RELIABILITY OF UPPER PHARYNGEAL AIRWAY ASSESSMENT USING DENTAL CBCT

by

Jason Noah Zimmerman

DDS, University of Western Ontario, 2014

BSc (Honours), University of Western Ontario, 2009

A THESIS SUBMIITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Craniofacial Science)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

August 2017

Abstract

Introduction: Upper airway analysis is an often-cited use of CBCT imaging in orthodontics, however the reliability of airway measurements using this technology is not fully understood. The purpose of this study was to determine the intra-examiner and inter-examiner reliability of the complete process of volumetric and cross-sectional area assessments of the upper airway using CBCT imaging.

Methods: Five examiners of varying levels of education and clinical experience performed manual orientation, slice and threshold selection, and measured nasopharyngeal, oropharyngeal, hypopharyngeal, and total upper pharyngeal airway volumes in addition to minimum cross-sectional area on the CBCT images of 10 patients. All measurements were repeated after 4-weeks. Intra and inter-examiner reliability was calculated using ICC and 95% CI.

Results: Threshold selection showed poor intra and inter-examiner reliability, while minimum cross-sectional area showed moderate intra and poor inter-examiner reliability. Intra-examiner reliability of volumetric measurements varied based on the anatomical region assessed with ICC ranging from 0.747-0.976, and was worst for hypopharynx and best for the oropharynx. Inter-examiner reliability of volume measurements was generally lower, with ICC ranging from 0.175-0.945, and was worst for nasopharynx and best for the oropharynx.

Conclusions: This study, for the first time, assessed the reliability of upper airway analysis with CBCT when all steps of image processing and measurement are performed by each examiner. Reliability improved with examiner experience, though was generally low for the hypopharynx and nasopharynx volumes and overall minimal cross sectional area. The oropharyngeal volume was the only parameter to have excellent intra-examiner and inter-examiner reliability.

Lay Summary

Study Question: How reliable is CBCT at assessing the upper airway's volume and minimum cross-sectional area?

Background: CBCT three-dimensional imaging is often used to look at the volume of the upper airway. A systematic review conducted by the authors of this study found major methodological flaws in the literature. Most significantly the reliability was only assessed for the examiners' ability to trace the upper airway, with many steps of the measurement process not considered.

Methods: Five examiners positioned the CBCT images of ten patients and measured the volumes of the entire upper airway and its individual sections, as well as minimum cross-sectional area. The examiners selected the threshold sensitivity value for each scan. After 4-weeks, all measurements were repeated and reliability was calculated.

Key Results: Threshold overall had poor reliability. Reliability greatly improved with experience of the examiner, with oropharyngeal volume being the only part to have generalized excellent reliability.

Preface

This thesis is an original intellectual product of the author, J.N. Zimmerman. The study and associated methods were approved by the University of British Columbia's Research Ethics Board certificate H12-00951.

A version of Chapter 2 has been published (Zimmerman JN, Lee J, Pliska BT. Reliability of upper pharyngeal airway assessment using dental CBCT: a systematic review. Eur J Orthod. 2016 Dec 20. pii: cjw079. doi: 10.1093/ejo/cjw079). I was the lead investigator, responsible for all major areas of concept formation, data collection and analysis, as well as manuscript composition. Lee J contributed to data collection. Pliska BT was the supervisory author on this project and was involved throughout the project in concept formation, data analysis, and manuscript composition.

A version of Chapter 3 has been submitted for publication. I was the lead investigator, responsible for all major areas of concept formation, data collection and analysis, as well as the majority of manuscript composition. Pliska BT was the supervisory author on this project and was involved throughout the project in concept formation, data analysis, and manuscript composition.

Table of Contents

Abstract …………………………………………………………………………………….. **ii**

Lay Summary ……………………………………………………………………………… **iii**

Preface ………………………………………………………………………...................... **iv**

Table of Contents ………………………………………………………………………….. **v**

List of Tables ……………………………………………………………………………… **viii**

List of Figures …………………………………………………………….......................... **ix**

List of Abbreviations ……………………………………………………………………… **x**

Acknowledgements …………………………………………………….......................... **xii**

Dedication …………………………………………………………………………………. **xiii**

Chapter 1: Introduction ……………………………………………………………….... **1**

1.1 Computed Tomography (CT) and Three-Dimensional Imaging ……………... **1**

1.1.1 Cone-beam CT and Fan-beam/Medical/Helical CT ……….................... 1

1.1.2 Hounsfield Scale and Grey Scale ……………………….................... 1

1.2 CBCT and Dentistry ……………………………………………………………... **2**

1.2.1 Advantages and Disadvantages of CBCT in Dentistry …….................... 2

1.3 CBCT and Orthodontics ……………………………………………………… **3**

1.3.1 Orthodontics and OSA …………………………………...................... 3

1.3.2 Obstructive Sleep Apnea …………………………………………………… 3

1.3.3 Pathophysiology of OSA …………………………………………………… 4

1.3.4 CBCT and Upper Pharyngeal Airway Assessment ……………………… 6

1.4 Objective …………………………………………………………………... 6

Chapter 2: Reliability of upper pharyngeal airway assessment using dental CBCT: A systematic review ……………………………………………………………………... 7

2.1 Introduction……………………………………………………….......................... 7

2.2 Material and Methods ………………………………………………………… 7

2.2.1 Protocol and Registration …………………………….......................... 7

2.2.2 Eligibility Criteria ………………………………………………………….. 8

2.2.3 Information Sources, Search Strategy, and Study Selection …………….. 8

2.2.4 Data Items and Collection………………………………………………... 8

2.2.5 Risk of Bias/Quality Assessment in Individual Studies …………….…. 9

2.2.6 Synthesis of Results and Risk of Bias/Quality Across Studies ………… 11

2.3 Results ……………………………………………………………………….. 11

2.3.1 Study Selection …………………………………………………………….. 11

2.3.2 Study Characteristics ……………………………………………………. 13

2.3.3 Risk of Bias/Quality of Studies …………………………………………. 13

2.3.4 Summary Description of the Studies ……………………………………. 16

2.3.5 Synthesis of the Results …………………………………………………. 16

2.3.6 Additional Analysis ……………………………………………………... 17

2.4 Discussion ……………………………………………………………………… 17

2.4.1 Limitations of the Available Evidence …………………………………... 18

2.4.2 Clinical Implications …………………………………………………….. 19

2.5 Conclusions ………………………………………………………………….. 20

2.6 Conflict of Interest ………………………………………………………….. 20

Chapter 3: Reliability of upper pharyngeal airway assessment using dental CBCT …… 21

3.1 Introduction ……………………………………………………………………… 21

3.2 Material and Methods …………………………………………………………… 21

3.3 Results ……………………………………………………………………………….. 37

3.4 Discussion ………………………………………………………………………… 44

3.5 Conclusions ……………………………………………………………………….. 46

Chapter 4: Should Dental CBCT Be Used Today For Quantitative Assessments of the Upper Pharyngeal Airway: Final Thoughts ………………………………………………… 47

4.1 Conclusion ………………………………………………………………………… 49

Bibliography …………………………………………………………………………………... 50

Appendix A ……………………………………………………………………………………. 61

List of Tables

Table 2.1 Evaluation scores of the included studies (N=42) …………………………………... 14

Table 3.1 Definitions of the anatomic boundaries for each region of the upper pharyngeal airway ……………………………………………………………………………………………... 36

Table 3.2 ICC values for intra-examiner and inter-examiner reliability for all scans …………. 40

Table 3.3 ICC values for intra-examiner and inter-examiner reliability for the fast scan protocol ……………………………………………………………………………………………... 41

Table 3.4 ICC values for intra-examiner and inter-examiner reliability for the slow scan protocol ……………………………………………………………………………………………... 42

Table 3.5 Examiner variance for all parameters...………………...……………………………. 43

Table A.1 Characters of the included studies in the systematic review (N=42) ……………….. 61

Table A.2 CBCT machine settings of the included studies in the systematic review (N=42) ... 110

Table A.3 Examination characteristics of the included studies in the systematic review (N=42) ……………………………………………………………………………………………. 113

Table A.4 Raw data for threshold value ...……………………………………………………. 115

Table A.5 Raw data for measured minimum cross-sectional area in mm^2…..…………………….116

Table A.6 Raw data for measured total upper pharyngeal airway volume in mm^3……………117

Table A.7 Raw data for measured nasopharyngeal airway volume in mm^3…….……………...118

Table A.8 Raw data for measured oropharyngeal airway volume in mm^3…..………………….119

Table A.9 Raw data for measured hypopharyngeal airway volume in mm^3….………………..120

List of Figures

Figure 2.1 Evaluation checklist for the included studies ………………………………………. 10

Figure 2.2 PRISMA flow diagram of the literature selection process …………………………. 12

Figure 3.1 Orientation of the axial plane using the lower border of the orbit landmarks ……… 24

Figure 3.2 Orientation of the coronal plane using the Frankfort Horizontal Plane (porion to orbitale) …………………………………………………………………………………… 25

Figure 3.3 Orientation of the midsagittal plane using the upper incisive foramen to opisthion ………………………………………………………………………………………... 26

Figure 3.4 Correct sensitivity thresholding where the upper pharyngeal airway is completely filled with no "fingers" projecting out from the airway ……………………………………….. 27

Figure 3.5 Incorrect sensitivity thresholding where the upper pharyngeal airway is under-filled ………………………………………………………………………………………... 28

Figure 3.6 Incorrect sensitivity thresholding where the upper pharyngeal airway is over-filled to the point where "fingers" can be seen projecting out of the airway …………………………… 29

Figure 3.7 Anatomic boundaries of the upper pharyngeal airway including its comprising regions ………………………………………………………………………………………... 30

Figure 3.8 Nasopharyngeal airway volume …………………………………………………. 31

Figure 3.9 Oropharyngeal airway volume …………………………………………………... 32

Figure 3.10 Hypopharyngeal airway volume ……………………………………………….. 33

Figure 3.11 Total upper pharyngeal airway volume …………………………………………... 34

Figure 3.12 Minimum cross-sectional area of the upper pharyngeal airway ………………….. 35

Figure A.1 Landmarks used for hard tissue orientation of the CBCT scans …………………..121

Figure A.2 Examiner data collection form ………………………………………………… 122

List of Abbreviations

CT: computed tomography

HCT: helical computed tomography

CBCT: cone-beam computer tomography

HU: Hounsfield Unit

FOV: field of view

3D: three-dimensional

DICOM: digital imaging and communications in medicine

CSA: central sleep apnea

OSA: obstructive sleep apnea

μSv: microsievert

AHI: apnea-hypopnea index

MRI: magnetic resonance imaging

BMI: body mass index

CNS: central nervous system

J.N.Z.: Jason N. Zimmerman

J.L.: Janson Lee

B.T.P.: Benjamin T. Pliska

mA: milliampere

kVp: peak kilovoltage

sec: seconds

mm: millimeters

ICC: intraclass correlation coefficient

#: number

ALARA: as low as reasonably achievable

HIPA: Health Information Protection Act

PNS: posterior nasal spine

CI: confidence interval

SDB: sleep-disordered breathing

UMN: University of Minnesota School of Dentistry, Division of Orthodontics

Acknowledgements

I would like to thank my supervisor Dr. Benjamin Pliska for his mentorship and guidance. His direction has furthered my research skills which will benefit me in clinical practice. I would also like to thank Dr. Fernanda Almeida, Dr. Nancy Ford, and Dr. Sid Vora for serving on my thesis committee and their support throughout the research process. Lastly, I would like to thank Dr. Walter Siqueira who first introduced me to the world of research and without whom I would not be here today.

Dedication

I dedicate this thesis to my family for their constant support and encouragement throughout my eleven years of university education. I would also like to dedicate this thesis to my soon-to-be fiancé Fernanda Barona. Your love, friendship, and wonderful sense of humour were the greatest gifts that this program has bestowed upon me and I cannot imagine going through these three years without you. You are all the motivation that keeps me going through the many long nights and I would be nothing without all of you. I love you all!

Chapter 1: Introduction

1.1 Computed Tomography (CT) and Three-Dimensional Imaging

At a time where medicine and dentistry were limited to using a two-dimensional radiograph to aid in diagnosis and treatment planning, the introduction of three-dimensional radiographic imaging through computed tomography (CT) has revolutionized radiography forever. Where once clinicians were forced to use a two-dimensional tool to assess a three-dimensional patient, CT can provide diagnostic information that can lead to more effective and efficient treatment. There are two principal types of CT; cone-beam CT (CBCT) and fan-beam/medical/helical (HCT).[1]

1.1.1 **Cone-beam CT and Fan-beam/Medical/Helical CT**

CBCT uses an x-ray beam that is in the shape of a cone and is the type of CT commonly used in dentistry.[1] CBCT images are generated using a rotating gantry with a fixed x-ray source that creates the cone-shaped beam of ionizing radiation. This beam is then directed through the centre of the patient onto an x-ray detector on the opposite side. Throughout the scan, the x-ray source and detector are rotating around the centre of the patient, producing multiple consecutive planar images of the field of view (FOV).[2] The scanning software collects these multiple images and reconstructs them into digital volume units called voxels containing anatomical data which can be displayed by said software.[3] Where a CBCT scan integrates the entire FOV, only one rotation of the gantry is required to produce an image.[2] One can achieve higher spatial resolution by applying scan settings using a longer scan time and a smaller voxel size.[4]

HCT, also known as medical CT uses an x-ray beam that is in the shape of a fan in a helical progression (Figure 1.1).[1] This produces a series of image slices of the FOV that are then sandwiched together to create a 3D image.[2]

1.1.2 **Hounsfield Scale and Grey Scale**

The Hounsfield Scale is a standardized quantitative scale for describing radiodensity. In a medical HCT scan, the Hounsfield Unit (HU) is proportional to the degree of x-ray attenuation by the tissue. This standardization means that a radiodensity value from a scan on one machine can be directly compared with a radiodensity value from a scan on a different machine. There is no such standardization in reconstructed grey density values from scans derived from CBCT machines.[5]

The Grey Scale is a measure of x-ray attenuation in dental CBCT. CBCT manufacturers have not introduced a standard system for displaying grey scale. Some studies have shown a strong linear relationship between HU and gray scale. However, grey scale differs from HU in that grey scale is associated with higher noise levels, increased scattered radiation, high heel effect, and beam hardening artifacts.[5]

1.2 CBCT and Dentistry

Two-dimensional radiographic imaging techniques have been conventionally used in dentistry for generations as a diagnostic aid to appropriately treat patients. Then, three-dimensional radiography through dental CBCT became readily available in the late 1990's and transformed clinical dentistry. Its interest has expanded in both dental research and clinical practice, among general dentists and specialists alike. Use of CBCT can range from traumatology and studying craniofacial anomalies to implantology.[6]

1.2.1 **Advantages and Disadvantages of CBCT in Dentistry**

Advantages of using CBCT include the ability to produce two-dimensional images from the 3D data, fewer metal artifacts, isotropic voxel size, has a smaller footprint and uses less energy than HCT, and is digital imaging and communications in medicine (DICOM) compatible.[6] CBCT is also less expensive for the dentist to operate and is more compact than HCT, allowing for in-office imaging. CBCT also has a lower radiation dose due to a shorter exposure time compared with HCT.[1]

However, limitations of CBCT include a limited detector size, low contrast range, a restricted FOV, limited inner soft tissue information, increased scatter radiation and reduced contrast resolution, and the inability to be used for the estimation of HU.[6]

Compared to conventional 2D radiography, CBCT is also associated with a higher level of radiation exposure for patients.[7] It is estimated that the effective dose of a conventional panoramic radiograph is 24.3 µSv and for a cephalometric radiograph is 5.6 µSv.[8] The effective dose of CBCT for a small FOV is 48-652 µSv and for a large FOV is 68-1073 µSv, which is relatively small compared to a conventional CT scan which is 534-2100 µSv.[9,10] Therefore the effective doses do differ significantly across CBCT machines and is substantially higher compared to conventional panoramic radiographs, but are still considerably less compared to conventional CT.[11]

Nonetheless, efforts have been made to reduce radiation dosage associated with CBCT including narrow collimation[12] and changing the rotational angle from 360^{o} to 180^{o} [13]. Some potential improvements to reduce patient radiation exposure associated with CBCT include reducing patient dose with high resolution, varying FOV, enhancing image quality, and reducing scan time.[2]

1.3 CBCT and Orthodontics

In the specialty of orthodontics, CBCT can have many uses such as assessing the location of supernumerary or impacted teeth and potential root resorption associated with these conditions.[6] Some studies have suggested the use of CBCT for fabrication of surgical guides for the placement of orthodontic mini-implants.[14] Furthermore, changes in the condyles, rami, chin, maxilla and dentition can be assessed by superimposing the CBCT scans taken before and after orthognathic surgery.[15] More recently, orthodontics has employed the use of CBCT for airway assessment.

