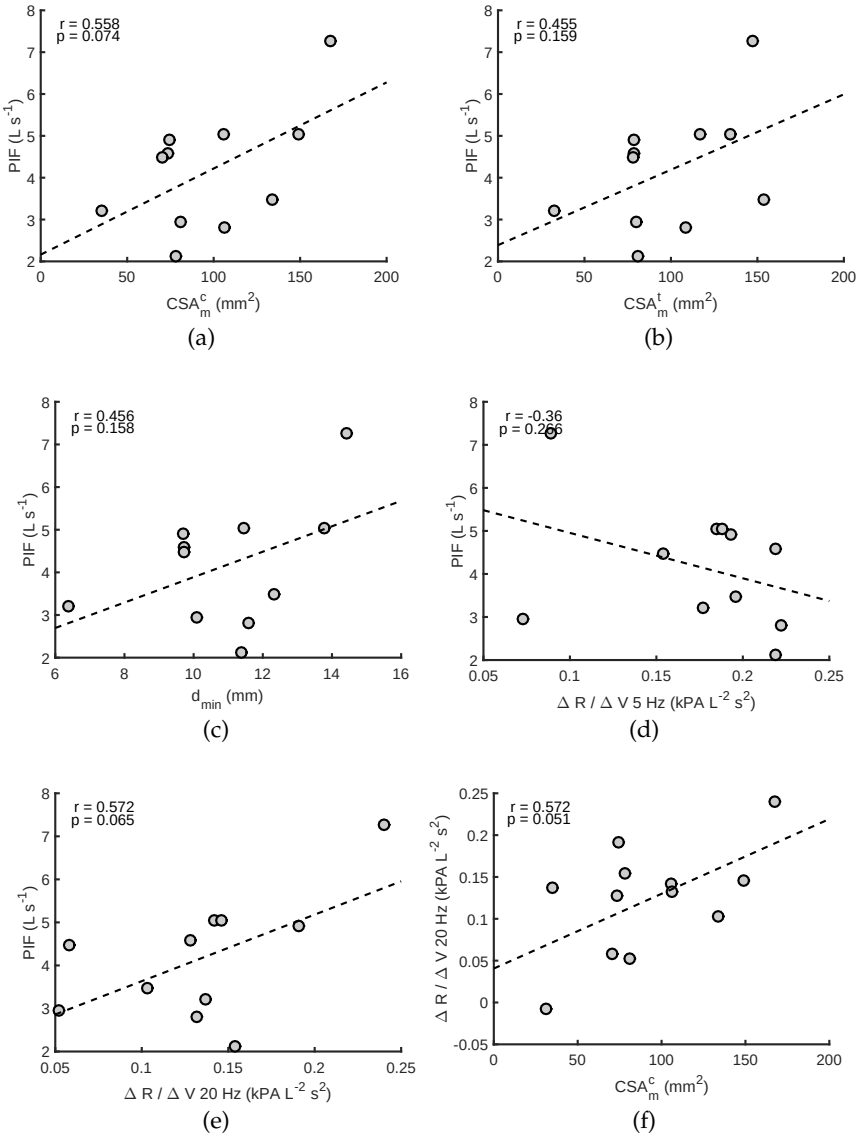


Figure 47

Correlation between CSA_m^c (r = 0.56, p = 0.074) and PIF was better than d_{min} (r = 0.46, p = 0.16) and PIF.
Next, $\frac{\Delta R}{\Delta V}$ at 5 and 20 Hz was compared to PIF. Hereby $\frac{\Delta R}{\Delta V}$ at 20 Hz (p = 0.065) showed better correlation than $\frac{\Delta R}{\Delta V}$ at 5 Hz (p = 0.27), see figure 47d and 47e.
As a countercheck, $\frac{\Delta R}{\Delta V}$ at 20 Hz was also compared to CSA_m^c, see figure 47f. This showed that $\frac{\Delta R}{\Delta V}$ at 20 Hz (p = 0.051) correlated better to CSA_m^c than PIF (p = 0.065). Based on these results, two hypothesis were established: (i) $\frac{\Delta R}{\Delta V}$ at 20 Hz is a better predictor for GITC than PIF, (ii) $\frac{\Delta R}{\Delta V}$ at 20 Hz shows better levels of significance because more samples were obtained for $\frac{\Delta R}{\Delta V}$ (n = 12) than PIF (n = 11). To assess these hypothesises, $\frac{\Delta R}{\Delta V}$ at 20 Hz was compared to CSA_m^c, whereby

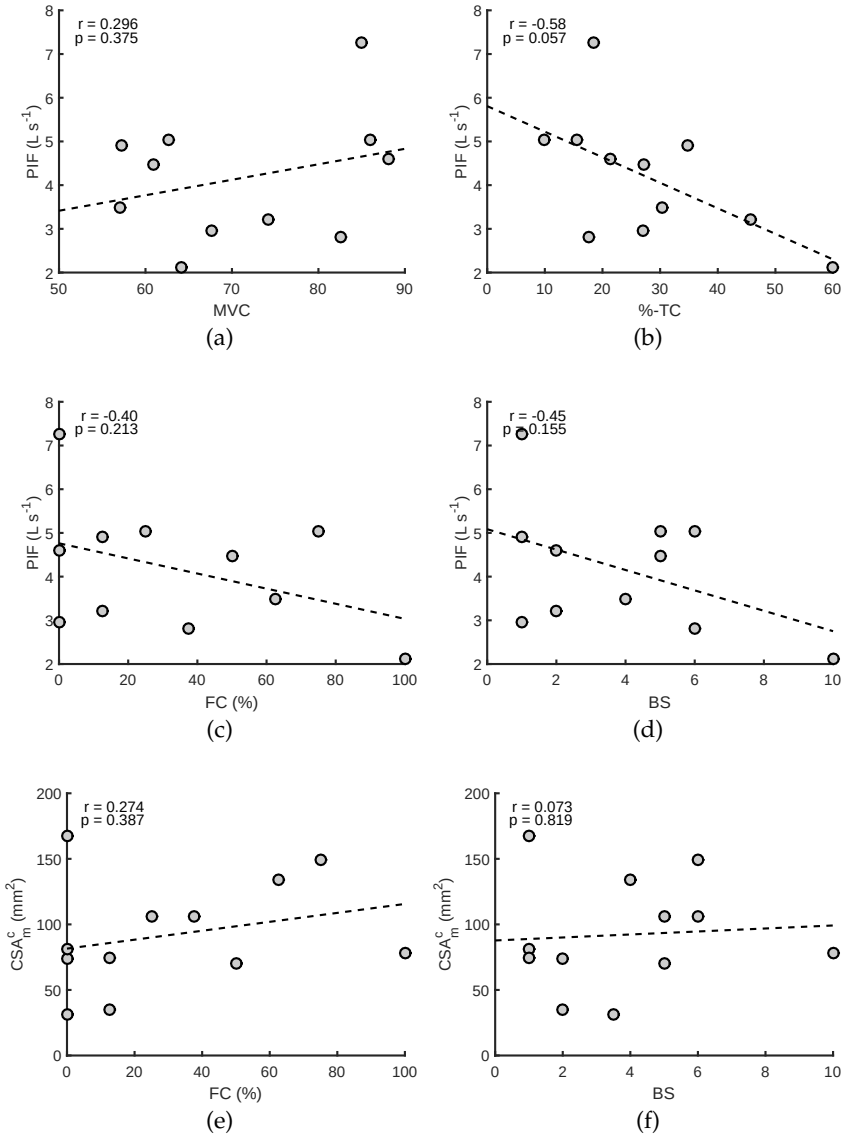


Figure 48

results for patient 1[1], were excluded.
After exclusion of this patient, significance for $\frac{\Delta R}{\Delta V}$ at 20 Hz was consistently lower ($p > 0.20$) than PIF.
Next, %-TC and MVC were compared to PIF. These results are provided in figure 48a and 48b. This showed that correlation between %-TC and PIF ($p = 0.057$) was better than MVC ($p = 0.375$) and CSA_m^c ($p = 0.074$) compared to PIF. Next PIF and CSA_m^c were used for the assessment of SRS. Hereby, it was assumed that Borg Score (BS) and Functional Consequences (FC) were the most relevant SRS for the assessment of GITC. The results of these analysis are provided in figure 48c, 48d, 48e, 48f, all these analysis reported poor correlations.

1 the patient who could not produce reliable FVL

7.3 DISCUSSION

7.3.1 *Comparative analysis*

Using a significance level of 5%, none of observed measures showed significant correlation, except for interrelations in SRS. Nevertheless, literature performing similar experiments, found significant correlations for some of these cases. One of these studies was performed by Bonnema *et al.*, hereby a significant correlation between CSA_m^t and Forced Inspiratory Flow at 50% of the vital capacity (which is highly correlated to PIF) was assessed [91]. Another study was performed by Verbanck *et al.*, which reported a correlation between $\frac{\Delta R}{\Delta \dot{V}}$ at 5 Hz and CSA_m^t [32]. Preliminary results of the current study suggest that the usage of CSA_m^c instead of CSA_m^t and $\frac{\Delta R}{\Delta \dot{V}}$ at 20 Hz and $\frac{\Delta R}{\Delta \dot{V}}$ at 5 Hz, yields even more relevant results in terms of ventilation impairment due to GITC. In view of these findings, it is likely that these non-significant correlations are caused by an insufficient sample size. Preliminary results seem to confirm some of the assumptions mentioned in section 7.0.1. However, these results should be carefully interpreted. First of all, none of the observed correlations between RP and PFT were significant. Furthermore, sample size was fairly low (n = 12). Additionally, only linear correlations were assessed. Since it is well-known that the correlation between airflow and trachea dimensions is not a linear one, it is likely that correlations will improve by using power-law functions to fit the data [68]. Nevertheless, employing such techniques increases the degree of freedom of such fits. In view of the small sample sizes of this study, it is likely that this will cause an over-fit of the data. Additionally, it was shown that the exclusion of patient 1, has drastic consequences for the significance level of some analysis. This further stresses the need for inclusion of more patients as well as post-surgery data for already included patients.
Another option to improve the reliability of the results is to correct for patient characteristics, such as age and gender. This was deliberately not done in the current study because it was assumed that this could lead to an over-fit of the data due to the low-sample size.
The positive correlation between $\frac{\Delta R}{\dot{V}}$ and CSA_m was remarkable, because Verbanck *et al.* reported a negative correlation between these two measures. A possible explanation for this phenomenon could be the manner in which $\frac{\Delta R}{\dot{V}}$ is defined. During this study, inspiration was defined as a positive change in flow ($\Delta \dot{V}$), whilst expiration was regarded negative, reversing these definitions can cause a negative correlations, as seen in the study of Verbanck *et al*[32].
Current results remain inconclusive about the significance of $\frac{\Delta R}{\dot{V}}$ compared to PIF. Nevertheless, $\frac{\Delta R}{\dot{V}}$ seems to have some benefits over the usage of PIF. First of all, the lack of effort-dependence in $\frac{\Delta R}{\dot{V}}$ means that some patients which might not be able to successfully

perform spirometry, might still be analysed using IOS, as illustrated by patient 1 [29]. Nevertheless, similar to spirometry, results still depend greatly on the investigator's level of familiarity with the device [52, 53, 54, 55, 56, 57, 58, 59, 60]. Additionally, the position of the patient can also greatly affect results, this makes that the ability of the investigator to correctly inform the patient how he should be seated as well as the patients ability to comprehend such instructions also affect the accuracy of this method. Furthermore, to our knowledge no normative data for $\frac{\Delta R}{V}$ at 20 Hz are readily available.

7.3.2 *MVC*

The poor correlation between the MVC and golden standard do not raise any major concerns regarding the relevance of the MVC. MVC is a measure of relative ventilation impair, i.e. an individuals percentage reduction in ventilation. This percentage ventilation reduction is assumed proportional to a percentage reduction in peak flow, whilst PIF-values measured in this study reflect an individuals absolute pre-surgery peak flow. The lack of normative data for PIF-values makes it difficult to predict an individuals healthy flow-capacity, thus hindering an accurate validation of the MVC. However, post-surgery PIF-values will most probably provide more insights on this matter. Ideally, this MVC is proportional to the ratio in pre and post-surgery PIF, thus:

$$MVC = \frac{PIF_{\text{pre-surgery}}}{PIF_{\text{post-surgery}}} \tag{33}$$

Hereby it is assumed that $PIF_{\text{post-surgery}} = PIF_{\text{healthy}}$. However, the lack of normative data for PIF makes it hard to validate this assumption, further stressing the need for accurate predictors of healthy-PIF data. That being sad, even if one would manage to confirm the MVC's ability to predict percentage reduction in flow, as described in equation 33, how does that help physicians in clinical practice? One of the issues physicians experience in clinical practice, is convincing patients of the potential benefits of intervention. The MVC's ability to express a patients presumed reduction in ventilation capacity in an objective measure is already a major leap forward from methods currently employed in clinical practice. However, how should this MVC be interpreted, does a MVC of 40% indicate that a patient is impaired in his abilities to tie his own shoes, or will it only impair his ability to perform high intensity exercise such as sprinting for a prolonged time? To answer these questions a system is necesarry, which illustrates how the MVC's should be interpreted. An attempt at such a system has been made and is covered in chapter 8.

7.3.3 *Questionnaire*

Although questionnaires might be an useful tool to assess consequences of a disease, they should be interpreted carefully. Information provided by patients can diverge significantly from reality [121]. Moreover, perception of symptoms per individual can differ greatly and might be influenced by a variety factors such as on the frame of reference, mood and personality. This was illustrated by patient 3 who reported a BORG-score of 10. For a lot of patients such a score would indicate extreme life-threatening situations, whereas the researchers conducting the questionnaire indicated that this patient's presence was relative calm and she did not appear to be in acute distress. A possible solution would be to ask the researcher to indicate the level of dyspnea based on his perception of the patient. However, the validity of such symptoms will also be questionably. Even if all questionnaires are evaluated by the same researcher, the manner in which patients express their symptoms both verbally as non-verbally will differ.

Another main drawback of the questionnaire is its inability to assess whether the reported symptoms actually originate from GITC or have other causes. Goitre, becomes more frequent with advanced age, similar to a lot of other diseases, making it quite likely that patients suffer from concomitant diseases, this is confirmed by the current study, whereby 25% of patients had such concomitant diseases.

A reliable method to monitor the consequences of GITC, should off course indicate an absence of such consequences in patients without or insignificant GITC. Thus to ensure that a questionnaire is a valid method to monitor GITC, it should also be tested in a variation of healthy individuals. Unfortunately, the validation of the current questionnaire in such a manner exceeded the scope of this study. Therefore it is recommended to perform additional research before making conclusive statements about its significance.

Furthermore, the poor levels of correlation with other measures observed in this study does not indicate that the questionnaire is not a reliable method for the assessment of SRS. Literature indicates that an individuals maximal minute ventilation decreases with advance age and also depends on gender [69]. Furthermore, the minute ventilation required for an individual to perform certain activities is not greatly affected by age, but does differ with gender and physical size [92]. Thus the fraction of a persons maximal minute ventilation required to perform certain activities differs with this persons age, gender and built. Additionally, the impairment in functional capabilities (FC) in this study was monitored on a linear scale, i. e. if an individual was unable to perform 25% of a variety of activities he got a FC score of 25%. However a doubling in FC, thus he would not be able to perform 50% of activities from that list, does not necessarily correspond to a

doubling in flow reduction. Thus the linear correlation used in this study to monitor the link between SRS and other measures are most probably a suboptimal method for this specific cause. This indicates that the interpretation of obtained SRS and how they are expected to interact with limitations in ventilatory capabilities should be defined for firmly before a conclusive statement about the relevance of this method can be made.

7.4 CONCLUSION

The current study seems proficient in the assessment of a variety of assumptions employed in proposed methods for the assessment of GITC, e. g. are trachea dimensions using the CLL-method more relevant than the transversal-method? Nevertheless, at the current moment none of the assessed correlations yield significant results, inclusion of more patients will probably resolve some of these issues. Additionally, the analysis employed in this study (assessment of linear correlation), are suboptimal for the assessment of some of these measures due to the different principles that underlie them, e. g. the relation between PIF and MVC. Assessment of post-surgery data, as well as more advanced methods (power-law fits instead of linear correlations), might possible benefit the assessment of these measures.

8

IMPLEMENTATION OF THE MVC

Goitre is an palpable enlargement of the thyroid which may cause trachea malformation. This malformation can raise airway airway resistance, thereby causing ventilatory impairment [3]. Various modes of treatment with different efficacies are available for the treatment of Goitre Induced Trachea Compression (GITC) [4, 5, 6, 7, 15, 17, 18, 19]. To decide on the most suitable therapy it is important to assess the clinical impact of GITC accurately.

Due to the slow progressive nature of the disease, patients are often unaware to what extent GITC impairs their physical capabilities, which makes Self Reported Symptoms (SRS) unreliable. An alternative is the assessment of Radiological Parameters (RP), but these often lack a firm objective basis and functional consequences remain unknown [61]. The current golden standard for assessment of GITC is whole-body plethysmography in combination with spirometry, also known as Pulmonary Function Tests (PFT) [28]. However, issues with standardisation, effort-dependence and functional interpretation make this a suboptimal measure for assessment of GITC [29, 35, 36, 68]. Moreover, sensitivity of PFT is low, and therefore only able to point out very severe cases of GITC [34, 68]. Physicians strive to intervene prior to severe escalation of GITC in order to keep patient discomfort to a minimum, therefore they resort to more sensitive but less reliable measures such as SRS and RP for assessment of GITC in clinical practice.

Another new and promising approach which uses CT-imaging and simplified computational fluid dynamics was introduced by Hoeben *et al.* (chapter 2). With this approach CT-derived anatomical data (basically RP) are translated into functional data that express the Maximal Ventilation Capacity (MVC) of the trachea. The MVC of the trachea in a patient with GITC ($MVC_{patient}$) is then expressed as the percentage of this individual's predicted normal MVC. This individual specific normal values is based on reference diameter (d_{ref}) observed in unaffected trachea areas.

This d_{ref} is extrapolated over the rest of the course of the trachea, to generate a prediction of healthy trachea dimensions, i. e. the presumed dimensions of the trachea if the patient would not have developed a goitre. Thereafter, the actual trachea dimensions are used to predict the magnitude of flow the patient is able to generate in this diseased state ($\Phi -_{disease}$). Next, the predicted healthy trachea dimensions are used to predict the magnitude of flow the patient would be able to generate in the healthy state ($\Phi -_{disease}$). The ratio between

these levels of flow are assumed to reflect the magnitude of ventilation impairment imposed by the GITC and is defined as the Maximal Ventilation Capacity (MVC) according to equation 34:

$$MVC = \frac{\Phi_{disease}}{\Phi_{healthy}} \times 100\% \tag{34}$$

This prediction of a patients percentage ventilation reduction based on pre-surgery data is already a major leap forward in assessment of GITC, since it is an objective and patient-specific measure. However, for an optimal understanding of the potential and expected benefits of intervention, it is important to translate such results into practical consequences which can easily be interpreted by patients and physicians. For instance, will a MVC of 42% only prohibit a patient from competing in high-intensity sports, or will this also effect his ability to do the dishes?
This chapter aims to define the interpretation the MVC and provide a hypothetical example.

8.1 MODEL

To define the interpretation of the MVC, it is important to establish the patients ventilatory requirements to perform certain activities. The Consolidated Human Activity Database (CHAD) contains ventilatory data for nearly 23.000 person-days for a wide variety of daily activities, ages and both genders, which makes it highly suitable for modelling purposes [92]. These ventilatory data in the CHAD are expressed in V_e, which represents minute ventilation in L min^{-1}. For the present model it was assumed that flow is proportional to ventilation (motivation in chapter 2). The CHAD shows that for all activity levels, ventilation in men is consequently higher compared to ventilation in women at the same age and depends on a persons physical size. Generally maximal ventilatory values decrease with rising age [69]. Thus the percentage of maximal ventilation used for certain activities is influenced by age, built and sex.
Thus, with these data, it is possible to predict which percentage of healthy ventilation a person requires, to perform certain activities based on his age, weight and gender. To quantify these predicted consequences, 5 classes were established. These classes define the predicted consequences of GITC based on calculated MVC-values and are given in table 11. A more user-friendly interpretation of the relationship between the MVC and predicted GITC-class is provided in figure 49.

MVC	Class	Predicted consequence(s)
$> a$	I	None
$a \leftrightarrow b$	II	Heavy exercise impaired: sprinting
$b \leftrightarrow c$	III	Mild exercise impaired: riding a bike
$c \leftrightarrow d$	IV	Light exercise impaired: playing golf, cleaning house
$< d$	V	Everyday activities impaired: talking, walking

Table 11: GITC-classes based on MVC and predicted consequences. Class boundaries (a, b, c and d) differ for gender and age and are provided in table 12.

	Male				Female			
Age (years)	a	b	c	d	a	b	c	d
21 to 30	41	22	10	3	46	24	11	4
31 to 40	45	25	11	5	41	21	10	4
41 to 50	51	29	13	5	54	27	13	5
51 to 60	53	29	13	5	67	36	17	7
61 to 70	56	31	14	7	63	34	17	8
71 to 80	71	39	19	9	76	40	21	9

Table 12: Parameters indicating class boundaries using ventilatory rates for several activity levels derived from the CHAD and maximal ventilation rates from Blackie *et al.* a equals an activity level over 6 Metabolic Equivalent of Task (MET), b equals 3.0 to 6.0 METS, c equals 1.5 to 3.0 METS and d under 1.5 METS. Data are based on an individual with a body weight of 70 kg. Note that age restrictions in data from Blackie *et al.* and the CHAD do not completly overlap.