1.3.1 **Orthodontics and OSA**

The relatively recent involvement of orthodontists with obstructive sleep apnea (OSA) in both children and adults has furthered the interest of CBCT in the assessment of the upper airway. Nasal obstruction and sleep disordered breathing has been shown to be associated with altered craniofacial growth in some patients.[16] More recently, common facial orthopaedic treatments have demonstrated effectiveness for paediatric OSA.[17] As such the relationship of upper airway anatomy to sleep disordered breathing development and treatment continues to be an area of ongoing research.

1.3.2 **Obstructive Sleep Apnea**

Sleep apnea is a life threatening condition with two subcategories; central sleep apnea (CSA) and obstructive sleep apnea (OSA). CSA is triggered by the brain temporarily ceases to send signals to the respiratory muscles which regulate breathing. This results in multiple cessation episodes of respiration during sleep.[18,19] OSA is a disorder that is a result of complete or partial collapse of the airway, leading to disturbances in respiratory parameters and abnormal sleep.[19]

In an adult, an apnea is described as a complete cessation of airflow for a minimum of 10 seconds. A hypopnea is characterised by a decrease in airflow below 70% for a minimum of 10 seconds with a 4% or greater blood oxygen desaturation. A hypopnea can instead be defined as a reduction of airflow below 50% for a minimum of 10 seconds with a 3% desaturation, or the event is associated with arousal.[20] The apnea-hypopnea index (AHI), which is the standard for diagnosing OSA, is the combined total number of apneas and hypopneas per one hour of sleep.[19,21] In adults, a diagnosis of OSA can be made when the AHI of a patient is 5 or greater and demonstrates symptoms of excessive daytime sleepiness, fatigue, disturbed sleep with choking or gasping, experiencing non-refreshing sleep, or if the bed partner reports loud snoring or pauses in respiration while the patient is sleeping.[22,23] A reported AHI of 5-15 events per hour is described as mild OSA, between 15-30 as moderate OSA, and 30 or greater as sever OSA.[24]

The prevalence of OSA has been reported to be 2% of women and 4% of men ages 30-60 years old, with OSA patients being most commonly middle-aged men who are also overweight.[25,26] However it is important to note that there may be many patients who go undiagnosed.[18,19] This is because OSA can be asymptomatic and the prevalence of these patients with OSA who do not exhibit a clinical syndrome can be up to 30% among the middle-aged population.[19,21]

Other risk factors for OSA seen in adult patients include an increased body mass index (BMI), an increased neck circumference, race, family history, alcohol use, smoking, use of sedatives, and nasal congestion.[27,28] However, since OSA is less common in women, other factors including neuromuscular pathways may contribute to protecting the airway from constriction.[29,30] OSA is not limited to patients who are overweight, but can also occur in those of normal bodyweight who have anatomic abnormalities.

1.3.3 **Pathophysiology of OSA**

The pharynx is a funnel-shaped tube which is fibromuscular, is approximately 15 centimeters in length, and functions as a conduit for food and air.[31,32] It is located superior to the larynx, esophagus, and trachea, and dorsal to the oral and nasal cavities.[32–36] The pharynx can be divided into 3 components; the nasopharynx, oropharynx, and hypopharynx from superior to inferior. The nasopharynx is located posterior to the nasal cavity while the oropharynx is posterior to the oral cavity.[32,37,38] The hypopharynx extends from the tip of the epiglottis to the lowest portion of the airway at the larynx. A large number of muscles affect this portion of the airway, often acting in concert with or opposition to other related muscles.[20]

A patient experiences an OSA event when the pharyngeal airway narrows or closes with respiratory effort during sleep. The pharyngeal airway is unique in that it has no rigid support, instead being muscle and ligament formed and supported. While the patient is awake, muscle tensions keep the lumen patent. While the patient is asleep however, the muscles relax and the pharyngeal walls become more flexible and more collapsible. Furthermore, in the reclined position the effects of gravity distort the pharyngeal walls, especially by retropositioning of the tongue while the patient is supine, resulting in a narrowed lumen. Since the required volume exchange of air remains the same, a higher velocity is necessary through the smaller passageway. This airflow is turbulent, causing vibration and flutter of the flexible walls and soft palate, resulting in (often loud) snoring. The narrower the lumen, the faster the velocity and the lower the pressure.[20]

Once a critical point is reached, this combination of physical conditions will result in an occluded airway. Although respiratory effort will continue, with the diaphragm contracting downward forcefully enough that the chest walls may be drawn inward, there will be no air exchanged until there is sufficient arousal (lighter level of sleep) to regain enough muscle tension and reopen the pharyngeal airway. This sequence of loud snoring, sudden silence, and loud resuscitative “snort” is not only virtually pathognomonic for OSA, but is frequently what drives the patients and their families to seek treatment.[20]

The resultant and repetitive apnea events characteristic of OSA can be associated with many symptoms including loud irregular snoring,[20] long pauses in breathing during sleep,[20] excessive daytime sleepiness,[20,22,23,26,39] obesity,[20] fatigue,[20] impotence,[20] morning headaches,[20] and changes in cognitive functions such as alertness, memory, personality, or behavior.[20] This can lead to motor vehicle,[20,21,25,26,40,41] reduced quality of life,[20,21,26] decreased work performance,[20] along with many consequences to the patients' health. These can include cardiovascular disease,[18,42–44] hypertension,[20,22,23,25,27] coronary artery disease,[20,25,26,39,45] deep vein thrombosis,[18,43,46,47] stroke,[18–20,27,46] and sudden death.[19–21,40,44] In patients with more chronic cases, OSA has been associated with cor pulmonale,[27,41] pulmonary hypertension,[47,48] polycythemia,[21,49,50] and metabolic syndromes.[48,51] Therefore OSA can be quite incapacitating and life-threatening.

According to imaging techniques including lateral cephalograms, magnetic resonance imaging (MRI), and CBCT, the airway constriction associated with OSA most often occurs in the retropalatal and retroglossal regions of the oropharynx.[33–36,52–54] This region of the pharynx is particularly vulnerable in OSA patients compared to normal control patients due to decreased collapsing pressures and airway dimensions, which is observed in patients under general anesthesia with complete muscle paralysis.[55–58]

Hard tissue craniofacial abnormalities commonly associated with OSA as revealed by radiography include a short anterior cranial base,[56–58] a retrognathic and retruded maxilla and mandible in relation to cranial base,[26,52,53,57,59–63] an increased mandibular plane angle,[26,53,57] a large gonial angle,[53,62,64,65] a decreased upper to lower facial height ratio,[57,64] an increased lower facial height,[53,58,66] and an inferior and counter-clockwise translation of the hyoid bone.[29,62,64,65] All of the above can result in the development of a compromise in airway dimension.

The soft tissue can also significantly contribute to upper pharyngeal airway risk factors related to OSA. Studies using lateral and posteroanterior cephalometry assessed restricted posterior airway space and discovered that thickening of the velum and velopharyngeal lumen can compromise the airway.[30,64,67] Furthermore, an increased tongue size, a longer soft palate, and lateral pharyngeal wall size can also contribute to OSA.[37,38,53,58–63,66,67]

Neurologic control of the upper pharyngeal airway also plays an interconnected part in OSA. The normal physiological process of respiration involves signals being sent from the medulla to the respiratory centres, which through the inspiratory phase stimulates the genioglossus to prepare the upper pharyngeal airway for the development of negative intrapharyngeal pressure. Airway patency is maintained by the pharyngeal abductor and dilator muscles.[30,67] Once the central nervous system (CNS) signals the upper pharyngeal airway and diaphragm the muscles go into a hypotonic state, and the size of the pharynx and soft tissue determine airway stability while the patient is sleeping.[53,67,68] Airway obstruction occurs if the negative intraluminal pressure created during inspiration surpasses the support of the soft tissues in the airway.[30,69] As a result, the CNS attempts to maintain airway patency by signaling the muscles to go into a hypertonic state to resume respiration, which leads to a lighter level of sleep.[30,70]

Patients with obstructive sleep apnea have been shown to have a significantly reduced total airway volume, airway area, airway width, and a significantly larger airway length compared to patients who do not suffer from obstructive sleep apnea.[71] This is important because the frequency of airway collapse increases in patients who have narrower and longer airways.[72] Obstructive sleep apnea patients have also been shown to have a significantly larger tongue for a given maxillomandibular size than patients who do not have obstructive sleep apnea.[73]

1.3.4 **CBCT and Upper Pharyngeal Airway Assessment**

The ability to assess the upper airway in three-dimensions and the lower radiation dose compared to medical CT imaging makes CBCT an attractive potential tool for the assessment of OSA patients.[74]

However it remains to be determined if CBCT can provide anything beyond a qualitative assessment of upper airway anatomy. In order for CBCT to become a resource for quantitative airway assessment, its reliability as a measurement tool must first be established. For the purpose of this thesis, reliability is defined as the agreement between measurements for the same examiner (intra-examiner) or between different examiners (inter-examiner).

1.4 Objective

The aim of this study was to determine the reliability of volumetric and cross-sectional area assessments of the upper pharyngeal airway using dental CBCT. This would be accomplished by first conducting a systematic review of the literature, followed by an original study to fill in the subsequently revealed knowledge gaps.

Chapter 2: Reliability of upper pharyngeal airway assessment using dental CBCT: A systematic review[75]

2.1 Introduction

Dental cone beam computed tomography (CBCT) became readily available in the late 1990's and revolutionized dental radiography. Its interest has expanded in both dental research and clinical practice, among general dentists and specialists alike. Use of CBCT can range from traumatology and studying craniofacial anomalies to implantology. In the specialty of orthodontics, CBCT can have many uses such as assessing the location of supernumerary or impacted teeth and potential root resorption associated with these conditions.[6]

The relatively recent involvement of orthodontists with obstructive sleep apnea (OSA) in both children and adults has furthered the interest of CBCT in the assessment of the upper pharyngeal airway. Nasal obstruction and sleep disordered breathing has been shown to be associated with altered craniofacial growth in some patients.[16] More recently, common facial orthopaedic treatments have demonstrated effectiveness for paediatric OSA.[17] As such the relationship of upper airway anatomy to sleep disordered breathing development and treatment continues to be an area of ongoing research. The ability to assess the upper pharyngeal airway in three-dimensions and the lower radiation dose compared to medical CT imaging makes CBCT an attractive potential tool for the assessment of OSA patients.[74] However it remains to be determined if CBCT can provide anything beyond a qualitative assessment of upper airway anatomy. In order for CBCT to become a resource for quantitative airway assessment, its reliability as a measurement tool must first be established. For the purpose of this review, reliability is defined as the agreement between measurements for the same examiner (intra-examiner) or between different examiners (inter-examiner).

Therefore, the purpose of this study is to systematically review the literature to evaluate the reliability of upper pharyngeal airway assessment using dental CBCT.

2.2 Material and Methods

2.2.1 **Protocol and Registration**

The protocol for the present systematic review was constructed a priori according to the Cochrane Handbook for Systematic Reviews of Interventions 5.1.0 and is available upon request. This systematic review follows the PRISMA statement,[76] its extension for abstracts,[77] and was not registered.

2.2.2 **Eligibility Criteria**

The following selection criteria were used for the systematic review:

1. Human studies involving patient data (not phantoms or simulated anatomy)
2. Use of CBCT imaging
3. Assessment of the upper pharyngeal airway
4. Reliability reported

2.2.3 **Information Sources, Search Strategy, and Study Selection**

The electronic databases of MEDLINE, EMBASE and Web of Science were searched through June 2015. The search tree used for the MEDLINE database is provided in Appendix 1, and similar trees were used for the subsequent databases. The studies included were restricted to those written in the English language. A limited gray literature search was conducted using Google Scholar by limiting the examination to the first 100 most relevant hits. Authors were contacted to identify unpublished literature or ongoing studies, and to clarify data as needed. The reference lists of the included studies were also searched for any relevant studies.

Assessment of the literature for inclusion in the systematic review, and the extraction of data were completed independently and in duplicate by two investigators (J.N.Z. and J.L.). Any discrepancies were resolved by consultation with the third author (B.T.P.). Risk of bias/quality assessment was also completed independently and in duplicate by two investigators (J.N.Z. and B.T.P.), with the third author (J.L.) resolving any discrepancies. The investigators were not blinded to the authors or the results of the research.

2.2.4 **Data Items and Collection**

Three different data extraction tables were developed. The first (Table A.1) recorded whether or not the study was randomized, sample size, age of the sample, whether or not the sample was syndromic, whether or not a control was used, if a gold standard was used, what kind of segmentation was used, the airway region measured, the measurements recorded (volume and/or minimum cross-sectional area), the reliability test used and statistics, imaging software used, and the threshold values used (if any).

The second data extraction table (Table A.2) recorded the CBCT machine used, field of view, tube current (mA), tube potential (kVp), exposure time (sec), and resolution/voxel size (mm).

The third data extraction table (Table A.3) recorded the number of examiners, the number of times the measurements were repeated, the time period between repeated measurements, and the qualifications of the examiner(s).

2.2.5 **Risk of Bias/Quality Assessment in Individual Studies**

Faced with a lack of an appropriately validated tool that is clearly indicated for risk of bias/quality assessment for reliability studies, it was decided to search for a method that was as systematic and objective as possible. A previously conducted systematic review on a similar topic was identified[78] and their assessment tool was used with minimal and appropriate adjustments to systematically assess the selected studies (Figure 2.1). There were three main parameters evaluated: study design, study measurements, and data analysis. Each of these three parameters were divided further into sub-sections.

Study design was divided into whether or not the sample was randomized, whether or not the sample size was greater than or equal to thirty subjects, whether or not a control was used, whether a human sample was used, and the method of segmentation. Study measurements was divided into the gold standard used, the portion of the airway studied, and the measurement assessed. Data analysis was divided into the type of reliability assessed and the statistical test used.

Each study was awarded a given number rating for fulfilling the sub-parameters, where each sub-parameter had a maximum rating that could be awarded.[78] If any of the sub-parameters were not fulfilled, then a zero was entered for that particular sub-parameter. The sum up to a maximum of 20 represented the overall quality of the study, with a higher rating signifying a higher quality of the study.

Figure 2.1 Evaluation checklist for the included studies

Parameters of evaluation				*Maximum score*
1.	Study design	(a)	Randomized sample (*)	1
		(b)	Sample size ≥30 (*)	1
		(c)	Control group included (*)	1
		(d)	Human sample (*)	1
		(e)	Method of segmentation:	1
			Algorithm (*)	
			Commercial software (*)	
2.	Study measurements	(f)	Gold standard:	4
			Physical model (***)	
			Manual segmentation (****)	
		(g)	Portion of airway:	4
			Nasopharynx (*)	
			Oropharynx (*)	
			Hypopharynx/Velopharynx (*)	
			Total upper pharyngeal airway (*)	
		(h)	Type of measurement:	3
			Volume (**)	
			Minimum cross-sectional area (*)	
3.	Data analysis	(i)	Reliability:	2
			Intra-examiner (*)	
			Inter-examiner (*)	
		(j)	Statistical test used:	2
			ICC (**)	
			Other appropriate statistical test (*)	
Total				20

2.2.6 **Synthesis of Results and Risk of Bias/Quality Across Studies**

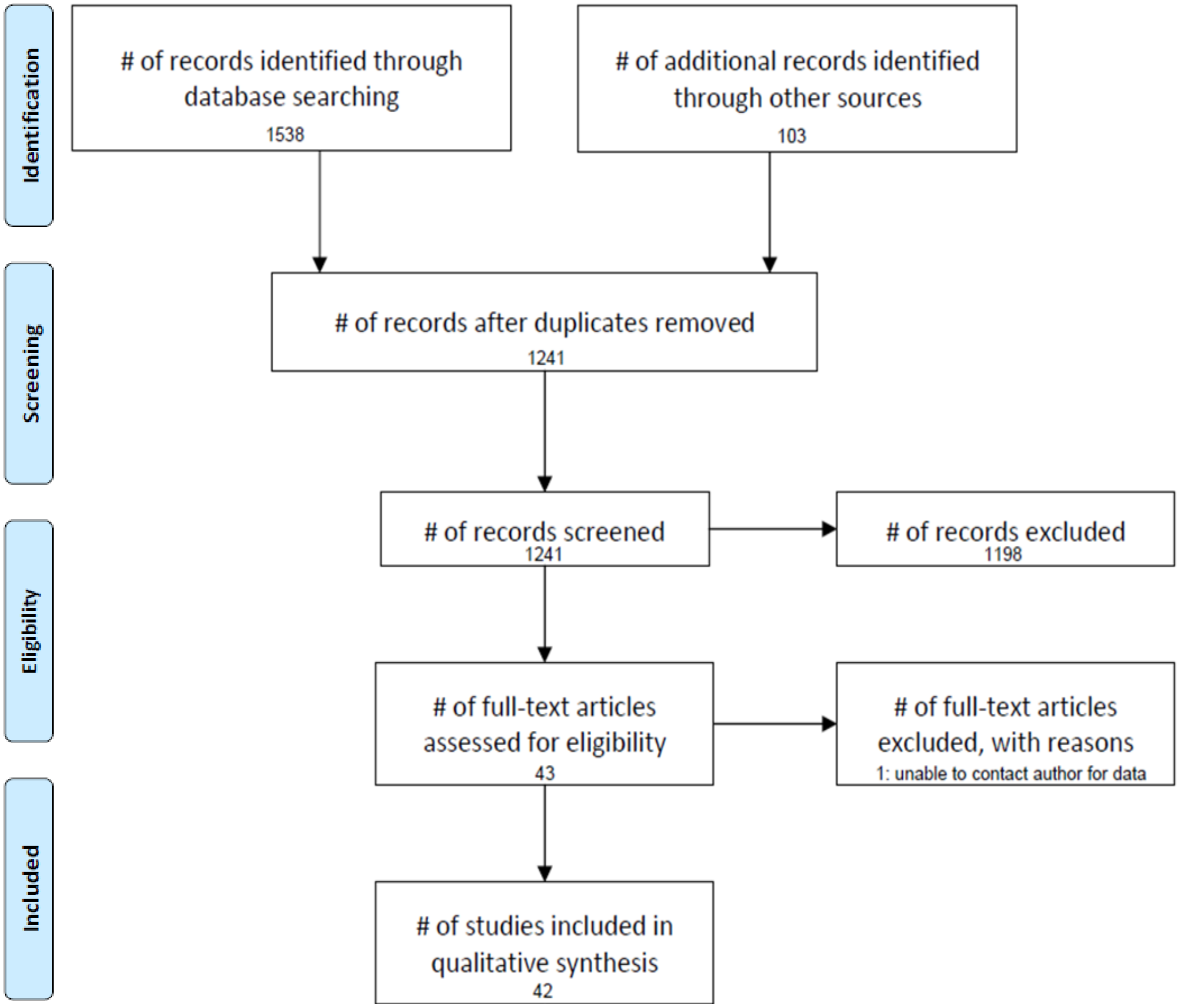
It was determined a priori that if the data extracted from each study was adequately homogeneous and the combination of the extracted data was valid, a meta-analysis would be conducted.

2.3 Results

2.3.1 **Study Selection**

Of the 1241 studies that were screened, 43 articles satisfied the inclusion criteria.[74,79–120] However due to the inability to make contact with the authors of one study[120] in order to obtain the required data, this study had to be excluded. A flowchart following the PRISMA format is provided (Figure 2.2), outlining the selection process employed.

Figure 2.2 PRISMA flow diagram of the literature selection process



2.3.2 **Study Characteristics**

The selected studies included the CBCT scans of 956 patients evaluated for reliability of upper pharyngeal airway assessment. The studies exhibited considerable variations in sample size (ranging from 4-71 scans), mean patient ages (ranging from 8-48 years old), imaging software, machine settings, and examiner protocols (Tables A.1-A.3). The assessed scans were of a wide spectrum of patients, including those with various syndromes and patients receiving orthodontic treatment (Table A.1). The studies also used examiners with an array of qualifications including dental students, general dentists, orthodontic residents, orthodontists, physicians, maxillofacial surgeons, and dental radiologists (Table A.3).

The most commonly used CBCT machine was i-CAT (Imaging Sciences International), and the most frequently used imaging software was Dolphin Imaging®. A majority of the studies used intraclass correlation coefficient (ICC) as the reliability statistic, followed by Dahlberg's formula being the next most common statistical analysis.

2.3.3 **Risk of Bias/Quality of Studies**

There were 42 studies that were assessed for methodological quality (Table 2.1). A score of ≥13/20 was deemed as a high quality study. Only 5 of the 42 studies fulfilled this criteria.[81,87,89,94,105]

The major methodological limitation was the lack of a gold standard used in the study. The next two biggest limitations were sample size and lack of a control group. Randomization of the sample was another key limitation indicating the potential risk of bias.

Table 2.1 Evaluation scores of the included studies (N=42)

	Parameters of scoring (x: maximum score)										
	Study design					*Study measurements*			*Data analysis*		*Total score, n (% out of 20)*
Studies evaluated	(a) = 1	(b) = 1	(c) = 1	(d) = 1	(e) = 1	(f) = 4	(g) = 4	(h) = 3	(i) = 2	(j) = 2	
Alves et al.[79]	1	0	0	1	1	0	1	3	1	2	10 (50)
Alves et al.[80]	1	0	0	1	1	0	1	3	1	2	10 (50)
Bandiera et al.[81]	1	1	1	1	1	0	3	3	1	2	14 (70)
Brunetto et al.[82]	0	0	0	1	1	0	4	3	1	2	12 (60)
Burkhard et al.[83]	1	0	0	1	1	0	1	3	2	1	10 (50)
Celikoglu et al. [84]	1	0	1	1	1	0	3	2	1	2	12 (60)
Chang et al.[85]	0	0	0	1	1	0	1	1	1	2	7 (35)
Cheung and Oberoi[86]	1	0	1	1	1	0	1	3	1	1	10 (50)
De Souza et al.[87]	0	1	0	1	1	0	3	3	2	2	13 (65)
Di Carlo et al.[88]	1	0	0	1	1	0	4	2	1	2	12 (60)
El and Palomo[89]	1	1	0	1	1	4	2	2	1	2	15 (75)
Enciso et al.[90]	1	0	1	1	1	0	1	3	1	2	11 (55)
Feng et al.[91]	1	0	0	1	1	0	1	2	2	2	10 (50)
Glupker et al.[92]	1	0	0	1	1	0	2	3	1	2	11 (55)
Grauer et al. [93]	1	0	0	1	1	0	1	2	1	1	8 (40)
Guijarro-Martinez and Swennen[94]	0	1	0	1	1	0	3	3	2	2	13 (65)
Hart et al.[95]	0	1	0	1	1	0	3	3	1	2	12 (60)
Hong et al.[96]	1	0	0	1	1	0	1	3	1	1	9 (45)
Iannetti et al.[97]	0	0	0	1	1	0	1	2	2	1	8 (40)
Iwasaki et al.[98]	1	0	1	1	1	0	3	2	1	2	12 (60)
Jiang et al.[99]	1	0	0	1	1	0	1	3	1	2	10 (50)

	Parameters of scoring (x: maximum score)										
	Study design					*Study measurements*			*Data analysis*		*Total score, n (% out of 20)*
Studies evaluated	(a) = 1	(b) = 1	(c) = 1	(d) = 1	(e) = 1	(f) = 4	(g) = 4	(h) = 3	(i) = 2	(j) = 2	
Kim et al.[100]	1	0	0	1	1	0	4	3	1	1	12 (60)
Kim et al.[101] (30)	1	0	0	1	1	0	4	3	1	1	12 (60)
Kochel et al.[102]	1	0	0	1	1	0	4	3	1	1	12 (60)
Lenza et al.[74]	1	0	0	1	1	0	4	3	1	1	12 (60)
Li, L. et al.[103]	0	1	1	1	1	0	1	1	1	1	8 (40)
Li, YM. et al.[104]	0	0	0	1	1	0	3	2	1	2	10 (50)
Mattos et al.[105]	1	0	0	1	1	0	4	3	2	2	14 (70)
Oh et al.[106]	0	1	0	1	1	0	1	3	1	2	10 (50)
Sears et al.[107]	1	0	0	1	1	0	3	2	1	1	10 (50)
Starbuck et al.[108]	1	0	0	1	1	0	1	2	1	2	9 (45)
Stefanovic et al.[109]	0	1	1	1	1	0	2	3	1	2	12 (60)
Valladares-Neto et al.[110]	1	0	0	1	1	0	3	3	1	1	11 (55)
Vizzotto et al.[111]	0	0	0	1	1	0	2	1	1	2	8 (40)
Weissheimer et al.[112]	0	1	0	1	1	3	1	2	1	2	12 (60)
Xu et al.[113]	0	1	1	1	1	0	1	3	2	2	12 (60)
Yoshihara et al.[114]	1	0	1	1	1	0	3	3	1	1	12 (60)
Zhao et al.[115] (44)	0	1	1	1	1	0	2	3	1	2	12 (60)
Zheng et al.[116]	1	0	0	1	1	0	4	3	1	1	12 (60)
Aboudara et al.[117]	1	0	0	1	1	0	1	3	1	2	10 (50)
Haskell et al.[118]	0	0	1	1	1	0	1	3	1	2	10 (50)
Iwasaki et al.[119]	1	0	1	1	1	0	2	3	1	2	12 (60)

2.3.4 **Summary Description of the Studies**

All of the included studies assessed intra-examiner reliability. However, only 7 of the 42 included studies (~17%)[83,87,91,94,97,105,113] and only 3 of the 5 high quality studies (60%)[87,94,105] assessed inter-examiner reliability. From the high quality studies, upper airway volume showed good to excellent intra-examiner reliability (0.880-0.990) and minimum cross-sectional area showed moderate to excellent intra-examiner reliability (0.780-0.999). Upper airway volume demonstrated excellent inter-examiner reliability (0.986-0.998) while minimum cross-sectional area demonstrated moderate to excellent inter-examiner reliability (0.696-0.988). Both intra- and inter-examiner reliability varied depending on which section of the upper pharyngeal airway was assessed.

According to the high quality studies, intra-examiner reliability for total airway volume ranged from 0.987-0.990, and inter-examiner reliability from 0.950-0.992. Intra-examiner reliability for nasopharyngeal airway volume ranged from 0.880-0.992 while inter-examiner reliability was 0.986. Intra-examiner reliability for oropharyngeal airway volume ranged from 0.990-0.999 and inter-examiner reliability was 0.998. Intra-examiner reliability for hypopharyngeal airway volume ranged from 0.994-0.996 and inter-examiner reliability was 0.994.

Only 19 of the 42 included studies (~45%)[74,85,88,89,92,94,95,97,98,104–106,108,110–112,114,115,117] identified the qualifications of the examiners, with only 2 of the 5 high quality studies (40%)[89,105] doing so. Furthermore, only 1 of the studies[105] used more than 2 examiners. The intra- and inter-examiner reliabilities of both airway volume and minimum cross-sectional area did vary depending on the qualifications of the examiners.

A majority of the studies did not assess the upper pharyngeal airway in its entirety, with only 8 of the 42 included studies (~19%)[74,82,88,100–102,105,116] and 1 of the 5 high quality studies (20%)[105] doing so. Additionally, many of the studies did not assess both airway volume and minimum cross-sectional area. Only 28 of the 42 included studies (~67%)[74,79–83,86,87,90,92,94–96,99–102,105,106,109,110,113–119] and 4 of the 5 high quality studies (80%)[81,87,94,105] measured both.

Most importantly, not a single study had the examiners orient the scan on their own. Equally as critical, none of the studies had the examiners assign the appropriate sensitivity threshold value for each scan on their own.

2.3.5 **Synthesis of the Results**

The studies generally show high intra-examiner reliability with lower inter-examiner reliability. Furthermore, airway volume demonstrated greater intra- and inter-examiner reliability than did minimum cross-sectional area. Many of the studies only assess intra-examiner reliability, and do not address inter-examiner reliability. A majority of the studies do not assess the upper pharyngeal airway in its entirety, and several of the studies do not evaluate both airway volume and minimum cross-sectional area. Less than half of the studies provide the qualifications of the examiners evaluating the scans. Furthermore, none of the studies allows for manual image orientation or manual selection of the airway sensitivity threshold by the examiners themselves.

2.3.6 **Additional Analysis**

Considering the significant heterogeneity between study protocols in terms of field of view, scan settings, indication for image acquisition and the machine type used, a meta-analysis of the results was not possible.