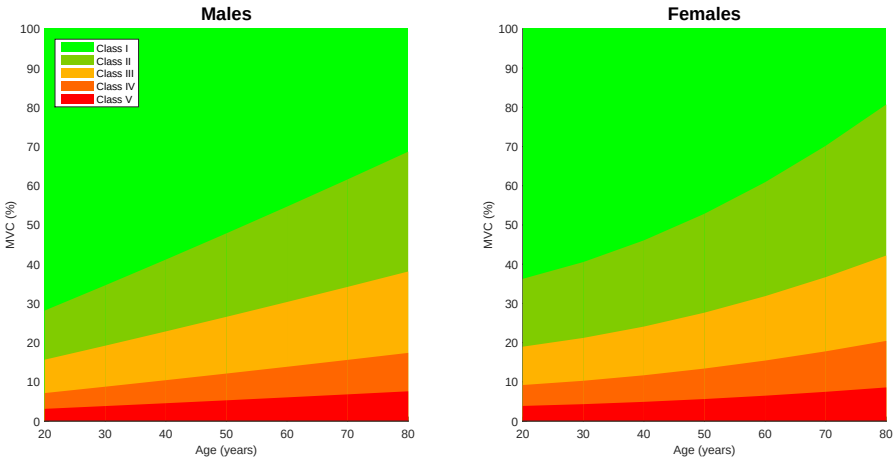


Figure 49: Classification of GITC based on MVC, gender and age: a visual interpretation of table 11 and 12. Data are based on a individual with a body weight of 70 kg.

8.1.1 *Hypothetical* *example*

Consider a 75 year old female of 70 kg with a MVC of 25%. Based on this MVC it is expected that she is able to perform everyday activities such as talking and walking, but her ability to perform light exercise was impaired, thus she would have trouble cleaning her house. If this person were to be operated, and her MVC would increase to for instance 65%. She would regain her ability to do the dishes and ride her bike to the supermarket, but she would not be able to sprint to catch a bus[1].

1 As far as that is possible for someone her age

9 IMPLEMENTATION

This chapter concludes this thesis by giving a brief point by point summary of all preceding chapters, thereby addressing their findings. Based on the conclusions of these chapters, a recommendation will be generated for additional research on the assessment of Goitre Induced Trachea Compression (GITC) in clinical practice.

9.1 SUMMARY

Chapter 1: Background

Goitre is a palpable enlargement of the thyroid which can compress the trachea, thereby raising airway resistance and limiting ventilation impairment. Several methods for the assessment of GITC are available, each with different efficacies and risks. To identify the most suited treatment, it is important to assess the magnitude of GITC accurately.

The golden standard for assessment of GITC are Pulmonary Function Tests (PFT), including the assessment of Flow Volume Loops (FVL) and plethysmography. However, many patients have difficulty performing the manoeuvres necessary for FVL. Furthermore, interpretation of these FVL relies on visual inspection and is therefore subjective. An alternative is the assessment of Pulmonary Function Values, which are objective predefined indices for the interpretation of FVL. However, GITC mainly hinders inspiration, making expiratory PFV not sensitive for the assessment of GITC. In addition to this, reference values for inspiratory PFV are lacking, making it difficult to interpret these values. Added to this, plethysmography is not standardised, making it difficult to compare the results of different studies. Furthermore, plethysmography assesses airway resistance in general, making it not specific for GITC. Another option is the assessment of Self Reported Symptoms (SRS). However, these SRS are very subjective. Furthermore, due the slow and progressive nature of the disease, patients are often not aware of the limitations imposed by GITC.

Assessment of Radiological Parameters (RP) is also common for the assessment of GITC. However, these parameters are often documented in rather subjective terms. Moreover, there is a significant lack of knowledge regarding the interpretation of these parameters.

Computational Fluid Dynamics (CFD) can simulate the practical consequences of RP, thereby providing a useful for the assessment of GITC. Detailed analysis of individual patients using these CFD can be very time-consuming. Simplification of these methods is possible, but this might comprise their reliability. Nevertheless, exploring the applicability of such a simplified CFD-method might be a worthwhile endeavour.

Chapter 2: Pilot study

A novel measure, which aims to quantify percentage ventilation impairment, was established and tested on a group of 10 patients suffering from goitre and 10 controls. This measure uses diseased trachea dimensions, and a prediction of presumed healthy trachea geometry to simulate the consequences of GITC. These trachea dimensions were obtained using a segmentation method which used a radio density threshold to segment the trachea lumen in transversal CT-images. These segmented slices were used to reconstruct a 3d-model of the trachea. Next a Centre Lumen Line (CLL) throughout the course of the trachea. Trachea dimensions orthogonal to this CLL were assumed to be more relevant in term of fluid mechanics because they were orthogonal to the mean direction of airflow.
A reference point near the vocal cords was used to predict healthy trachea dimensions. It was assumed that trachea dimensions near this landmark were not affected by GITC and thus a good indication for healthy trachea dimensions. These predicted healthy trachea dimensions and the actual pre-surgery trachea geometry were used to simulate the magnitude of peak flow through the trachea using simplified fluid mechanics. The ratio between the flow in the diseased state and healthy state was believed to represent the remaining ventilation capacity of the patient and was defined as the Maximal Ventilation Capacity.

Chapter 3: Retrospective analysis

Using the segmentation method devised during the pilot study, the trachea dimensions of a larger cohort of 74 patients and 55 control subjects were quantified. Hereby Minimal Cross-Sectional Area (CSA_m), Percentage Trachea Constriction (%-TC) and the MVC were quantified for all subjects. These analysis generated reference values for healthy patients for all three measures. Additionally these measures were compared to SRS. Despite the suboptimal quality of these SRS, a significant correlation between SRS and %-TC was identified.

Additionally, a portion of pre-surgery patients were diagnosed with GITC based on the reference values of the CSA_m, MVC and %-TC. During this analysis, %-TC was able to diagnose the largest percentage of pre-surgery patients.

Chapter 4: Prospective exploration

The suboptimal quality of SRS in retrospective data and lack of data for additional methods such as PFT indicated the need for a prospective study. First an exploratory analysis was performed to identify potential opportunities and issues of such a prospective study. This included the detailed analysis of a single patient. Hereby, SRS, RP and LFT were assessed. For the assessment of SRS, a novel questionnaire was devised, which was aimed specifically at the assessment of several aspects of GITC.
PFT included commonly used methods such as PFV and FVL, as well as potential novel indices, such as the assessment of frequency specific resistance, fundamental frequency and flow dependent resistance using Impulse Oscillometry Spirometry (IOS). Preliminary results from the assessment of these measures identify flow dependent resistance as the most promising, due to the large inter-individual variation of normal values for fundamental frequency and frequency specific resistance.
Additionally RP, were assessed. During this analysis an issue with the automated segmentation procedure was identified, whereas the procedure failed to quantify trachea dimensions in geometrically complex portions of the trachea. As an alternative, dimensions were assessed in the transversal plane, rather than orthogonal to the CLL as described in chapter 2. All of the assessed RP (CSA_m, %-TC and MVC) fell outside of the normal ranges defined in chapter 3, and returned to normal after surgery.

Chapter 5: Assessment of the segmentation procedure

To verify whether the issue with the CLL-method identified in chapter 4 was an isolated case, or whether this was is a recurring phenomenon was assessed by quantifying the dimensions of a predefined landmark (CSA_m) in a collection CT-scans of patients with varying degrees of GITC using the CLL- and transversal-method. This transversal-method assessed trachea dimensions in isolated segmented trachea slices, rather than a 3d-reconstruction of the trachea geometry as in the CLL-method. The relevance of this analysis was two-sided. On the one hand, it aimed to quantify the prevalence of technical issues of the CLL-method, and on the other hand it aimed to verify whether there

is an actual added value in more computationally complex methods such as the CLL-method.
Failure of the CLL-method, as reported in chapter 4 was observed in 7 other cases, which was less than 10% of the total population. These cases could easily be identified through visual inspection and comparison to the transversal-method.
Correlation between both methods was good, however, for some cases substantial differences were observed (60 mm^2). Using the relationship between Forced Inspiratory Flow at 50% of the vital capacity ($FIF_{50\%}$) and CSA_m documented in literature as a guideline, it was concluded that such differences could have substantial consequences for the assessment of GITC. Therefore it was indicated that the assessment of RP during further analysis should be done using both methods, to determine which of them yields the most relevant results.

Chapter 6: Automated versus manual assessment

To verify whether manual assessment of RP is sufficient for an effective assessment of GITC in clinical practice, it is important to quantify the accuracy and reliability of this method. Furthermore, for the interpretation of results obtained using this manual method, it should be known, how its results relate to the automated quantification of trachea dimensions.
Radiologist were provided with a protocol for assessment of minimal trachea diameter, based on the assumption that this will yield better inter-observer agreement and reproducibly compared the radiologist's personal interpretation of this parameter. The algorithm for the automated assessment of these parameters was modified to replicate the methods described in the protocol. Measurements were performed twice to enable assessment of intra-individual agreement.
Results showed that the magnitude of inter- and intra-individual variability can amount to several millimetres. Generally, radiologist were not consistent in the slice-selection in which they assessed these diameters. However, even when similar slices were selected, inter- and intra-individual variability amounted to several millimetres. Additionally, the automated method generated a consistent under estimation of these dimensions.
This provided insight in the accuracy of manual assessment of trachea dimensions, nevertheless its results are not conclusive regarding the relevance of this method in the assessment of GITC. For instance, it is unknown whether assessment of CSA instead of diameter will provide similar levels of unreliability. Furthermore,the significance of RP for the assessment of GITC has not been firmly established yet, i. e. , the practical consequences observed errors are

unknown. Therefore additional research regarding the manual assessment of trachea diameters is necessary, as well as further prospective analysis to establish the relevance of RP in the assessment of GITC.

Chapter 7: Prospective analysis

Following the exploratory study in described in chapter 4, a follow-up was devised whereby additional patients were included. At the current moment 12 patients suffering from goitre were included. Methods for the assessment of GITC in these patients were devised using the recommendations put forth in chapter 4. This included the assessment of SRS using a questionnaire, $\frac{\Delta R}{\Delta V}$ by Impulse Oscillometry Spirometry, PIF by traditional spirometry and various RP using an automated segmentation tool.
Generally, no significant correlations were found between the assessed methods. However, the presumed correlation between most of these measures are non-linear, whilst they were assessed using a linear fit. Therefore it is not remarkable that these data did not produce any significant results yet. Furthermore, number of included patients were quite low. It is expected that the inclusion of additional patients and employing more relevant fits will provide more insights in the relevance of the variously assessed methods for the management of GITC.