2.4 Discussion

This systematic review was performed to assess the reliability of CBCT measurement of the upper airway, a process that has become increasing more common in the field of orthodontics. The practical aspects of airway analysis of a DICOM file generated from a CBCT scan of a patient generally involves several steps, each with its own potential for error. Orientation of the image is typically the first step following opening of the file in a software program used for the analysis. As the boundaries for the airway are most commonly based on lines parallel to horizontal plane of the image instead of internal landmarks, a standardized method of orientating the field of view in the frontal, sagittal and coronal planes is essential to consistent measurement. Following image orientation, the appropriate slice on which the airway boundaries are identified is chosen. The second step of the process requires the landmarks defining the boundaries of the airway to be then identified. Either of these initial steps is subject to some level of variability and operator error and should be accounted for when assessing method error. Indeed, in their study of CBCT software accuracy for airway analysis Weissheimer et al.[112] used a predefined and orientated airway segment in order to "eliminate variability introduced by using different imaging software programs to define the oropharyngeal airway".

The final step in airway measurement typically is to then choose the sensitivity threshold value at which the software program will differentiate soft tissue from air within the patient's anatomy. This value is selected on a sliding scale and it allows for the software to distinguish between soft tissue and airway by their radiodensities at the level of each voxel. The examiner does this by increasing the threshold value along the scale until the entire airway is shaded in by the software. It should be noted that the same threshold sensitivity value cannot be assigned to all patient scans as you can under- or over-fill the airway, thereby risking under- or overestimating the airway volume.[121] It is the authors' experience that this last step of choosing a threshold value is the most subjective and prone to effecting measurement accuracy and reliability. This has been also been discussed by others.[89,112]

The search strategy for this review was designed to include all studies that reported the method error or reliability of airway measurement as part of the study protocol. However three studies investigated reliability of CBCT in airway measurement as the specific aim of the study. The first of these studies was conducted by Guijarro-Martinez and Swennen[94], who assessed 35 non-syndromic patients between 23-35 years of age. Two examiners assessed the patient scans twice separated 4 weeks apart. They found that airway volume had excellent reliability, with an intra-examiner reliability of 0.981-0.999 and inter-examiner reliability of 0.986-0.998. Furthermore they found that minimum cross-sectional area had good-to-excellent reliability, with an intra-examiner reliability of 0.780-0.937 and an inter-examiner reliability of 0.839-0.876. Intra-examiner reliability varied depending on the specific part of the airway being assessed and the educational background of the examiner. Some limitations of this study are that the total airway volume was not assessed, only two examiners were used, image orientation was not specified to be performed by the examiners, and manual selection of the sensitivity threshold value was not indicated to have been used in the final assessment.

The second study was conducted by De Souza et al.[87], who assessed 60 non-syndromic patients with a mean age of 17.86 years. Two examiners assessed the patient scans twice separated by a two week interval. They found that total airway volume had excellent reliability, with an intra-examiner reliability of 0.99 and an inter-examiner reliability of 0.95. Nasopharyngeal minimum cross-sectional area had good-to-excellent reliability, with an intra-examiner reliability of 0.93-0.98 and an inter-examiner reliability of 0.88. Oropharyngeal minimum cross-sectional area had excellent reliability, with an intra-examiner reliability of 0.98-0.99 and inter-examiner reliability of 0.98. One limitation of this study is that the authors did not assess the reliability of each section of the upper airway in regards to volume. Also, the hypopharynx was not assessed at all on its own for reliability of volume or minimum cross-sectional area assessment. Furthermore there was no mention in the study as to whether or not image orientation and selection of the sensitivity threshold values was conducted manually. Lastly, only two examiners were used and their educational backgrounds or experience levels with the process were not provided.

The third study was conducted by Mattos et al.[105], who assessed 12 non-syndromic patients of unspecified age. Three examiners assessed the patient scans twice separated two weeks apart. They found that airway volume had excellent reliability, with an intra-examiner reliability of 0.987-0.995 and an inter-examiner reliability of 0.992. Minimum cross-sectional area had moderate to excellent reliability, with an intra-examiner reliability of 0.869-0.999 and an inter-examiner reliability of 0.696-0.988. Intra-examiner reliability depended on the specific location of the upper airway assessed and on the educational background of the examiners. Inter-examiner reliability depended on the specific location of the upper airway assessed. One limitation of this study is that the authors did not assess the reliability of each section of the upper airway in regards to volume. Furthermore, image orientation and sensitivity threshold value selection was not conducted by the examiners.

2.4.1 **Limitations of the Available Evidence**

In order to truly assess the reliability of CBCT as a tool to quantitatively measure the airway, the entire procedure of image processing from image orientation, to segmentation of the airway and the selection of threshold value must be evaluated as all three steps are fraught with subjectivity on the part of the examiner. However, the results of this systematic review have demonstrated that the reliability and method error reported in the included studies have only assessed the examiners' ability to reliably segment the airway. None of the studies have allowed for the examiners to orient the image or select the sensitivity threshold value manually despite this being essential to the process. Therefore, even though the studies indicate moderate to excellent reliability, two-thirds of the airway measurement protocol have been largely unexamined in the included studies.

Furthermore, the majority of the studies limited their assessment to intra-examiner reliability and did not consider inter-examiner reliability. Inter-examiner reliability is just as important as intra-examiner reliability as diagnostic consistency is not only essential within one professional, but amongst professionals as well. There is a wide range of healthcare professionals that would assess the airway of patients with CBCT, and operator experience has been previously shown to influence airway measurement reliability.[122] Often a team of professionals spanning different disciplines form a sleep team treating affected patients. It is also important for reliability amongst healthcare professionals with different backgrounds and training and this is something not readily addressed in the current literature.

For the above reasons, combined with the fact that many studies do not assess the upper pharyngeal airway in its entirety, the reliability of CBCT to assess the upper airway has not been adequately established. Further studies taking all sources of variability into account are still required to truly determine how reliably CBCT scans of patients can assess volume and minimum cross-sectional area of the upper pharyngeal airway.

2.4.2 **Clinical Implications**

It is important to note that ALARA principles and SedentexCT guidelines condemn the indiscriminate use of CBCT, stating that its use should be reserved for selected orthodontic cases where conventional radiography cannot provide necessary diagnostic information.[123] Therefore, not only should radiation exposure be kept to a minimum, but the use of CBCT examinations for any particular orthodontic patient should be justified.

The CBCT assessment of airway has become commonplace in many areas of orthodontic research, with anatomical linear and volumetric measurements being used to assess the effect of various orthodontic and surgical treatments. This is despite the fact that a validated and optimized CBCT protocol for airway imaging remains elusive.[78] The first step toward this goal would be to determine CBCT's reliability for upper airway assessment. Although the current literature suggests that there is moderate to excellent reliability, careful examination of the limitations of the current evidence implies that this question is still unanswered.

Future research should be directed at improving the quality of evidence by addressing both intra-examiner and inter-examiner reliability, while using ICC to describe the variation in measurement.[78] Furthermore, reliability should be assessed for both volume and minimum cross-sectional not only for the total upper pharyngeal airway but also for its component sections; the nasopharynx, oropharynx, and hypopharynx. The anatomical boundaries for each section of the upper pharyngeal airway should also be clearly defined and standardized. Having many examiners conducting such an assessment would be beneficial, along with assessing if and how reliability changes depending on the examiners' educational background and clinical experience. There was not sufficient data in the high quality studies to compare reliability between pediatric and adult patients, but such a study could be beneficial. Lastly, a meaningful study will allow the examiners to manually perform all steps actually required for assessing the upper pharyngeal airway including image orientation, landmark identification, and selection of the threshold sensitivity for the DICOM file.

2.5 Conclusions

Based on the current and limited evidence, upper pharyngeal airway assessment with CBCT demonstrated moderate to excellent intra- and inter-examiner reliability for volume and minimum cross-sectional area. However caution is warranted in interpreting these findings as CBCT reliability has only been examined under controlled conditions, which artificially restricts potential sources of variability. Furthermore, airway volume demonstrated greater intra- and inter-examiner reliability than did minimum cross-sectional area. However, limitations of the current evidence suggest that more research needs to be conducted to adequately determine the reliability of upper pharyngeal airway assessment using dental CBCT.

2.6 Conflict of Interest

There was no conflict of interest present for conducting this systematic review.

Chapter 3: Reliability of upper pharyngeal airway assessment using dental CBCT

3.1 Introduction

Dental radiography was revolutionized when cone beam computed tomography (CBCT) became readily available in the late 1990's. Since then its interest has rapidly increased in dental research and clinical practice among general dentists and specialists alike.[6]

The field of orthodontics is no exception and the relatively recent and increased awareness in obstructive sleep apnea (OSA) in children and adults has driven the assessment of the upper pharyngeal airway using CBCT to the forefront of academic and clinical interest. More specifically, the ability to perform a three-dimensional evaluation of the upper airway coupled with the lower radiation dose compared to medical CT imaging makes CBCT a potentially attractive tool for the assessment of airway anatomy in OSA patients.[16]

Before CBCT is employed to quantitatively assess the airway, it is crucial that we establish its reliability as a measurement tool. While the quantitative assessment of the airway is semi-automated with contemporary software programs, the operator must initially process the DICOM file through several steps including image orientation and selection of threshold sensitivity before measurements are made. These steps have the potential to introduce a level of subjectivity and negatively affect reliability of the airway analysis. A recent systematic review on the subject has highlighted the significant methodological limitations of the current literature.[75] Most significantly, the reliability and method error reported in the literature have only assessed the examiners' ability to reliably segment and trace the upper pharyngeal airway. None of the available studies allowed for the manual orientation of the CBCT images and selection of slice and threshold sensitivity by the examiners in the study protocols. Furthermore, there is not a single study that assesses the upper airway in its entirety or evaluates both inter-examiner and intra-examiner reliability.[75] Therefore, this suggests that reliability of upper pharyngeal airway assessment using CBCT has not been adequately established.

The purpose of this study was to determine the intra-examiner and inter-examiner reliability of the complete process of volumetric and cross-sectional area assessments of the upper airway using CBCT.

3.2 Material and Methods

The sample size was determined following the recommendations of Walter et al[124] for reliability studies. The parameters included $\rho_O = 0.5$ (minimum acceptable level of reliability), $\rho 1 = 0.9$ (expected level of reliability), $\alpha = 0.05$, $\beta = 0.2$ (implying a power test of 80%).[105] For this study it was decided that n = 2 (intra-examiner) and n = 6 (inter-examiner). Considering these factors, it was determined that a sample of CBCT images from a minimum of 9 patients would be sufficient.

The initial de-identified DICOM files of 10 adult patients treated at a university based orthodontic clinic were randomly selected from the orthodontic records database of previously treated patients. Patients younger than 18 years of age, or with clefts, craniofacial syndromes, detectable airway pathology, or those with previous orthognathic or craniofacial surgery were excluded from selection. This study adheres with the Health Information Protection Act (HIPA), and was accepted by the Research Ethics Board at the University of British Columbia (H12-00951).

The CBCT scans were taken by one operator using the same I-CAT tomograph (Imaging Sciences International, Hatfield, Pa). The patients were positioned ensuring that the Frankfort horizontal plane was parallel to the floor. They were instructed to occlude in maximum intercuspation with their tongue touching the palate, and were refrained from swallowing during the scanning period. Five of the scans were taken using the fast scan protocol and five scans were taken using the slow scan protocol. The slow scan protocol included 13 X 17 field of view, 0.3 mm voxel size, 17.8 second scan time, 120 kVp tube voltage, and 37.1 mA tube current. The fast scan protocol included 13 X 17 field of view, 0.4 mm voxel size, 8.9 second scan time, 120 kVp tube voltage, and 18.5 mA tube current. Images were saved in DICOM files which were uploaded into Dolphin Imaging software (version 11.5; Dolphin Imaging and Management Systems, Chats- worth, Calif) to obtain the primary reconstructed images and the 3D reconstructions.

An oral and maxillofacial radiologist, an academic orthodontist, an academic orthodontists with additional study in airway and sleep apnea, a private practice orthodontist, a senior orthodontic resident, and a junior orthodontic resident were orientated, trained, and calibrated as examiners for upper pharyngeal airway analysis using CBCT images not included in the study. The calibration protocol included an explanation of the 3D measurement tools in the Dolphin Imaging software and a demonstration of the measurements to be made for this study. A video and manual were also provided to train the examiners in manual scan orientation, slice selection, landmark identification, and threshold sensitivity selection for upper pharyngeal airway analysis.

Once calibration was complete, the examiners proceeded with the airway analysis protocol for each of the ten sample patients. This began with the examiners independently and manually orienting the patient 3D image in the coronal, sagittal and transverse planes (Figures 3.1-3.3). Then they selected the slice in the mid sagittal plane to be traced, and proceeded to trace the upper pharyngeal airway. The threshold sensitivity value for the software to discriminate soft tissue from air space was then manually selected and adjusted so that the software completely fills in the airway space, without under or over-filling (Figures 3.4-3.6). After all required parameters were set, the software processed the measurements of the airway. The selected threshold sensitivity value, minimum cross-sectional area, total upper airway volume, nasopharyngeal airway volume, oropharyngeal airway volume, and hypopharyngeal airway volume were then recorded. This process was then repeated with the same scans in reverse order with a 4-week interval between assessment periods. The examiners did not have access to their previous assessments at the second analysis period, and the scans were randomly analyzed to allow for a blinded assessment. The total upper pharyngeal airway and its components can be seen in Figures 3.7-3.11 and the corresponding landmarks in Table 3.1. The determination of the minimum cross-sectional area can be seen in Figure 3.12.