Chapter 8: Implementation of the MVC

It is recognised that providing standardised methods for the assessment of RP can already aid the assessment GITC using these RP. However, for an optimal understanding of the potential and expected benefits of intervention, it is important to translate such results into practical consequences which can easily be interpreted by patients and physicians. This chapter illustrates how the MVC can be used to this extent.

9.2 RECOMMENDATIONS

Various studies suggested that radiological parameters can not be successfully linked to ventilation impairment due to goitre. However, preliminary results of this study suggest that these radiological parameters can be useful in the assessment of GITC. Nevertheless, one should critically evaluate the manner in which such analysis are performed. Results indicated that manual assessment of RP showed significant intra- and inter-observer variability. Automated assessment of these RP shows superior reliability, however a superior validity could not be proven. Comparing these results to a

golden standard such as a phantom with known dimensions or validated software is therefore highly recommended.
Additionally, it seems that commonly used RP could be improved upon in various areas, such as defining a reference diameter in the cranial portion of the trachea rather than caudally, assessing dimensions orthogonal to the morphological skeleton of the trachea rather than the traditional transversal plane and using a circular equivalent diameter rather than cross-sectional area. However, the prospective study initiated to verify these assumptions could not generate conclusive results regarding these theories just yet. This is probably caused by an insufficient sample size.
Based on current results it is not yet possible whether the MVC is a more valuable tool than other radiological parameters such as minimal cross-sectional area and percentage trachea constriction. Nevertheless, the MVC has a distinct advantage over these other radiological parameters because it can predict patient specific consequences imposed by the GITC regardless of concomitant diseases.
Additionally, it is important to realise that the MVC in its current form is a proper simplification of real-life fluid dynamics. Thus, even if it turns out that the MVC in its current form is unable to accurately predict the consequences of GITC, its concept, of predicting the patient-specific consequences of GITC, might be used for an improved version whereby the various assumptions made in the establishment of the MVC are re-evaluated and modified if necessary.
Other measures, such as pulmonary function testing, also show a lack of knowledge regarding their patient-specific interpretation. Nevertheless, these techniques do show some promising new measures, such as $\frac{\Delta R}{\Delta \dot{V}}$, which have not been tested to their full extent yet. The relevance of these tools, for the assessment of GITC, should therefore further be explored.

BIBLIOGRAPHY

[1] David S. Cooper, Gerard M. Doherty, Bryan R. Haugen, Bryan R. Hauger, Richard T. Kloos, Stephanie L. Lee, Susan J. Mandel, Ernest L. Mazzaferri, Bryan McIver, Furio Pacini, Martin Schlumberger, Steven I. Sherman, David L. Steward, and R Michael Tuttle. Revised american thyroid association management guidelines for patients with thyroid nodules and differentiated thyroid cancer. *Thyroid*, 19(11):1167–1214, Nov 2009. doi: 10.1089/thy.2009.0110. URL http://dx.doi.org/10.1089/thy.2009.0110.

[2] Chung-Yau Lo Johnathan Hubbard, William B Inabnet, editor. *Endocrine Surgery*. Springer, London, 2009.

[3] Sunil K. Menon, Varsha S. Jagtap, Vijaya Sarathi, Anurag R. Lila, Tushar R. Bandgar, Padmavathy S. Menon, and Nalini S. Shah. Prevalence of upper airway obstruction in patients with apparently asymptomatic euthyroid multi nodular goitre. *Indian J Endocrinol Metab*, 15(Suppl 2):S127–S131, Jul 2011. doi: 10.4103/2230-8210.83351. URL http://dx.doi.org/10.4103/2230-8210.83351.

[4] B. Nygaard, L. Hegedüs, M. Gervil, H. Hjalgrim, P. Søe-Jensen, and J. M. Hansen. Radioiodine treatment of multinodular nontoxic goitre. *BMJ*, 307(6908):828–832, Oct 1993.

[5] M. Dietlein, B. Dederichs, C. Kobe, P. Theissen, M. Schmidt, and H. Schicha. [therapy for non-toxic multinodular goiter: radioiodine therapy as attractive alternative to surgery]. *Nuklearmedizin*, 45(1):21–34; quiz N1–2, 2006. doi: 10.1267/NUKL06010021. URL http://dx.doi.org/10.1267/NUKL06010021.

[6] Steen J. Bonnema, Viveque E. Nielsen, and Laszlo Hegedüs. Radioiodine therapy in non-toxic multinodular goitre. the possibility of effect-amplification with recombinant human tsh (rhtsh). *Acta Oncol*, 45(8):1051–1058, 2006. doi: 10.1080/02841860600617084. URL http://dx.doi.org/10.1080/02841860600617084.

[7] Søren Fast, Viveque Egsgaard Nielsen, Steen Joop Bonnema, and Laszlo Hegedüs. Time to reconsider nonsurgical therapy of benign non-toxic multinodular goitre: focus on recombinant human tsh augmented radioiodine therapy. *Eur J Endocrinol*, 160(4):517–528, Apr 2009. doi: 10.1530/EJE-08-0779. URL http://dx.doi.org/10.1530/EJE-08-0779.

[8] Woo Kyoung Jeong, Jung Hwan Baek, Hyunchul Rhim, Yoon Suk Kim, Min Sook Kwak, Hyun Jo Jeong, and Ducky Lee. Radiofrequency ablation of benign thyroid nodules: safety and imaging follow-up in 236 patients. *Eur Radiol*, 18(6):1244–1250, Jun 2008. doi: 10.1007/s00330-008-0880-6. URL http://dx.doi.org/10.1007/s00330-008-0880-6.

[9] R. Lencioni, D. Cioni, and C. Bartolozzi. Percutaneous radiofrequency thermal ablation of liver malignancies: techniques, indications, imaging findings, and clinical results. *Abdom Imaging*, 26(4):345–360, 2001.

[10] Jung Hwan Baek, Yoon Suk Kim, Ducky Lee, Jung Yin Huh, and Jeong Hyun Lee. Benign predominantly solid thyroid nodules: prospective study of efficacy of sonographically guided radiofrequency ablation versus control condition. *AJR Am J Roentgenol*, 194(4):1137–1142, Apr 2010. doi: 10.2214/AJR.09.3372. URL http://dx.doi.org/10.2214/AJR.09.3372.

[11] Jung Hwan Baek, Won-Jin Moon, Yoon Suk Kim, Jeong Hyun Lee, and Ducky Lee. Radiofrequency ablation for the treatment of autonomously functioning thyroid nodules. *World J Surg*, 33(9):1971–1977, Sep 2009. doi: 10.1007/s00268-009-0130-3. URL http://dx.doi.org/10.1007/s00268-009-0130-3.

[12] Maurilio Deandrea, Paolo Limone, Edoardo Basso, Alberto Mormile, Federico Ragazzoni, Elena Gamarra, Stefano Spiezia, Antongiulio Faggiano, Annamaria Colao, Filippo Molinari, and Roberto Garberoglio. Us-guided percutaneous radiofrequency thermal ablation for the treatment of solid benign hyperfunctioning or compressive thyroid nodules. *Ultrasound Med Biol*, 34(5):784–791, May 2008. doi: 10.1016/j.ultrasmedbio.2007.10.018. URL http://dx.doi.org/10.1016/j.ultrasmedbio.2007.10.018.

[13] Stefano Spiezia, Roberto Garberoglio, Francesco Milone, Valeria Ramundo, Corrado Caiazzo, Angelo Pio Assanti, Maurilio Deandrea, Paolo P. Limone, Paolo E. Macchia, Gaetano Lombardi, Annamaria Colao, and Antongiulio Faggiano. Thyroid nodules and related symptoms are stably controlled two years after radiofrequency thermal ablation. *Thyroid*, 19(3):219–225, Mar 2009. doi: 10.1089/thy.2008.0202. URL http://dx.doi.org/10.1089/thy.2008.0202.

[14] Young-Sun Kim, Hyunchul Rhim, Kyung Tae, Dong Woo Park, and Sung Tae Kim. Radiofrequency ablation of benign cold thyroid nodules: initial clinical experience. *Thyroid*, 16(4):361–367, Apr 2006. doi: 10.1089/thy.2006.16.361. URL http://dx.doi.org/10.1089/thy.2006.16.361.

[15] R. Le Moli, M. F. Wesche, M. M. Tiel-Van Buul, and W. M. Wiersinga. Determinants of longterm outcome of radioiodine therapy of sporadic non-toxic goitre. *Clin Endocrinol (Oxf)*, 50 (6):783–789, Jun 1999.

[16] Jung Hwan Baek, Jeong Hyun Lee, Roberto Valcavi, Claudio M. Pacella, Hyunchul Rhim, and Dong Gyu Na. Thermal ablation for benign thyroid nodules: radiofrequency and laser. *Korean J Radiol*, 12(5):525–540, 2011. doi: 10.3348/kjr.2011.12.5.525. URL http://dx.doi.org/10.3348/kjr.2011.12.5.525.

[17] O. Thomusch, C. Sekulla, and H. Dralle. [is primary total thyroidectomy justified in benign multinodular goiter? results of a prospective quality assurance study of 45 hospitals offering different levels of care]. *Chirurg*, 74(5):437–443, May 2003. doi: 10.1007/s00104-002-0605-3. URL http://dx.doi.org/10.1007/s00104-002-0605-3.

[18] Won-Kyeong Kim, Ji Hoon Shin, Jin Hyoung Kim, Jin Woo Song, and Ho-Young Song. Management of tracheal obstruction caused by benign or malignant thyroid disease using covered retrievable self-expandable nitinol stents. *Acta Radiol*, 51 (7):768–774, Sep 2010. doi: 10.3109/02841851.2010.491093. URL http://dx.doi.org/10.3109/02841851.2010.491093.

[19] James E. Wiseman, Matthew Mossanen, Philip H G. Ituarte, Jonathan M T. Bath, and Michael W. Yeh. An algorithm informed by the parathyroid hormone level reduces hypocalcemic complications of thyroidectomy. *World J Surg*, 34(3):532–537, Mar 2010. doi: 10.1007/s00268-009-0348-0. URL http://dx.doi.org/10.1007/s00268-009-0348-0.

[20] Gregory W. Randolph, Jennifer J. Shin, Hermes C. Grillo, Doug Mathisen, Mark R. Katlic, Dipti Kamani, and David Zurakowski. The surgical management of goiter: Part ii. surgical treatment and results. *Laryngoscope*, 121(1):68–76, Jan 2011. doi: 10.1002/lary.21091. URL http://dx.doi.org/10.1002/lary.21091.

[21] Jennifer J. Shin, Hermes C. Grillo, Doug Mathisen, Mark R. Katlic, David Zurakowski, Dipti Kamani, and Gregory W. Randolph. The surgical management of goiter: Part i. preoperative evaluation. *Laryngoscope*, 121(1):60–67, Jan 2011. doi: 10.1002/lary.21084. URL http://dx.doi.org/10.1002/lary.21084.

[22] A. R. Shaha. Surgery for benign thyroid disease causing tracheoesophageal compression. *Otolaryngol Clin North Am*, 23(3): 391–401, Jun 1990.

[23] Wen T. Shen, Electron Kebebew, Quan-Yang Duh, and Orlo H. Clark. Predictors of airway complications after thyroidectomy for substernal goiter. *Arch Surg*, 139(6):656–9; discussion 659–60, Jun 2004. doi: 10.1001/archsurg.139.6.656. URL http://dx.doi.org/10.1001/archsurg.139.6.656.

[24] Michael T. Stang, Michaele J. Armstrong, Jennifer B. Ogilvie, Linwah Yip, Kelly L. McCoy, Christopher N. Faber, and Sally E. Carty. Positional dyspnea and tracheal compression as indications for goiter resection. *Arch Surg*, 147(7):621–626, Jul 2012. doi: 10.1001/archsurg.2012.96. URL http://dx.doi.org/10.1001/archsurg.2012.96.

[25] Roberto Torchio, Carlo Gulotta, Alberto Perboni, Claudio Ciacco, Marco Guglielmo, Fabio Orlandi, and Joseph Milic-Emili. Orthopnea and tidal expiratory flow limitation in patients with euthyroid goiter. *Chest*, 124(1):133–140, Jul 2003.

[26] Amy Y. Chen, Victor J. Bernet, Sally E. Carty, Terry F. Davies, Ian Ganly, William B Inabnet, 3rd, Ashok R. Shaha, and Surgical Affairs Committee of the American Thyroid Association . American thyroid association statement on optimal surgical management of goiter. *Thyroid*, 24(2):181–189, Feb 2014.

[27] Steen J. Bonnema, Viveque E. Nielsen, and Laszlo Hegedüs. Long-term effects of radioiodine on thyroid function, size and patient satisfaction in non-toxic diffuse goitre. *Eur J Endocrinol*, 150(4):439–445, Apr 2004.

[28] Porntiva Poorisrisak, Carsten Vrang, Jorn Molgaard Henriksen, Bent Klug, Birgitte Hanel, and Hans Bisgaard. Accuracy of whole-body plethysmography requires biological calibration. *Chest*, 135(6):1476–1480, Jun 2009. doi: 10.1378/chest.08-1555. URL http://dx.doi.org/10.1378/chest.08-1555.

[29] W.T. Ulmer M.D. Goldman, H.J. Smith. Whole-body plethysmography. *European Respiratory Society Monograph*, 31:15–43, 2005. doi: 10.1183/1025448x.ERM3105.

[30] N. J. Gittoes, M. R. Miller, J. Daykin, M. C. Sheppard, and J. A. Franklyn. Upper airways obstruction in 153 consecutive patients presenting with thyroid enlargement. *BMJ*, 312(7029):484, Feb 1996.

[31] M. R. Miller, A. C. Pincock, G. D. Oates, R. Wilkinson, and H. Skene-Smith. Upper airway obstruction due to goitre: detection, prevalence and results of surgical management. *Q J Med*, 74(274):177–188, Feb 1990.

[32] Sylvia Verbanck, Tom de Keukeleire, Daniël Schuermans, Marc Meysman, Walter Vincken, and Bruce Thompson. Detecting upper airway obstruction in patients with tracheal stenosis. *J Appl Physiol (1985)*, 109(1):47–52, Jul 2010. doi: 10.1152/japplphysiol.01103.2009. URL http://dx.doi.org/10.1152/japplphysiol.01103.2009.

[33] W. W. Lunn and J. R. Sheller. Flow volume loops in the evaluation of upper airway obstruction. *Otolaryngol Clin North Am*, 28 (4):721–729, Aug 1995.

[34] M. Albareda, J. Viguera, C. Santiveri, P. Lozano, A. Mestrón, N. Bengoa, R. Calvet, A. Roger, D. Pardillo, E. Delgado, and Ll Vila. Upper airway obstruction in patients with endothoracic goiter enlargement: no relationship between flow-volume loops and radiological tests. *Eur J Endocrinol*, 163(4):665–669, Oct 2010. doi: 10.1530/EJE-10-0235. URL http://dx.doi.org/10.1530/EJE-10-0235.

[35] Giovanna Cantarella, Valter Fasano, Enrica Bucchioni, Elena Domenichini, and Bruno M. Cesana. Spirometric and plethysmographic assessment of upper airway obstruction in laryngeal hemiplegia. *Ann Otol Rhinol Laryngol*, 112(12):1014–1020, Dec 2003.

[36] J. G. Geraghty, E. C. Coveney, M. Kiernan, and N. J. O'Higgins. Flow volume loops in patients with goiters. *Ann Surg*, 215(1): 83–86, Jan 1992.

[37] R. D. Miller and R. E. Hyatt. Obstructing lesions of the larynx and trachea: clinical and physiologic characteristics. *Mayo Clin Proc*, 44(3):145–161, Mar 1969.

[38] R. D. Miller and R. E. Hyatt. Evaluation of obstructing lesions of the trachea and larynx by flow-volume loops. *Am Rev Respir Dis*, 108(3):475–481, Sep 1973.

[39] D. L. FRY and R. E. HYATT. Pulmonary mechanics. a unified analysis of the relationship between pressure, volume and gasflow in the lungs of normal and diseased human subjects. *Am J Med*, 29:672–689, Oct 1960.

[40] J. M. Bogaard, K. H. Pauw, and A. Versprille. Flow limitation in upper-airway obstruction. theoretical analysis. *ORL J Otorhinolaryngol Relat Spec*, 49(1):42–47, 1987.

[41] R. E. HYATT, D. P. SCHILDER, and D. L. FRY. Relationship between maximum expiratory flow and degree of lung inflation. *J Appl Physiol*, 13(3):331–336, Nov 1958.

[42] M. Kryger, F. Bode, R. Antic, and N. Anthonisen. Diagnosis of obstruction of the upper and central airways. *Am J Med*, 61(1): 85–93, Jul 1976.

[43] H. H. Rotman, H. P. Liss, and J. G. Weg. Diagnosis of upper airway obstruction by pulmonary function testing. *Chest*, 68(6): 796–799, Dec 1975.

[44] D. W. Empey. Assessment of upper airways obstruction. *Br Med J*, 3(5825):503–505, Aug 1972.

[45] J. Jordanoglou and N. B. Pride. A comparison of maximum inspiratory and expiratory flow in health and in lung disease. *Thorax*, 23(1):38–45, Jan 1968.

[46] J. I. Jensen, S. Lyager, and O. F. Pedersen. The relationship between maximal ventilation, breathing pattern and mechanical limitation of ventilation. *J Physiol*, 309:521–532, Dec 1980.

[47] R. Pellegrino, G. Viegi, V. Brusasco, R. O. Crapo, F. Burgos, R. Casaburi, A. Coates, C P M. van der Grinten, P. Gustafsson, J. Hankinson, R. Jensen, D. C. Johnson, N. MacIntyre, R. McKay, M. R. Miller, D. Navajas, O. F. Pedersen, and J. Wanger. Interpretative strategies for lung function tests. *Eur Respir J*, 26(5): 948–968, Nov 2005. doi: 10.1183/09031936.05.00035205. URL http://dx.doi.org/10.1183/09031936.05.00035205.

[48] Harre Bass. The flow volume loop: Normal standards and abnormalities in chronic obstructive pulmonary disease. *Chest*, 63(2):171–176, 1973. doi: 10.1378/chest.63.2.171. URL +http://dx.doi.org/10.1378/chest.63.2.171.

[49] T. Engel, J. H. Heinig, F. Madsen, and K. Nikander. Peak inspiratory flow and inspiratory vital capacity of patients with asthma measured with and without a new dry-powder inhaler device (turbuhaler). *Eur Respir J*, 3(9):1037–1041, Oct 1990.

[50] M.D. Goldman H.J. Smith, P. Reinhold. Forced oscillation technique and impulse oscillometry. *European Respiratory Society Monograph*, 31:72–105, 2005. doi: 10.1183/1025448x.ERM3105.

[51] M. D. Goldman. Clinical application of forced oscillation. *Pulm Pharmacol Ther*, 14(5):341–350, 2001. doi: 10.1006/pupt.2001.0310. URL http://dx.doi.org/10.1006/pupt.2001.0310.

[52] C. E. Cuijpers, G. J. Wesseling, A. G. Kessels, G. M. Swaen, P. L. Mertens, M. E. de Kok, J. Broer, F. Sturmans, and E. F. Wouters. Low diagnostic value of respiratory impedance measurements in children. *Eur Respir J*, 10(1):88–93, Jan 1997.

[53] C. Delacourt, H. Lorino, M. Herve-Guillot, P. Reinert, A. Harf, and B. Housset. Use of the forced oscillation technique to assess airway obstruction and reversibility in children. *Am J Respir Crit Care Med*, 161(3 Pt 1):730–736, Mar 2000. doi: 10.1164/ajrccm.161.3.9904081. URL http://dx.doi.org/10.1164/ajrccm.161.3.9904081.

[54] J. Hellinckx, K. De Boeck, J. Bande-Knops, M. van der Poel, and M. Demedts. Bronchodilator response in 3-6.5 years old healthy and stable asthmatic children. *Eur Respir J*, 12(2):438–443, Aug 1998.

[55] J. Hellinckx, K. De Boeck, and M. Demedts. No paradoxical bronchodilator response with forced oscillation technique in children with cystic fibrosis. *Chest*, 113(1):55–59, Jan 1998.

[56] B. Klug, K. G. Nielsen, and H. Bisgaard. Observer variability of lung function measurements in 2-6-yr-old children. *Eur Respir J*, 16(3):472–475, Sep 2000.

[57] H. J. Pennings and E. F. Wouters. Effect of inhaled beclomethasone dipropionate on isocapnic hyperventilation with cold air in asthmatics, measured with forced oscillation technique. *Eur Respir J*, 10(3):665–671, Mar 1997.

[58] J. A. Van Noord, J. Clément, K. P. Van de Woestijne, and M. Demedts. Total respiratory resistance and reactance in patients with asthma, chronic bronchitis, and emphysema. *Am Rev Respir Dis*, 143(5 Pt 1):922–927, May 1991. doi: 10.1164/ajrccm/143.5_Pt_1.922. URL http://dx.doi.org/10.1164/ajrccm/143.5_Pt_1.922.