Figure 3.1 Orientation of the axial plane using the lower border of the orbit landmarks

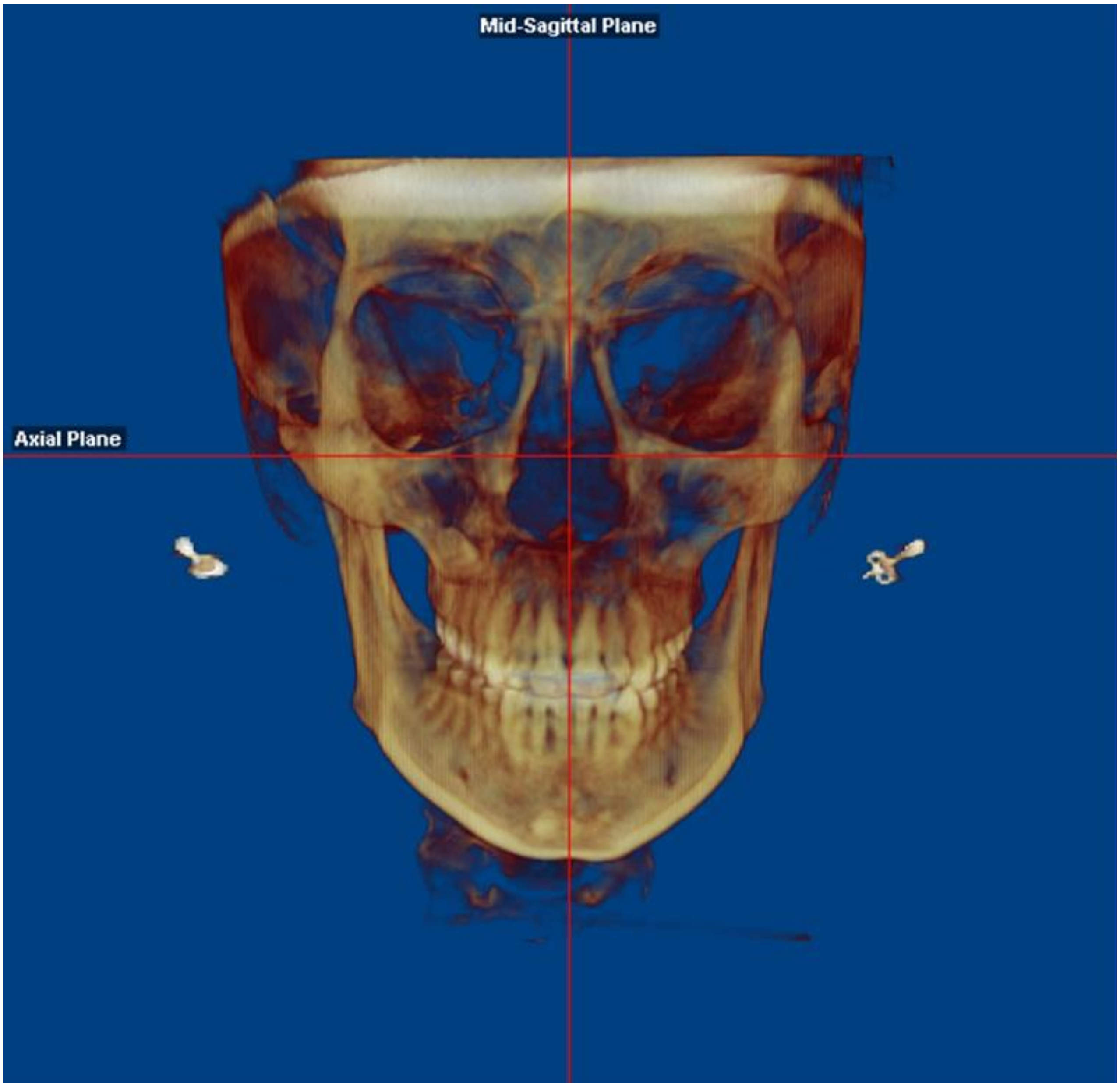


Figure 3.2 Orientation of the coronal plane using the Frankfort Horizontal Plane (porion to orbitale)

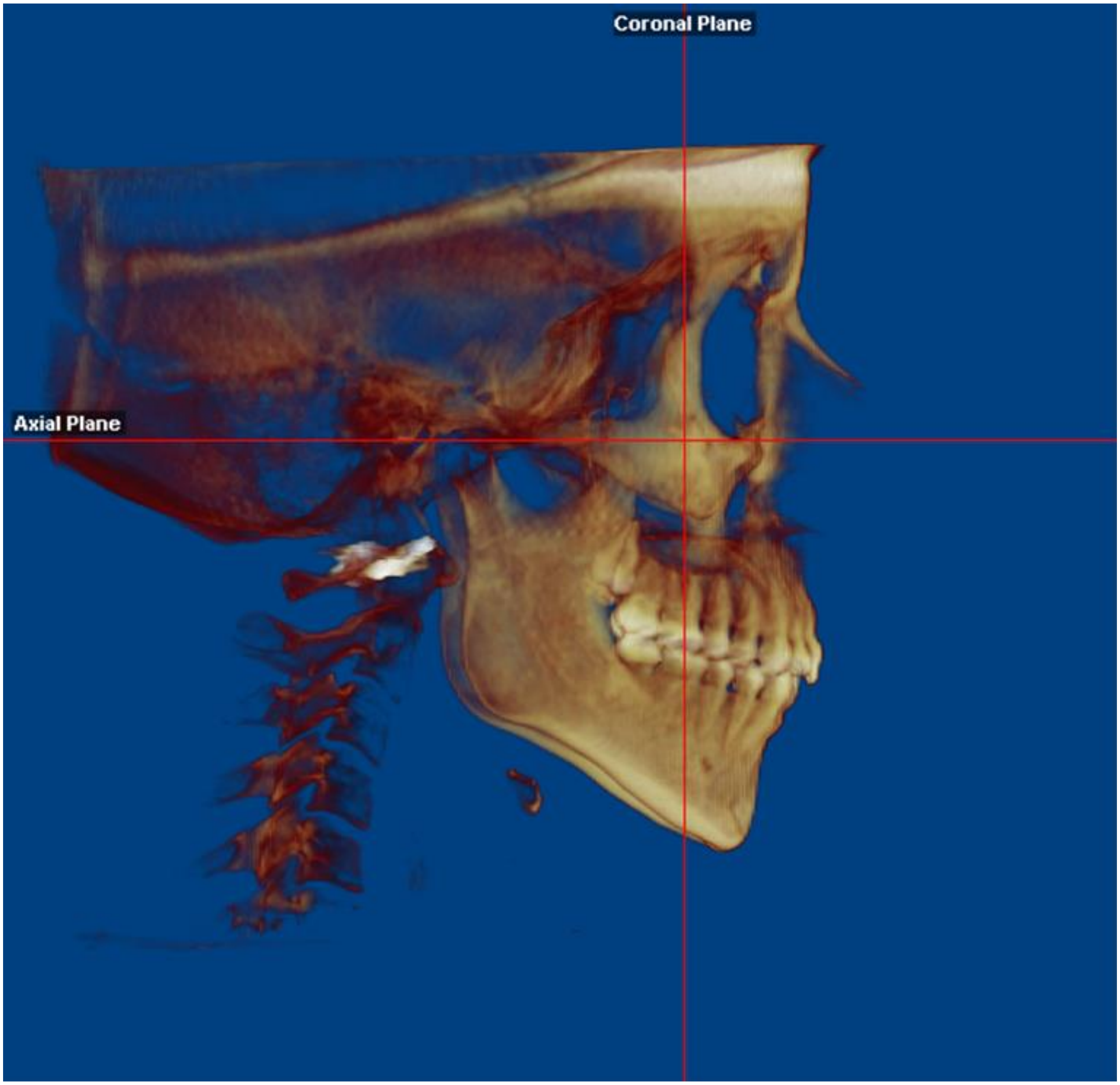


Figure 3.3 Orientation of the midsagittal plane using the upper incisive foramen to opisthion

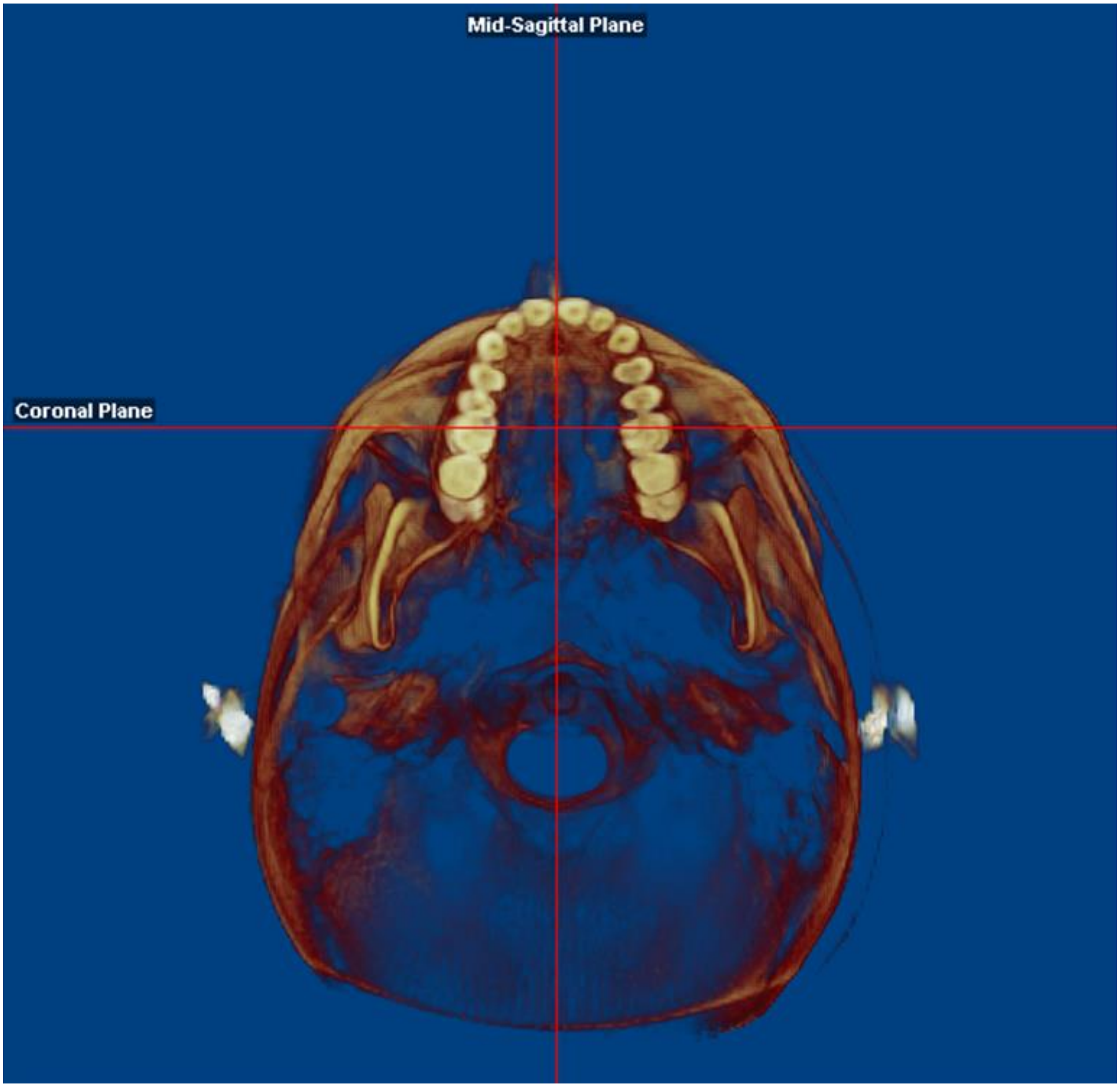


Figure 3.4 Correct sensitivity thresholding where the upper pharyngeal airway is completely filled with no "fingers" projecting out from the airway

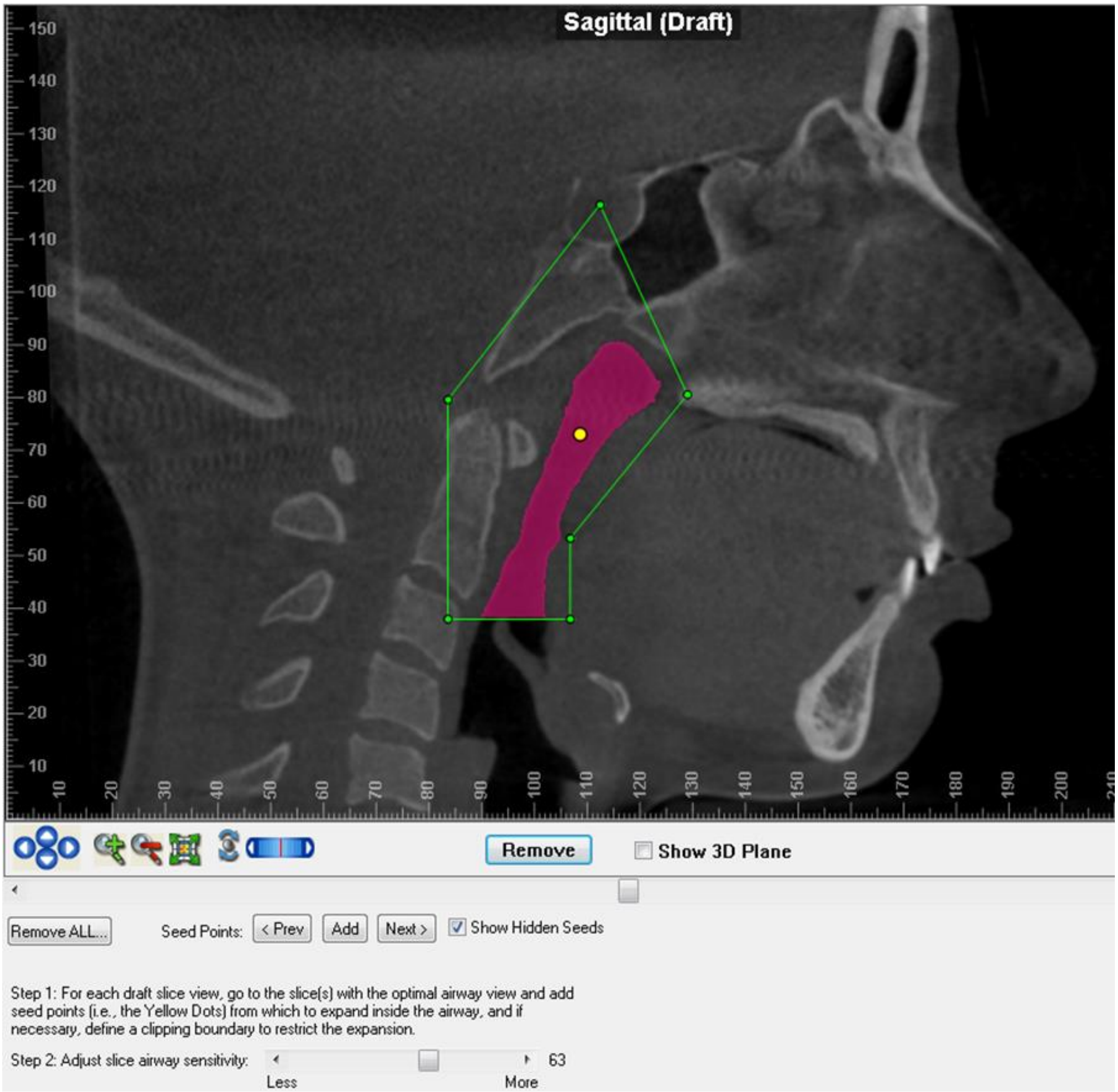


Figure 3.5 Incorrect sensitivity thresholding where the upper pharyngeal airway is under-filled