[59] J. A. Van Noord, J. Smeets, J. Clément, K. P. Van de Woestijne, and M. Demedts. Assessment of reversibility of airflow obstruction. *Am J Respir Crit Care Med*, 150(2):551–554, Aug 1994. doi: 10.1164/ajrccm.150.2.8049845. URL http://dx.doi.org/10.1164/ajrccm.150.2.8049845.

[60] F. Zerah, A. M. Lorino, H. Lorino, A. Harf, and I. Macquin-Mavier. Forced oscillation technique vs spirometry to assess bronchodilatation in patients with asthma and copd. *Chest*, 108 (1):41–47, Jul 1995.

[61] A. R. Shaha, C. Burnett, A. Alfonso, and B. M. Jaffe. Goiters and airway problems. *Am J Surg*, 158(4):378–80; discussion 380–1, Oct 1989.

[62] A. H. Kendrick, J. F. O'Reilly, and G. Laszlo. Lung function and exercise performance in hyperthyroidism before and after treatment. *Q J Med*, 68(256):615–627, Aug 1988.

[63] Raniero Di Giovambattista. Hyperthyroidism as a reversible cause of right ventricular overload and congestive heart failure. *Cardiovasc Ultrasound*, 6:29, 2008. doi: 10.1186/1476-7120-6-29. URL http://dx.doi.org/10.1186/1476-7120-6-29.

[64] G. J. Kahaly, J. Nieswandt, S. Wagner, J. Schlegel, S. Mohr-Kahaly, and G. Hommel. Ineffective cardiorespiratory function in hyperthyroidism. *J Clin Endocrinol Metab*, 83(11):4075–4078, Nov 1998. doi: 10.1210/jcem.83.11.5275. URL http://dx.doi.org/10.1210/jcem.83.11.5275.

[65] Sailaja Vallabhajosula, Saba Radhi, Cihan Cevik, Raed Alalawi, Rishi Raj, and Kenneth Nugent. Hyperthyroidism and pulmonary hypertension: an important association. *Am J Med Sci*, 342(6):507–512, Dec 2011. doi: 10.1097/MAJ.0b013e31821790f4. URL http://dx.doi.org/10.1097/MAJ.0b013e31821790f4.

[66] S. M. Hashmi, D. J. Premachandra, A M D. Bennett, and W. Parry. Management of retrosternal goitres: results of early surgical intervention to prevent airway morbidity, and a review of the english literature. *J Laryngol Otol*, 120(8):644–649, Aug 2006. doi: 10.1017/S0022215106000995. URL http://dx.doi.org/10.1017/S0022215106000995.

[67] J. M. Rodriguez, Q. Hernandez, A. Piñero, S. Ortiz, T. Soria, P. Ramirez, and P. Parrilla. Substernal goiter: clinical experience of 72 cases. *Ann Otol Rhinol Laryngol*, 108(5):501–504, May 1999.

[68] Mark Brouns, Santhosh T. Jayaraju, Chris Lacor, Johan De Mey, Marc Noppen, Walter Vincken, and Sylvia Verbanck. Tracheal stenosis: a flow dynamics study. *J Appl Physiol (1985)*, 102 (3):1178–1184, Mar 2007. doi: 10.1152/japplphysiol.01063.2006. URL http://dx.doi.org/10.1152/japplphysiol.01063.2006.

[69] S. P. Blackie, M. S. Fairbarn, N. G. McElvaney, P. G. Wilcox, N. J. Morrison, and R. L. Pardy. Normal values and ranges for ventilation and breathing pattern at maximal exercise. *Chest*, 100(1):136–142, Jul 1991.

[70] Elias Greenbaum, editor. *Computational Fluid and Particle Dynamics in the Human Respiratory System*. Springer, 2013.

[71] V. Fasano, L. Raiteri, E. Bucchioni, S. Guerra, G. Cantarella, M. G. Massari, B. M. Cesana, and L. Allegra. Increased frequency dependence of specific airway resistance in patients with laryngeal hemiplegia. *Eur Respir J*, 18(6):1003–1008, Dec 2001.

[72] S. INGELSTEDT and N. G. TOREMALM. Air flow patterns and heat transfer within the respiratory tract. a new method

for experimental studies on models. *Acta Physiol Scand*, 51:204–217, 1961. doi: 10.1111/j.1748-1716.1961.tb02129.x. URL http://dx.doi.org/10.1111/j.1748-1716.1961.tb02129.x.

[73] N. T. Griscom and M. E. Wohl. Dimensions of the growing trachea related to age and gender. *AJR Am J Roentgenol*, 146 (2):233–237, Feb 1986. doi: 10.2214/ajr.146.2.233. URL http://dx.doi.org/10.2214/ajr.146.2.233.

[74] H. Nakano, K. Mishima, Y. Ueda, A. Matsushita, H. Suga, Y. Miyawaki, T. Mano, Y. Mori, and Y. Ueyama. A new method for determining the optimal ct threshold for extracting the upper airway. *Dentomaxillofac Radiol*, 42(3):26397438, 2013. doi: 10.1259/dmfr/26397438. URL http://dx.doi.org/10.1259/dmfr/26397438.

[75] J. P. Williamson, A. L. James, M. J. Phillips, D. D. Sampson, D. R. Hillman, and P. R. Eastwood. Quantifying tracheobronchial tree dimensions: methods, limitations and emerging techniques. *Eur Respir J*, 34(1):42–55, Jul 2009. doi: 10.1183/09031936.00020408. URL http://dx.doi.org/10.1183/09031936.00020408.

[76] D. Heyt. Pressure drop in flat-oval; spiral ducts. *ASHRAE Transactions*, 81(2), 1975.

[77] N. T. Griscom. Ct measurement of the tracheal lumen in children and adolescents. *AJR Am J Roentgenol*, 156(2):371–372, Feb 1991. doi: 10.2214/ajr.156.2.1898817. URL http://dx.doi.org/10.2214/ajr.156.2.1898817.

[78] E. Breatnach, G. C. Abbott, and R. G. Fraser. Dimensions of the normal human trachea. *AJR Am J Roentgenol*, 142(5):903–906, May 1984. doi: 10.2214/ajr.142.5.903. URL http://dx.doi.org/10.2214/ajr.142.5.903.

[79] A. Zapletal, E. K. Motoyama, K. P. Van De Woestijne, V. R. Hunt, and A. Bouhuys. Maximum expiratory flow-volume curves and airway conductance in children and adolescents. *J Appl Physiol*, 26(3):308–316, Mar 1969.

[80] Jennifer Welter, editor. *A Brief Introduction to Fluid Mechanics*, volume 1. Don Fowley, 2011.

[81] Myer Kutz, editor. *Biomedical Engineering and Design Handbook*, volume 1. The MCGraw-Hill Companies, 2009.

[82] T. E. Corcoran and N. Chigier. Characterization of the laryngeal jet using phase doppler interferometry. *J Aerosol Med*, 13(2):125–137, 2000. doi: 10.1089/089426800418659. URL http://dx.doi.org/10.1089/089426800418659.

[83] C. Kleinstreuer and Z. Zhang. Laminar-to-turbulent fluid-particle flows in a human airway model. *International Journal of Multiphase Flow*, 29(2):271 – 289, 2003. ISSN 0301-9322. doi: http://dx.doi.org/10.1016/S0301-9322(02)00131-3. URL http://www.sciencedirect.com/science/article/pii/S0301932202001313.

[84] P. H. Jarreau, B. Louis, G. Dassieu, L. Desfrere, P. W. Blanchard, G. Moriette, D. Isabey, and A. Harf. Estimation of inspiratory pressure drop in neonatal and pediatric endotracheal tubes. *J Appl Physiol (1985)*, 87(1):36–46, Jul 1999.

[85] A. S. Green. Modelling of peak-flow wall shear stress in major airways of the lung. *J Biomech*, 37(5):661–667, May 2004. doi: 10.1016/j.jbiomech.2003.09.024. URL http://dx.doi.org/10.1016/j.jbiomech.2003.09.024.

[86] Kristina Kairaitis, Manisha Verma, Jason Amatoury, John R. Wheatley, David P. White, and Terence C. Amis. A threshold lung volume for optimal mechanical effects on upper airway airflow dynamics: studies in an anesthetized rabbit model. *J Appl Physiol (1985)*, 112(7):1197–1205, Apr 2012. doi: 10.1152/japplphysiol.01286.2011. URL http://dx.doi.org/10.1152/japplphysiol.01286.2011.

[87] Malcolm J. McPherson, editor. *Subsurface Ventilation and Environmental Engineering*. Springer, 1993.

[88] C. F. Melissant, S. J. Smith, R. Perlberger, J. Verschakelen, J. W. Lammers, and M. Demedts. Lung function, ct-scan and x-ray in upper airway obstruction due to thyroid goitre. *Eur Respir J*, 7(10):1782–1787, Oct 1994.

[89] J. M. Fouke and K. P. Strohl. Effect of position and lung volume on upper airway geometry. *J Appl Physiol (1985)*, 63(1):375–380, Jul 1987.

[90] V. Hoffstein. Relationship between lung volume, maximal expiratory flow, forced expiratory volume in one second, and tracheal area in normal men and women. *Am Rev Respir Dis*, 134 (5):956–961, Nov 1986.

[91] S. J. Bonnema, H. Bertelsen, J. Mortensen, P. B. Andersen, D. U. Knudsen, L. Bastholt, and L. Hegedüs. 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. *J Clin Endocrinol Metab*, 84 (10):3636–3641, Oct 1999. doi: 10.1210/jcem.84.10.6052. URL http://dx.doi.org/10.1210/jcem.84.10.6052.

[92] T. McCurdy, G. Glen, L. Smith, and Y. Lakkadi. The national exposure research laboratory's consolidated human activity database. *J Expo Anal Environ Epidemiol*, 10(6 Pt 1):566–578, 2000.

[93] H.C. Grillo. *Surgery of the Trachea and Bronchi*. Pmph USA Ltd Series. BC Decker, 2004. ISBN 9781550090581. URL https://books.google.nl/books?id=8vODpl_05wYC.

[94] B. Tandler, J. M. Sherman, T. F. Boat, and R. E. Wood. Surface architecture of the mucosal epithelium of the cat trachea: Ii. structure and dynamics of the membranous portion. *Am J Anat*, 168(2):133–144, Oct 1983. doi: 10.1002/aja.1001680203. URL http://dx.doi.org/10.1002/aja.1001680203.

[95] Ann Bowling. Mode of questionnaire administration can have serious effects on data quality. *J Public Health (Oxf)*, 27(3):281–291, Sep 2005. doi: 10.1093/pubmed/fdi031. URL http://dx.doi.org/10.1093/pubmed/fdi031.