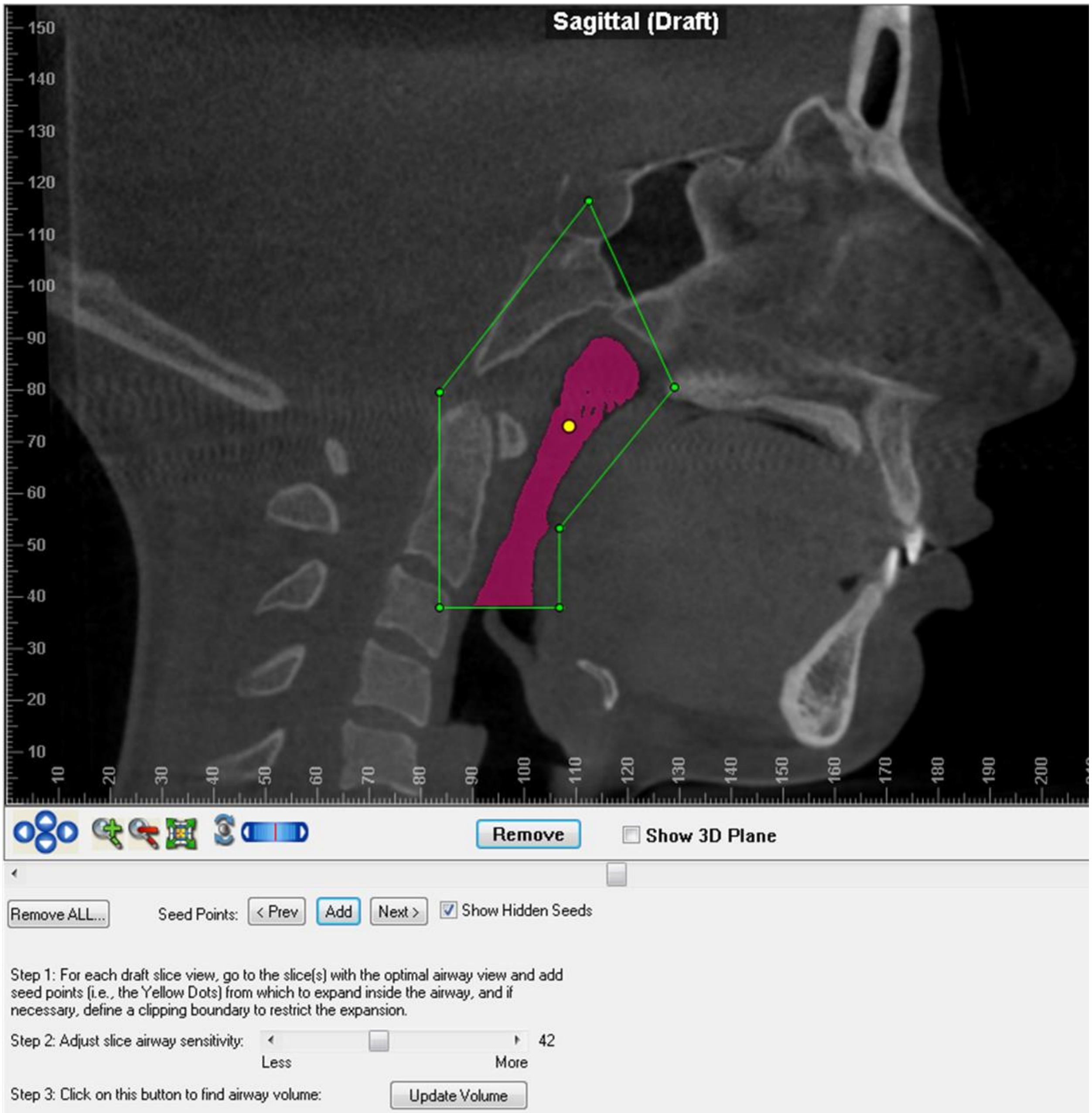


Figure 3.6 Incorrect sensitivity thresholding where the upper pharyngeal airway is over-filled to the point where "fingers" can be seen projecting out of the airway

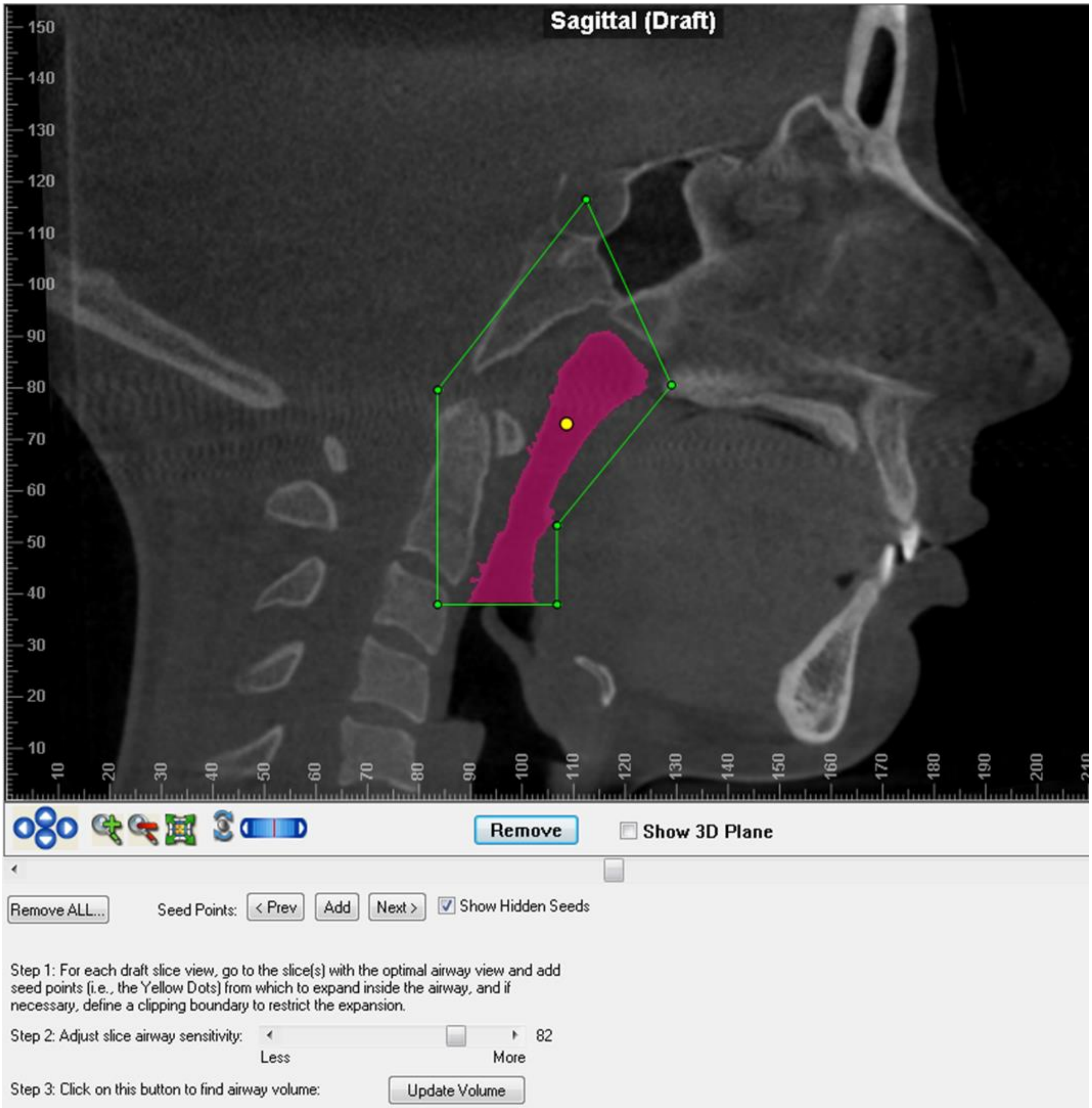


Figure 3.7 Anatomic boundaries of the upper pharyngeal airway including its comprising regions

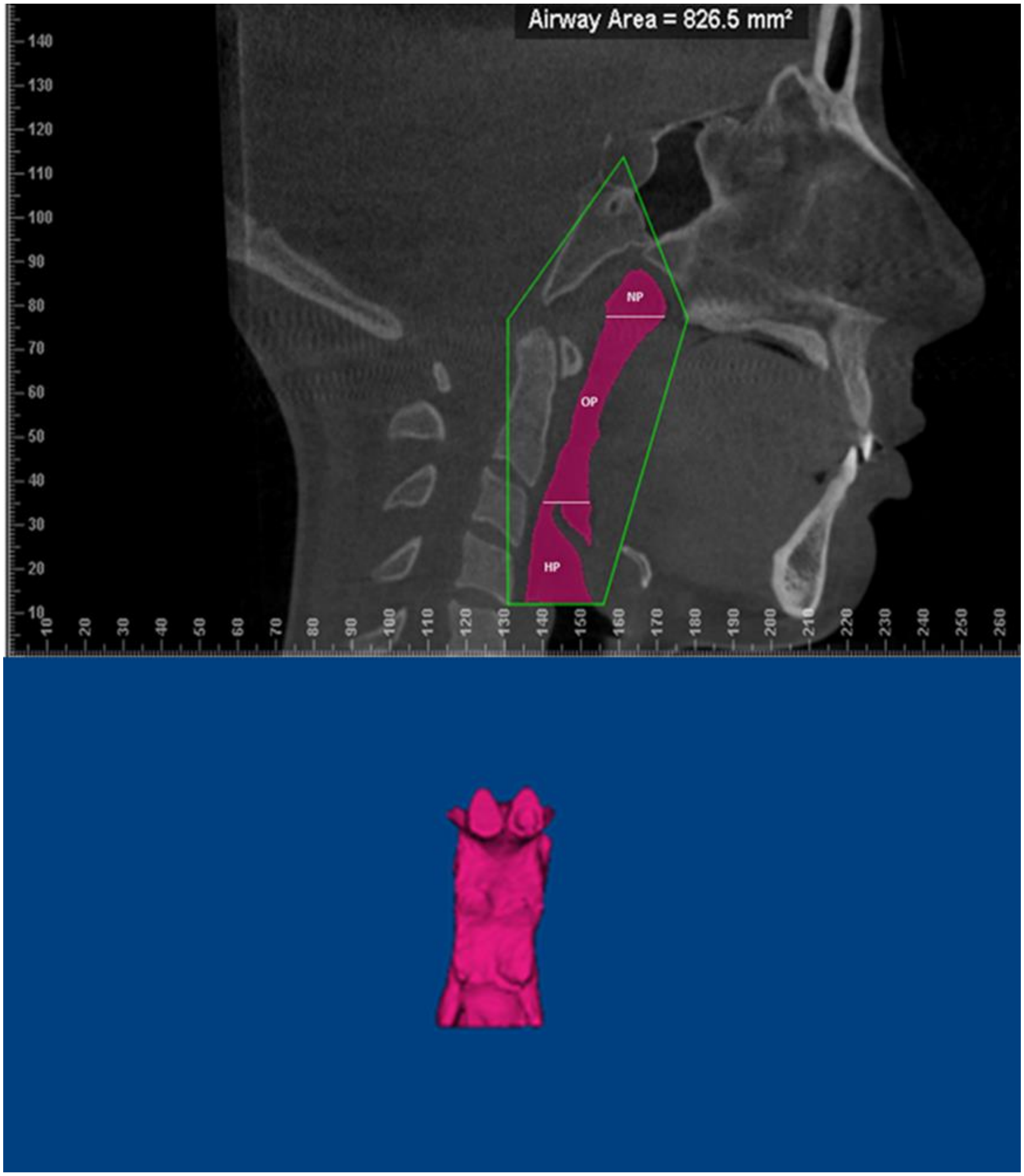


Figure 3.8 Nasopharyngeal airway volume

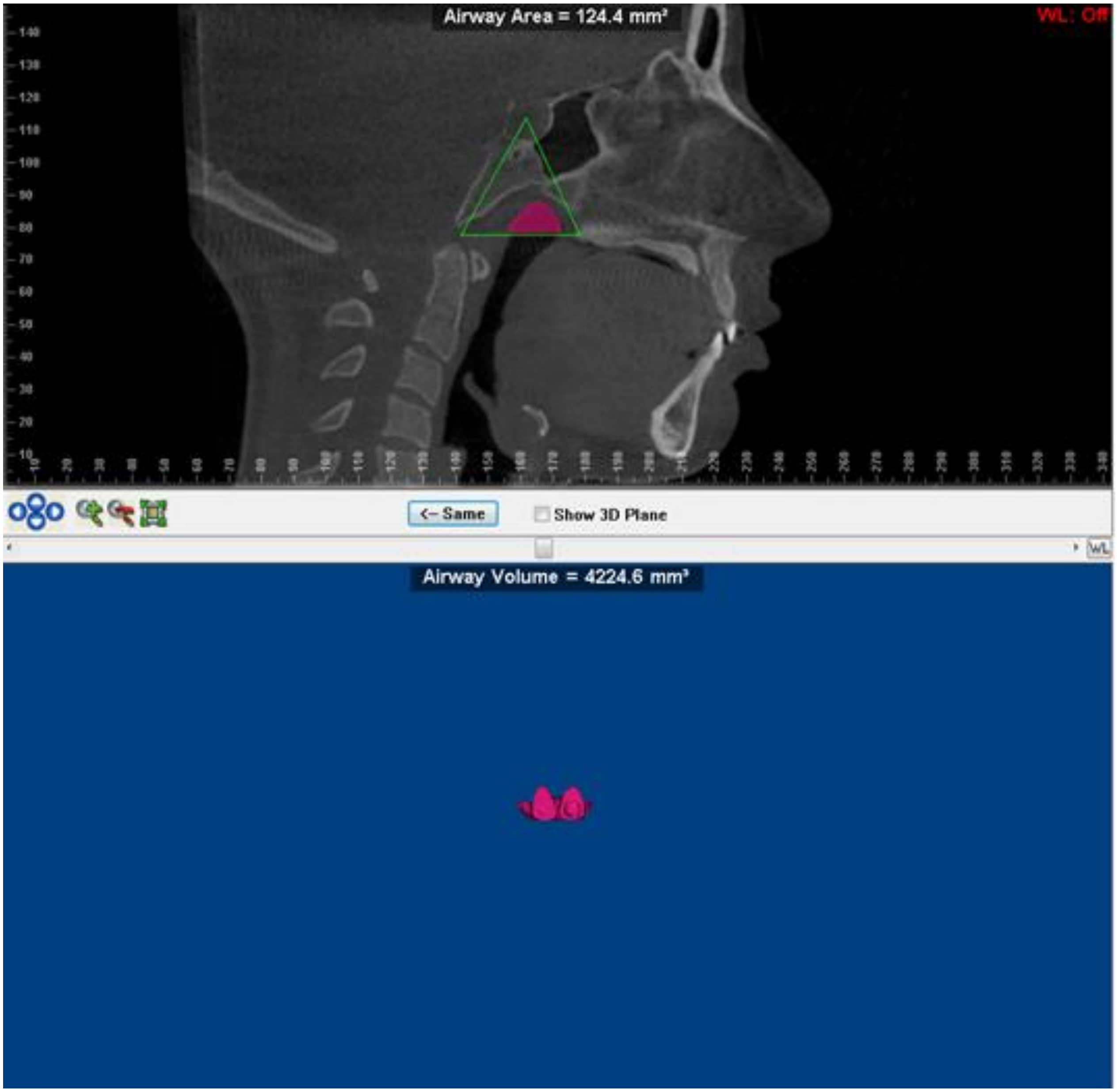


Figure 3.9 Oropharyngeal airway volume

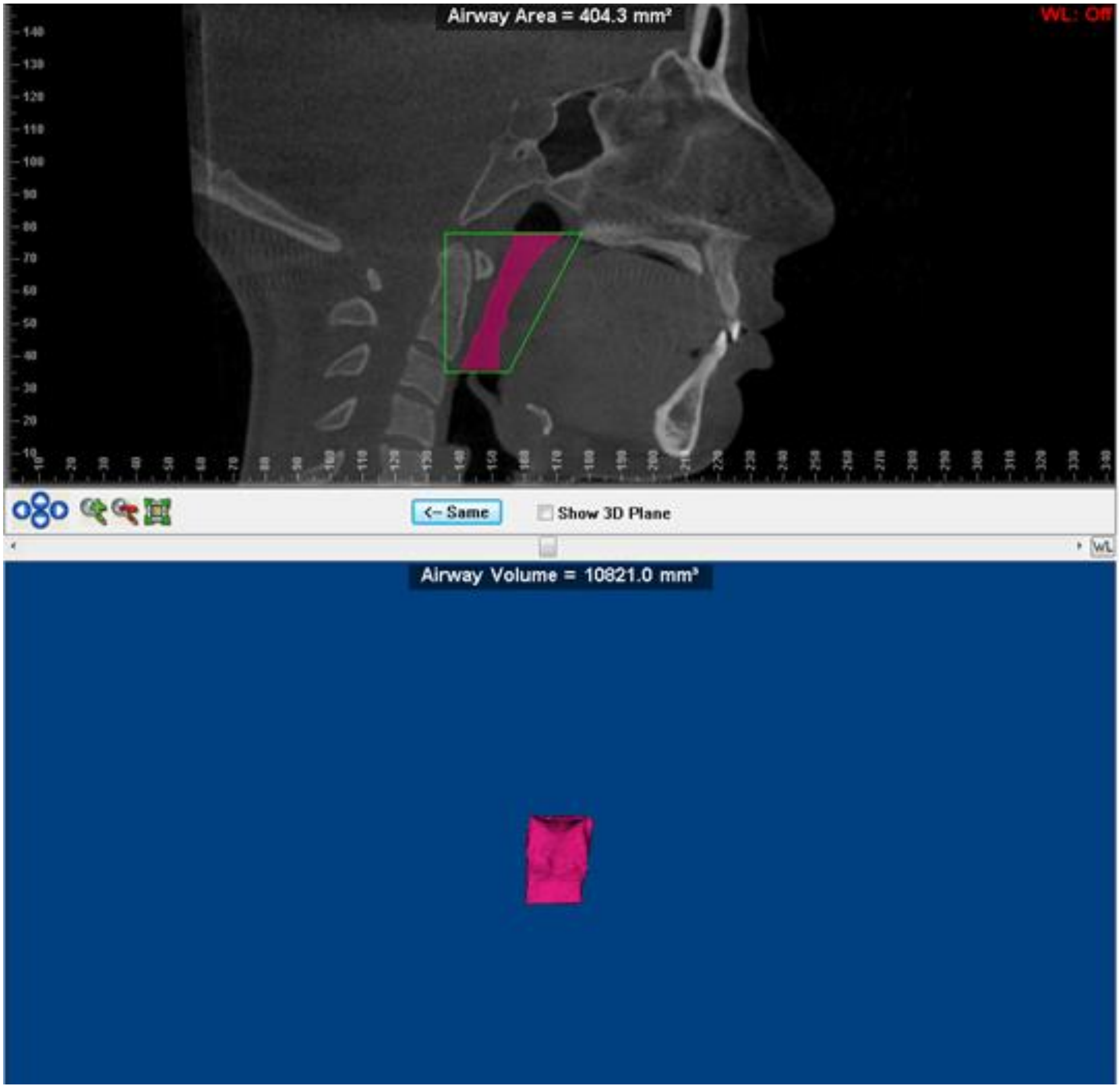


Figure 3.10 Hypopharyngeal airway volume

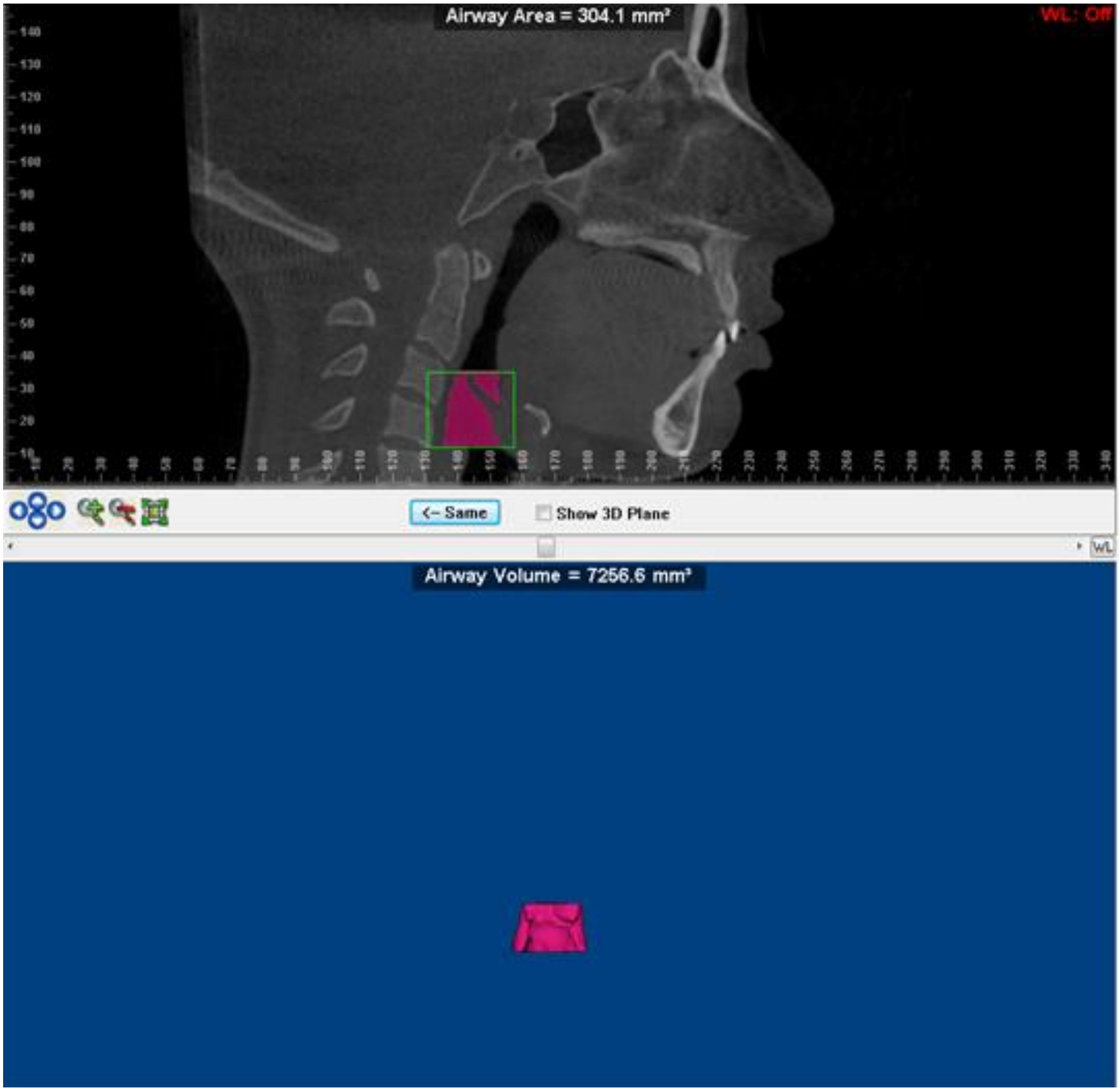


Figure 3.11 Total upper pharyngeal airway volume

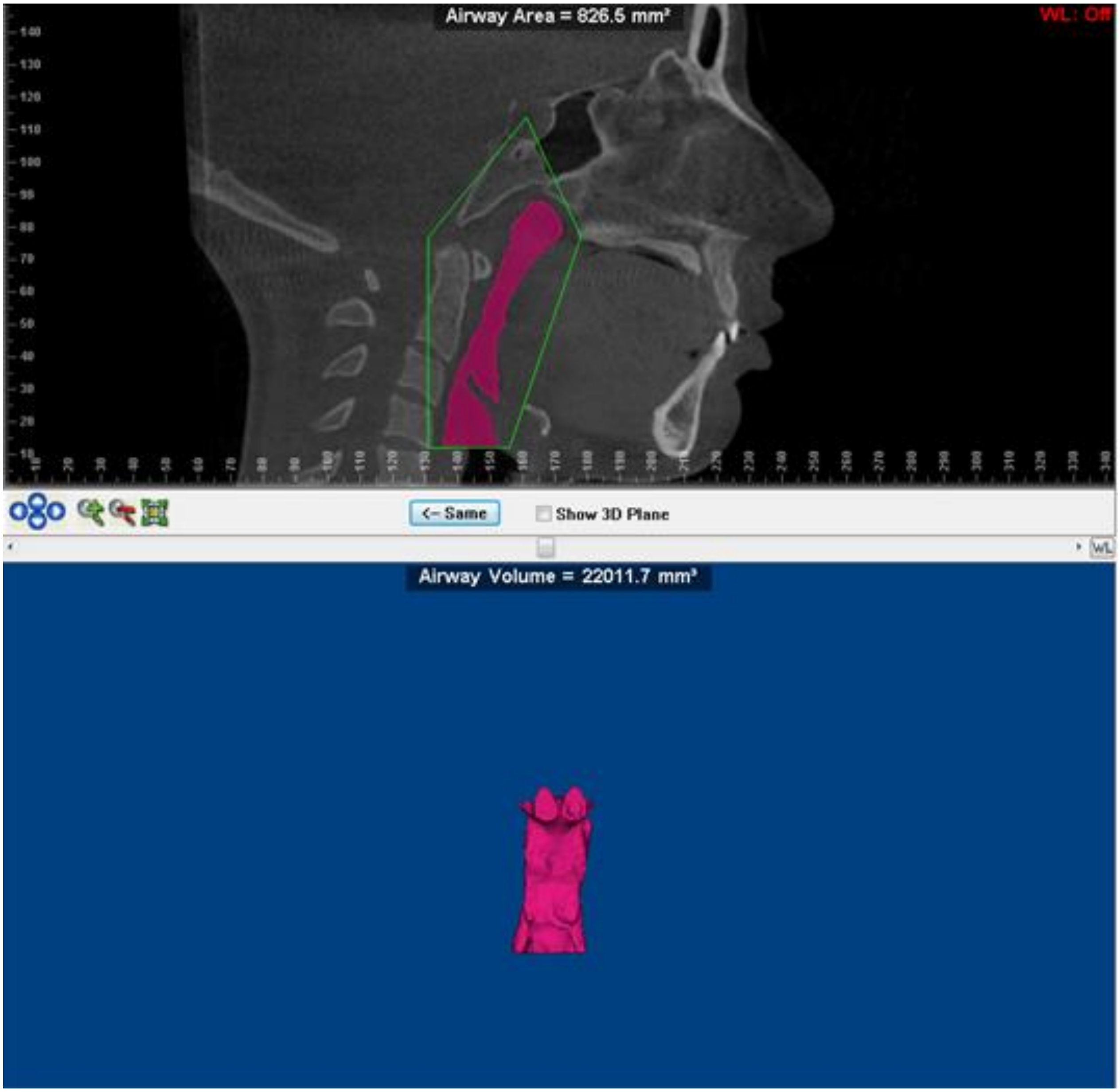


Figure 3.12 Minimum cross-sectional area of the upper pharyngeal airway

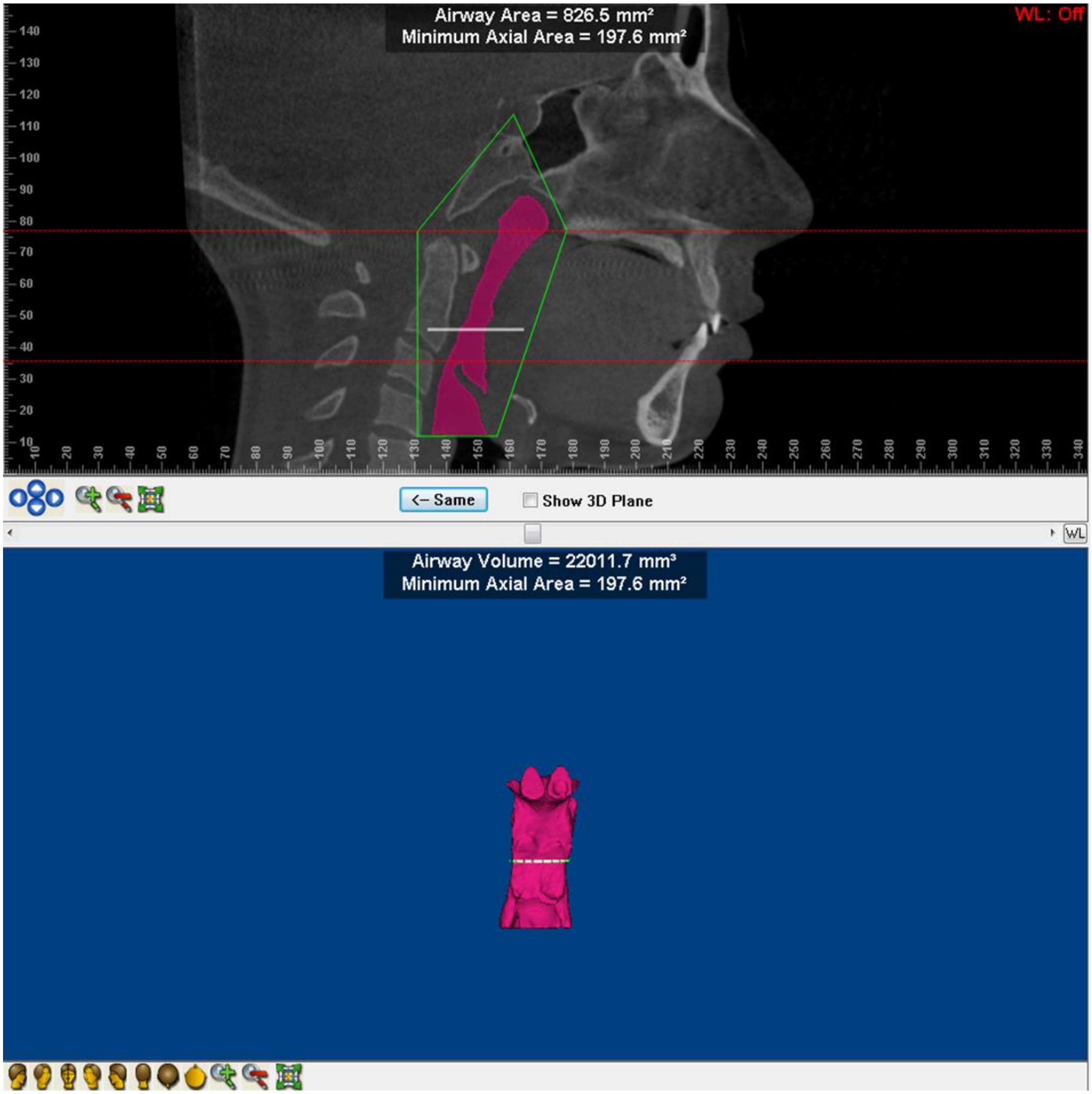


Table 3.1 Definitions of the anatomic boundaries for each region of the upper pharyngeal airway

	Anterior boundary	Posterior boundary	Superior boundary	Inferior boundary
Total Airway	Line extending from Sella to the posterior nasal spine (PNS) to the tip of the epiglottis to the base of the epiglottis and entrance to the esophagus	Line extending from Sella to the superior pharyngeal wall to the inferior pharyngeal wall	Sella point	Line extending from the base of the epiglottis and entrance to the esophagus to the posterior inferior pharyngeal wall
Nasopharynx	Line extending from Sella to the posterior nasal spine (PNS)	Line extending from Sella to the posterior pharyngeal wall	Sella point	Line extending from the posterior nasal spine (PNS) to the posterior superior pharyngeal wall
Oropharynx	Line extending from the posterior nasal spine (PNS) to the tip of the epiglottis	Line extending from the posterior superior pharyngeal wall to the posterior middle pharyngeal wall	Line extending from the posterior nasal spine (PNS) to the posterior superior pharyngeal wall	Line extending from the tip of the epiglottis to the posterior middle pharyngeal wall
Hypopharynx	Line extending from the tip of the epiglottis to the base of the epiglottis and entrance to the esophagus	Line extending from the posterior middle pharyngeal wall to the posterior inferior pharyngeal wall	Line extending from the tip of the epiglottis to the posterior middle pharyngeal wall	Line extending from the base of the epiglottis and entrance of the esophagus to the posterior inferior pharyngeal wall

Intra-examiner and inter-examiner reliability was calculated using ICC for the measurements obtained by each examiner at both assessment periods. Using SPSS version 24 (SPSS Inc, Chicago, IL), ICC values along with 95% confidence interval were also used to assess inter- examiner reliability by comparing their first and second assessments. Reliability was ranked according to the ICC value and considered excellent when it was above 0.9, good when it was between 0.75 and 0.9, moderate when it was between 0.5 and 0.75, and poor when it was below 0.5.[105] In addition, examiner variation was calculated as the absolute value of the difference between the two recordings made for each parameter. Median examiner variation along with quartiles 1 and 3, as well as the mean examiner variation as a percentage of the mean values were calculated for each parameter. Furthermore, the method error using Dahlberg's formula was calculated using the examiner with the highest ICC for each parameter.