[96] Lois Ruth HARRIS and Gavin Thomas Lumsden BROWN. Mixing interview and questionnaire methods: Practical problems in aligning data. *Practical Assessment, Research & Evaluation*, 2010.

[97] Torquil Watt, Per Cramon, Laszlo Hegedüs, Jakob Bue Bjorner, Steen Joop Bonnema, Åse Krogh Rasmussen, Ulla Feldt-Rasmussen, and Mogens Groenvold. The thyroid-related quality of life measure thypro has good responsiveness and ability to detect relevant treatment effects. *J Clin Endocrinol Metab*, 99(10):3708–3717, Oct 2014. doi: 10.1210/jc.2014-1322. URL http://dx.doi.org/10.1210/jc.2014-1322.

[98] Torquil Watt, Giuseppe Barbesino, Jakob Bue Bjorner, Steen Joop Bonnema, Branka Bukvic, Russell Drummond, Mogens Groenvold, Laszlo Hegedüs, Valeska Kantzer, Kathryn E. Lasch, Claudio Marcocci, Anjali Mishra, Romana Netea-Maier, Merel Ekker, Ivan Paunovic, Terence J. Quinn, Åse Krogh Rasmussen, Audrey Russell, Mayilvaganan Sabaretnam, Johannes Smit, Ove Törring, Vladan Zivaljevic, and Ulla Feldt-Rasmussen. Cross-cultural validity of the thyroid-specific quality-of-life patient-reported outcome measure, thypro. *Qual Life Res*, 24(3):769–780, Mar 2015. doi: 10.1007/s11136-014-0798-1. URL http://dx.doi.org/10.1007/s11136-014-0798-1.

[99] Khaled Al-shair, Graham T W. Atherton, Deborah Kennedy, Georgina Powell, David W. Denning, and Ann Caress. Validity and reliability of the st. george's respiratory questionnaire in assessing health status in patients with chronic pulmonary aspergillosis. *Chest*, 144(2):623–631, Aug 2013. doi: 10.1378/chest.12-0014. URL http://dx.doi.org/10.1378/chest.12-0014.

[100] Y-J. Bae, Y. S. Kim, C-S. Park, Y. S. Lee, Y-S. Chang, Y. S. Cho, A-S. Jang, S-H. Cho, B. W. Choi, S-G. Kim, H-B. Moon, T-B. Kim, and C. O. R. E. A Study Group . Reliability and validity of the st george's respiratory questionnaire for asthma. *Int J Tuberc Lung Dis*, 15(7):966–971, Jul 2011. doi: 10.5588/ijtld.10.0603. URL http://dx.doi.org/10.5588/ijtld.10.0603.

[101] Jeffrey J. Swigris, Hye-Seung Lee, Marsha Cohen, Yoshikazu Inoue, Joel Moss, Lianne G. Singer, Lisa R. Young, Francis X. McCormack, National Institutes of Health Rare Lung Diseases Consortium , the Multicenter International Lymphangioleiomyomatosis Efficacy, and Safety of Sirolimus Trial Group. St. george's respiratory questionnaire has longitudinal construct validity in lymphangioleiomyomatosis. *Chest*, 143(6): 1671–1678, Jun 2013.

[102] David A. Kaminsky. What does airway resistance tell us about lung function? *Respir Care*, 57(1):85–96; discussion 96–9, Jan 2012. doi: 10.4187/respcare.01411. URL http://dx.doi.org/10.4187/respcare.01411.

[103] Hirsh D. Komarow, Ian A. Myles, Ashraf Uzzaman, and Dean D. Metcalfe. Impulse oscillometry in the evaluation of diseases of the airways in children. *Ann Allergy Asthma Immunol*, 106(3):191–199, Mar 2011. doi: 10.1016/j.anai.2010.11.011. URL http://dx.doi.org/10.1016/j.anai.2010.11.011.

[104] A. B. DUBOIS, A. W. BRODY, D. H. LEWIS, and BF BURGESS, Jr. Oscillation mechanics of lungs and chest in man. *J Appl Physiol*, 8(6):587–594, May 1956.

[105] T. Horan, S. Mateus, P. Beraldo, L. Araújo, J. Urschel, E. Urmenyi, and F. Santiago. Forced oscillation technique to evaluate tracheostenosis in patients with neurologic injury. *Chest*, 120(1): 69–73, Jul 2001.

[106] N. B. Pride. Forced oscillation techniques for measuring mechanical properties of the respiratory system. *Thorax*, 47(4):317–320, Apr 1992.

[107] M. Cauberghs and K. P. Van de Woestijne. Changes of respiratory input impedance during breathing in humans. *J Appl Physiol (1985)*, 73(6):2355–2362, Dec 1992.

[108] J. Clément, F. J. Làndsér, and K. P. Van de Woestijne. Total resistance and reactance in patients with respiratory complaints with and without airways obstruction. *Chest*, 83(2):215–220, Feb 1983.

[109] Holger Schulz, Claudia Flexeder, Jürgen Behr, Margit Heier, Rolf Holle, Rudolf M. Huber, Rudolf A. Jörres, Dennis Nowak, Annette Peters, H-Erich Wichmann, Joachim Heinrich, Stefan Karrasch, and K. O. R. A Study Group . Reference values of impulse oscillometric lung function indices in adults of advanced age. *PLoS One*, 8(5):e63366, 2013.

[110] Scott Bickel, Jonathan Popler, Burton Lesnick, and Nemr Eid. Impulse oscillometry: interpretation and practical applications. *Chest*, 146(3):841–847, Sep 2014. doi: 10.1378/chest.13-1875. URL http://dx.doi.org/10.1378/chest.13-1875.

[111] F. J. Làndsér, J. Clément, and K. P. Van de Woestijne. Normal values of total respiratory resistance and reactance determined by forced oscillations: influence of smoking. *Chest*, 81(5):586–591, May 1982.

[112] A. M. Lorino, F. Zerah, C. Mariette, A. Harf, and H. Lorino. Respiratory resistive impedance in obstructive patients: linear regression analysis vs viscoelastic modelling. *Eur Respir J*, 10 (1):150–155, Jan 1997.

[113] B. Daróczy and Z. Hantos. Generation of optimum pseudo-random signals for respiratory impedance measurements. *Int J Biomed Comput*, 25(1):21–31, Feb 1990.

[114] A. Berghout, W. M. Wiersinga, N. J. Smits, and J. L. Touber. Interrelationships between age, thyroid volume, thyroid nodularity, and thyroid function in patients with sporadic nontoxic goiter. *Am J Med*, 89(5):602–608, Nov 1990.

[115] Barbara A. Jereczek-Fossa, Daniela Alterio, Jacek Jassem, Bianca Gibelli, Nicoletta Tradati, and Roberto Orecchia. Radiotherapy-induced thyroid disorders. *Cancer Treat Rev*, 30 (4):369–384, Jun 2004. doi: 10.1016/j.ctrv.2003.12.003. URL http://dx.doi.org/10.1016/j.ctrv.2003.12.003.

[116] George van der Watt, David Haarburger, and Peter Berman. Euthyroid patient with elevated serum free thyroxine. *Clin Chem*, 54(7):1239–1241, Jul 2008. doi: 10.1373/clinchem.2007.101428. URL http://dx.doi.org/10.1373/clinchem.2007.101428.

[117] Tao Le, editor. *First aid for the USMLE - Step 1*. McGrawHill Education, 2014.

[118] FW Ewald, Jr, M. F. Tenholder, and R. F. Waller. Analysis of the inspiratory flow-volume curve. should it always precede the forced expiratory maneuver? *Chest*, 106(3):814–818, Sep 1994.

[119] A. G. LIDDELOW and A. H. CAMPBELL. Widening of the membranous wall and flattening of the trachea and main bronchi. *Br J Dis Chest*, 58:56–60, Apr 1964.

[120] S. K. Ramlal, F. J. Visser, W C J. Hop, P N R. Dekhuijzen, and Y. F. Heijdra. Effects of long-acting bronchodilators and prednisolone on inspiratory lung function parameters in stable copd. *Pulm Pharmacol Ther*, 28(2):158–164, Aug 2014. doi: 10.1016/j.pupt.2013.09.002. URL http://dx.doi.org/10.1016/j.pupt.2013.09.002.

[121] A. Falconer, C. Oldman, and P. Helms. Poor agreement between reported and recorded nocturnal cough in asthma. *Pediatr Pulmonol*, 15(4):209–211, Apr 1993.

Part I

APPENDIX

	Patient											
	1	2	3	4	5	6	7	8	9	10	11	12
VC (L)	-	3.25	2.71	3.11	3.75	4.03	3.4	2.77	3.68	3.59	2.62	3.83
FVC_{in} (L)	-	2.52	2.57	2.58	3.63	3.96	3.00	2.58	3.29	2.97	2.4	3.68
FVC_{out} (L)	-	2.66	2.71	3.11	3.69	4.03	3.4	2.76	3.43	3.46	2.45	3.4
PIF (L s^{-1})	-	4.59	2.12	2.95	4.91	7.27	5.04	2.81	3.48	3.21	4.48	5.04
PEF (L s^{-1})	-	6.26	3.56	6.85	6.05	7.19	7.92	7.14	8.63	6.17	4.74	7.92
FEV_1 (L)	-	2.2	2.22	2.57	2.98	2.77	3.07	2.5	2.82	3.15	1.77	2.5
FIV_1 (L)	-	2.52	2.01	2.47	3.6	3.96	3.00	2.48	3.16	2.86	2.34	2.76
$\frac{FEV_1}{FVC}$	-	0.827	0.819	0.826	0.808	0.687	0.903	0.9058	0.822	0.910	0.722	0.879
$\frac{FIV_1}{FVC_{in}}$	-	1.000	0.782	0.957	1.000	1.000	1.000	0.9612	0.960	0.963	0.975	0.957
R 5 Hz ($\frac{kPa}{Ls^{-1}}$)	0.43	0.35	0.51	0.36	0.52	0.54	0.35	0.40	0.28	0.66	0.54	0.52
R 20 Hz ($\frac{kPa}{Ls^{-1}}$)	0.27	0.33	0.42	0.33	0.46	0.48	0.31	0.31	0.23	0.59	0.36	0.45
X 5 Hz	-0.12	-0.06	-0.10	-0.07	-0.14	-0.17	-0.12	-0.16	-0.10	-0.21	-0.19	-0.10
f_0 (Hz)	18.69	9.01	11.41	9.65	12.68	13.59	10.18	14.52	14.31	17.38	24.93	10.71
$\frac{\Delta R}{\Delta V}(\frac{kPa}{Ls^{-1}})$5Hz	0.021	0.219	0.219	0.073	0.193	0.089	0.185	0.222	0.196	0.177	0.154	0.188
$\frac{\Delta R}{\Delta V}(\frac{kPa}{Ls^{-1}})$20Hz	-0.008	0.128	0.154	0.052	0.191	0.024	0.142	0.132	0.103	0.137	0.058	0.146
CSA_m^c (mm^2)	31.03	73.62	78.16	80.87	74.54	167.36	105.86	106.23	133.74	35.14	70.44	149.03
%-TC	51.64	21.39	59.89	27.07	34.73	18.40	25.51	17.68	30.29	45.65	27.11	9.91
d_{min}	5.98	9.71	11.37	10.08	9.69	14.42	11.45	11.58	12.32	6.38	9.72	13.77
MVC	49.04	88.07	64.18	67.69	57.21	84.96	62.71	82.54	57.05	74.17	60.96	85.99

Table 13: Lung function results and radiological parameters for prospective patients. $MCSA_t$-values were obtaining using the transversal-method, whereas $MCSA_c$ were obtained using the CLL-method.

					Iteration											
	Measure	Unit(s)	Pred.	LLN	1	2	3	4	5	6	7	8	9	10	μ	σ
Pre-surgery	FEV_1	L	3.04	2.21						2.93	2.93	2.84	2.98	2.87	2.91	0.055
	VC_{max}	L	4.09	3.17			4.73	4.77	4.66						4.72	0.056
	FVC	L	3.94	2.94						3.91	3.72	3.68	3.73	3.74	3.76	0.089
	$\frac{FEV_1}{VC_{max}}$	%	75.0	63.2						61.4	61.5	59.5	62.6	60.2	61.0	1.21
	PEF	$L\,s^{-1}$	7.94	5.96						3.86	3.77	3.65	4.30	3.67	3.85	0.27
	FVC_{in}	L	4.09	3.17	2.32	1.53	0.98	2.91	2.84	4.05	2.98	2.88	2.43	2.78	2.57	0.84
	FET	s								4.09	4.17	4.00	4.83	4.09	4.24	0.34
Post-surgery	FEV_1	L	3.04	2.21	3.92		3.93	3.90	4.01						3.69	0.51
	VC_{max}	L	4.09	3.17	5.00	5.31	5.13								5.15	0.16
	FVC	L	3.94	2.94	4.05		5.49	5.42	5.68						5.16	0.75
	$\frac{FEV_1}{VC_{max}}$	%	75.0	63.2	68.9		69.1	68.6	70.6						69.3	0.89
	PEF	$L\,s^{-1}$	7.94	5.96	8.00		8.33	8.36	7.82						8.13	0.16
	FVC_{in}	L	4.09	3.17		3.73	5.04	4.97	4.85						4.35	0.85
	FET	s			1.47		12.50	10.74	11.73						9.11	5.14

Table 14: Pre- and post-surgery pulmonary function values of patient X

Vragenlijst: Struma

Hartelijk dank voor uw medewerking. Onderstaande vragenlijst helpt ons om te begrijpen welke invloed het struma heeft op uw dagelijks leven.

Leest u alstublieft de instructies zorgvuldig door een geef het aan wanneer u iets niet begrijpt. Denk niet te lang na over uw antwoorden. Let op, dit formulier heeft ook een achterzijde.

1. **Leeftijd:** ___ jaar
2. **Geslacht:** □ man □ vrouw
3. **Mijn gezondheid vind ik:**
   Zeer slecht □—□—□—□—□ Zeer goed
4. **Mijn ademhalings problemen begonnen ___ jaar geleden.**
5. **De ernst van mijn ademhalingsproblemen:**
   - □ blijft constant
   - □ neemt toe
   - □ neemt af
   - □ ik heb geen ademhalingsproblemen
6. **Hoestbuien heb ik:**
   - □ 0 - 5 keer per dag
   - □ 5 - 10 keer per dag
   - □ $>$ 10 keer per dag
7. **Problemen met slikken heb ik (meerdere antwoorden mogelijk):**
   - □ Nooit
   - □ Bij het eten van grote vaste stukken
   - □ Bij het eten van kleine vaste stukken
   - □ Bij het eten van pap
   - □ Bij drinken
8. **Een drukkend gevoel in de keel heb ik (meerdere antwoorden mogelijk):**
   - □ Nooit
   - □ Soms
   - □ Bij het heffen van mijn armen
   - □ Bij liggen op mijn linker zij
   - □ Bij liggen op mijn rechter zij
   - □ Altijd
9. **Een piepende ademhaling heb ik:**
   - □ Enkele keren per dag
   - □ Enkele keren per week
   - □ Enkele keren per maand
   - □ Nooit
10. **Heesheid heb ik:**
    - □ Enkele keren per dag
    - □ Enkele keren per week
    - □ Enkele keren per maand
    - □ Nooit
11. **Welke situaties zijn op u van toepasing? (meerdere antwoorden mogelijk)**
    - □ Door mijn ademhaling doe ik er lang over om mij te wassen of aan te kleden

- ☐ Door mijn ademhaling kan ik niet baden of douchen, of ik doe er erg lang over
- ☐ Door mijn ademhaling loop ik langzamer dan anderen, of ik stop om te rusten
- ☐ Door mijn ademhaling, loop ik langzaam een trap op of stop ik tussendoor om te rusten
- ☐ Door mijn ademhaling ga ik langzaam of ik stop tussendoor om te rusten als ik mij moet haasten of snel wil lopen
- ☐ Mijn ademhaling maakt het moeilijk om dingen te doen als objecten een trap optillen, licht werk in de tuin, dansen, bowlen of golfen
- ☐ Mijn ademhaling maakt het moeilijk om dingen te doen als zware dingen tillen, zwaar werk in de tuin, lopen bij 8 km/uur, tennisen of zwemmen
- ☐ Mijn ademhaling maakt het moeilijk om dingen te doen als zware lichamelijk arbeid, rennen, fietsen, snel zwemmen of intensieve sporten beoefenen

12. Welke situaties zijn op u van toepasing? (meerdere antwoorden mogelijk)

- ☐ Ik schaam mij voor mijn ademhaling
- ☐ Mijn ademhaling zorgt voor overlast bij familie, vrienden en buren
- ☐ Ik ben bang of raak in paniek bij ademtekort
- ☐ Ik heb het gevoel dat ik mijn ademhalingsprobleem niet onder controle heb
- ☐ Ik verwacht niet dat mijn ademhaling beter zal worden
- ☐ Mijn ademhaling maakt mij zwak en hulpbehoevend
- ☐ Lichaamsbeweging is niet veilig voor mij
- ☐ Alles lijkt teveel moeite te kosten

13. In welke mate ervaart u kortademigheid op een schaal van 0-10?

- ☐ 0 In het geheel geen
- ☐ 0.5 Juist merkbaar
- ☐ 1 Zeer gering
- ☐ 2 Gering
- ☐ 3 Licht tot matig
- ☐ 4 Matig hevig
- ☐ 5 Hevig
- ☐ 6 -
- ☐ 7 Zeer hevig
- ☐ 8 -
- ☐ 9 Zeer hevig/bijna maximaal
- ☐ 10 Maximaal

14. Bent u bekend met hart- en/of longaandoeningen?

- ☐ Nee
- ☐ Ja, namelijk: _____

Einde van de vragenlijst.

Questionaire: Goitre

Thank you very much for you cooperation. This questionnaire helps us to understand to what extent goitre influences your daily life.

Please read the instructions carefully and indicate if things are unclear. Do not think too long about your answers. Note that this form also has a backside.

1. **Age:** ___ year
2. **Gender:** □ male □ female
3. **My general health is:**
   Very bad □—□—□—□—□ very good
4. **My breathing problems started ___ years ago.**
5. **The severity of my breathing problems:**
   - □ remain constant
   - □ are increasing
   - □ are decreasing
   - □ I do not have any breathing problems
6. **I have to cough:**
   - □ 0 - 5 times a day
   - □ 5 - 10 times a day
   - □ $>$ 10 times a day
7. **I expierence problems with swallowing (multiple answers possible):**
   - □ Never
   - □ Whilst eating bulky solid pieces
   - □ Whilst eating small solid pieces
   - □ Whilst eating porridge
   - □ Whilst drinking
8. **I expierence a globus sensation (multiple answers possible):**
   - □ Never
   - □ Sometimes
   - □ When raising my arms
   - □ Lying on my left side
   - □ Lying on my right side
   - □ Always
9. **I expierence wheezing:**
   - □ Several times a day
   - □ Several times a week
   - □ Several times a month
   - □ Never
10. **I expierence hoarseness:**
    - □ Several times a day
    - □ Several times a week
    - □ Several times a month
    - □ Never
11. **Which situations apply to you? (multiple answers possible)**
    - □ Because of my breathing problems I take longer to wash myself or to get dressed

- ☐ Because of my breathing problems I can't bathe or shower, or I take a very long time
- ☐ Because of my breathing problems I walk slower than others, or I take periods of rest
- ☐ Because of my breathing problems I slow down or take periods of rest when I am in a hurry of want to walk quickly
- ☐ Because of my breathing problems I experience troubles when carrying objects up a flight of stairs, during light garden work, dancing, bowling or golfing
- ☐ Because of my breathing problems I experience troubles when carrying heavy objects up a flight of stairs, during heavy garden work, walking at 8 km/h, tennis or swimming
- ☐ Because of my breathing problems I experience troubles when performing heavy psychical labour, running, riding a bike, swimming at a hight pace or performing intensive sports

12. Which situations apply to you? (multiple answers possible)

- ☐ I am ashamed of my breathing
- ☐ My breathing is a nuisance for family, friends and neighbours
- ☐ I am afraid or I panic during shortness of breath
- ☐ I feel that I do not control my breathing problem
- ☐ I do not expect my breathing to get better
- ☐ My breathing makes me weak and helpless
- ☐ Physical exercise is not safe for me
- ☐ Everything seems to cost to much effort

13. To what extent do you experience shortness of breath on a scale from 0-10?

- ☐ 0 Not at all
- ☐ 0.5 Barely noticable
- ☐ 1 very little
- ☐ 2 little
- ☐ 3 Light to moderate
- ☐ 4 Moderate to severe
- ☐ 5 Severe
- ☐ 6 -
- ☐ 7 Very severe
- ☐ 8 -
- ☐ 9 Very severe/almost maximal
- ☐ 10 Maximal

14. Do you have any cardiovacular or pulmonary diseases?

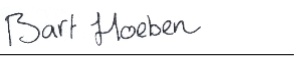
- ☐ No
- ☐ Yes, namely: _____

End of this questionnaire.

DECLARATION

I, Bart Wilhelmus Hoeben, declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own research.

Arnhem, November 6, 2015



Bart Hoeben, November 6, 2015

COLOPHON

This thesis has been realised by a collaboration between the departments of internal medicine, surgery and radiology at the Rijnstate Hospital group and the department of Technical Mechine at the University of Twente.