3.3 Results

Intra-examiner and inter-examiner reliabilities estimated by ICC for each parameter are shown in Table II for all 10 scans. The selection of threshold sensitivity value showed poor intra-examiner (mean ICC 0.473) and poor inter-examiner (ICC 0.100; CI 0.000-0.380) reliability. Minimum cross-sectional area showed moderate intra-examiner (mean ICC 0.591) and poor inter-examiner (ICC 0.223; CI 0.029-0.581) reliability. Total airway volume showed good (mean ICC 0.819) and poor inter-examiner (ICC 0.175; CI 0.000-0.533) reliability. Nasopharyngeal airway volume showed good intra-examiner (mean ICC 0.777) and poor inter-examiner (ICC 0.350; CI 0.124-0.690) reliability. Oropharyngeal airway volume showed excellent intra-examiner (mean ICC 0.976) and excellent inter-examiner (ICC 0.945; CI 0.849-0.985) reliability. Lastly, hypopharyngeal airway volume showed moderate intra-examiner (mean ICC 0.747) and moderate inter-examiner (ICC 0.550; CI 0.297-0.822) reliability.

However it should be noted that intra-examiner reliability varied greatly with education and experience level as seen in the difference between the minimum and maximum ICC for each parameter. Intra-examiner reliability for threshold sensitivity value ranged from 0.260-0.741, minimum cross-sectional area from 0.000-0.983, total airway volume from 0.160-0.992, nasopharyngeal airway volume from 0.279-0.979, oropharyngeal airway volume from 0.930-0.996, and hypopharyngeal airway volume from 0.679-0.811. The more educated and experienced examiners generally showed considerably higher intra-examiner reliability. Inter-examiner reliability also greatly increased with more educated and experienced examiners for most parameters as seen in Table 3.2.

Tables 3.3 and 3.4 highlight the differences in intra-examiner and inter-examiner reliabilities between the fast and slow scan protocols respectively. The slow scan protocol demonstrated generally a higher intra-examiner reliability than the fast scan protocol. However, the differences between the two protocols was relatively minor for intra-examiner reliability compared to inter-examiner reliability. The slow scan protocol displayed considerably higher inter-examiner reliability compared to the fast scan protocol.

The median examiner variation is shown in Table 3.5 along with first and third quartiles. To further represent the observer error in our study we also calculated the mean examiner variation as a percentage of the mean values obtained in each parameter, also shown in Table 3.5. Table 3.5 also includes the method error using Dahlberg's formula which was calculated using the examiner with the highest ICC for each parameter. The examiner with the highest ICC was used to provide the best case scenario.

Also shown in Table 3.5 is the range of measured values for each parameter. Threshold value ranged from 44-82, minimum cross-sectional area from 67.90-1960.30 mm^2, total upper airway volume from 17433.70-217481.90 mm^3, nasopharyngeal airway volume from 3216.50-17922.60 mm^3, oropharyngeal airway volume from 6985.20-40242.30 mm^3, and hypopharyngeal airway volume from 1949.80-11835.00 mm^3. The raw data can be found in Tables A.4-A.9. The volumetric and cross-sectional data from this study is relatively consistent with the previous literature.[94]

Table 3.2 ICC values for intra-examiner and inter-examiner reliability for all scans

	Intra-Examiner Reliability							Inter-Examiner Reliability			
	Op A	Op B	Op C	Op D	Op E	Op F	Mean ICC	Ops A,B,C,D,E,F	95% CI	Ops A,B,C,D	Ops E,F
Threshold Value	0.358	0.690	0.260	0.533	0.254	0.741	0.473	0.100	0.000, 0.380	0.501	0.059
Minimum Cross-Sectional Area	0.928	0.983	0.818	0.898	0.124	0.000	0.591	0.223	0.029, 0.581	0.868	0.116
Total Airway Volume	0.992	0.987	0.967	0.991	0.160	0.819	0.819	0.175	0.000, 0.533	0.956	0.107
Nasopharyngeal Airway Volume	0.954	0.979	0.874	0.976	0.602	0.279	0.777	0.350	0.124, 0.690	0.827	0.228
Oropharyngeal Airway Volume	0.996	0.993	0.983	0.965	0.988	0.930	0.976	0.945	0.849, 0.985	0.985	0.950
Hypopharyngeal Airway Volume	0.810	0.811	0.729	0.747	0.706	0.679	0.747	0.550	0.297, 0.822	0.517	0.663

Op A = Oral and maxillofacial radiologist

Op B = Academic orthodontists with additional study in airway and sleep apnea

Op C = Private practice orthodontist

Op D = Academic orthodontist

Op E = Senior orthodontic resident

Op F = Junior orthodontic resident

Table 3.3 ICC values for intra-examiner and inter-examiner reliability for the fast scan protocol

	Intra-Examiner Reliability							Inter-Examiner Reliability			
	Op A	Op B	Op C	Op D	Op E	Op F	Mean ICC	Ops A,B,C,D,E,F	95% CI	Ops A,B,C,D	Ops E,F
Threshold Value	0.000	0.703	0.531	0.000	0.000	0.897	0.245	0.000	0.000, 0.156	0.459	0.000
Minimum Cross-Sectional Area	0.873	0.996	0.956	0.988	0.104	0.114	0.672	0.152	0.000, 0.732	0.986	0.106
Total Airway Volume	0.983	0.979	0.958	0.984	0.167	0.767	0.806	0.142	0.000, 0.727	0.924	0.127
Nasopharyngeal Airway Volume	0.925	0.989	0.929	0.921	0.862	0.386	0.835	0.187	0.000, 0.754	0.798	0.202
Oropharyngeal Airway Volume	0.993	0.98	0.951	0.985	0.962	0.922	0.966	0.908	0.712, 0.988	0.974	0.905
Hypopharyngeal Airway Volume	0.598	0.495	0.846	0.939	0.461	0.674	0.669	0.489	0.145, 0.903	0.421	0.645

Op A = Oral and maxillofacial radiologist

Op B = Academic orthodontists with additional study in airway and sleep apnea

Op C = Private practice orthodontist

Op D = Academic orthodontist

Op E = Senior orthodontic resident

Op F = Junior orthodontic resident

Table 3.4 ICC values for intra-examiner and inter-examiner reliability for the slow scan protocol

	Intra-Examiner Reliability							Inter-Examiner Reliability			
	Op A	Op B	Op C	Op D	Op E	Op F	Mean ICC	Ops A,B,C,D,E,F	95% CI	Ops A,B,C,D	Ops E,F
Threshold Value	0.609	0.653	0.000	0.748	0.673	0.620	0.516	0.291	0.045, 0.809	0.551	0.150
Minimum Cross-Sectional Area	0.946	0.980	0.763	0.872	0.994	0.000	0.704	0.824	0.552, 0.976	0.831	0.849
Total Airway Volume	0.996	0.991	0.975	0.996	0.996	0.859	0.969	0.917	0.739, 0.990	0.976	0.870
Nasopharyngeal Airway Volume	0.974	0.953	0.765	0.991	0.444	0.038	0.694	0.652	0.311, 0.945	0.814	0.500
Oropharyngeal Airway Volume	0.996	0.997	0.993	0.965	0.995	0.936	0.980	0.958	0.852, 0.995	0.988	0.963
Hypopharyngeal Airway Volume	0.966	0.902	0.540	0.531	0.936	0.764	0.773	0.678	0.314, 0.950	0.721	0.619

Op A = Oral and maxillofacial radiologist

Op B = Academic orthodontists with additional study in airway and sleep apnea

Op C = Private practice orthodontist

Op D = Academic orthodontist

Op E = Senior orthodontic resident

Op F = Junior orthodontic resident

Table 3.5 Examiner variance for all parameters

	Units	Mean value	Range of data	Median observer variance	Q1	Q3	Mean observer variance as percent of the mean value (%)	Method Error Using Dahlberg's formula
Threshold	N/A	58.30	44-82	2	1	4	5.34	2.46
Minimum Cross-Sectional Area	mm^2	260.10	67.90-1960.30	12.10	6.10	61.15	27.23	15.56
Total Airway Volume	mm^3	31277.80	17433.70-217481.90	1100.55	429.28	2635.08	15.09	784.20
Nasopharyngeal Airway Volume	mm^3	6159.90	3216.50-17922.60	416.00	193.22	785.02	12.86	225.16
Oropharyngeal Airway Volume	mm^3	18213.40	6985.20-40242.30	730.00	248.30	1335.52	6.24	542.44
Hypopharyngeal Airway Volume	mm^3	5972.90	1949.80-11835.00	710.15	277.20	1427.60	17.81	730.58

N/A = Not applicable

3.4 Discussion

The recent systematic review of the literature on this area of research revealed that there were significant methodological limitations in previous assessments of upper airway anatomy using CBCT imaging.[75] More specifically, none of the available studies allowed for the manual orientation, mid-sagittal plane slice selection of the CBCT images and selection of threshold sensitivity by the examiners in the study protocols, despite the fact that these steps are fraught with subjectivity and have the potential to affect reliability. This is the first study to determine the reliability of upper airway assessment using CBCT which considers the above limitations, combined with examiner experience/qualifications and fast versus slow scan protocols.

Overall, the oropharynx is the only region of the upper pharyngeal airway to exhibit excellent intra-examiner and inter-examiner reliability. This was independent of examiner education and experience, and selected threshold sensitivity value. This is consistent with previous studies by El et al.[89], and Guijarro-Martinez et al.[94] showing that the oropharynx was the region with the highest reliability. Potential explanations can include that the nasopharynx and hypopharynx are either more sensitive to threshold selection which in itself has poor reliability, or that landmark identification for these regions are more challenging. Alsufyani et al.[78] provides another possible explanation in that the shape of the oropharynx three-dimensionally is essentially similar to that of a tube, being completely hollow. This allows for relatively straight-forward segmentation and processing by the imaging software. However, the anatomy of the nasopharynx is more complicated due to the narrow and tortuous pathways of the eustachian tubes and choanae. The same can be said about the hypopharynx due to the presence of the epiglottis. This combined with potentially noisy CBCT images results in an extremely challenging segmentation process, owing to difficulties encountered in defining the boundaries and grey level thresholding. They further conclude that studies which only focus on the oropharyngeal airway will likely over-represent the reliability of the evaluated tools.[78]

Selection of threshold sensitivity value for the airway displayed poor intra-examiner and poor inter-examiner reliability. Previously, Alves et al.[121] conducted a study to determine the optimal threshold value on Dolphin Imaging software to measure airway volume. They reported that a threshold value of 73 was most accurate, and that values of 70, 71, 72, 74, and 75 had no statistically significant differences in measurement outcomes. This study however was conducted using airway replicas of only the oropharynx made of silicone to determine the optimal threshold value, which can likely over-estimate the reliability as previously stated by Alsufyani.[78] Furthermore it is clear from the current study and others[85,91] that threshold values for segmenting the airway in silicone models may have little applicability to the values required in scans of actual patients.

It is interesting to note that selection of the threshold sensitivity value showed poor reliability even amongst the educated and experienced examiners, but their intra-examiner reliability in the other parameters was still relatively high. This could mean that threshold selection may not have a major effect on reliability, but maybe threshold selection combined with the manual orientation and mid-sagittal plane slice selection all play minor roles that when combined can have a more significant effect on reliability, especially with less experienced examiners.

The slow scan protocol generally displayed higher reliability than the fast scan protocol, however this trend was much more pronounced for inter-examiner reliability than for intra-examiner reliability. This could be explained in that the increased scan time, decreased voxel size, and increased tube current provided for greater resolution in the CBCT image.[78] However increasing scan time is not always desirable. Firstly the slow scan protocol comes at an associated cost with an increased radiation dosage, with the slow scan protocol having an effective dose of 127.3 µSv whereas the fast scan protocol has an effective dose of only 64.7 µSv.[7] A further limitation is if the scan time is too long then the patient can undergo multiple breathing cycles. This can result in some motion artifact which can affect the resolution of the airway boundaries.[78]

The protocol of this study mimics the upper airway assessment process as would be performed in a clinical setting. Random human error is inevitably introduced with each manual step, thereby affecting reliability. Perhaps the most noteworthy finding from this study is that the intra-examiner and inter-examiner reliability for all parameters were lower than previously reported in the literature.[75] This can be due to the fact that in this study the examiners had to perform each step of the assessment process manually, whereas the previous studies essentially only assessed the ability of the examiners to reliably trace the upper pharyngeal airway. Between manual orientation, mid-sagittal plane slice selection, and selection of the threshold sensitivity value, these are all steps in the assessment process that introduce an element of subjectivity and are therefore burdened with potential to introduce error. To demonstrate the magnitude of this error, inter-observer error was presented as the median observer error along with the first and third quartiles for each parameter as seen in previous studies.[125,126] The mean inter-observer difference as a percentage of the average values obtained in each parameter was also provided. Indeed, in CBCT studies which report changes in airway anatomy less than the values of mean percentage error presented in Table V (6% for oropharynx, 27% for minimal cross sectional area), this may in fact be due to measurement error rather than treatment effect.

It is clear from this study that education and experience level of the examiner has a significant effect on both intra-examiner and inter-examiner reliability. The findings are positive in that the examiners which demonstrated the greatest reliability are those who would be readily assessing the upper pharyngeal airway of patients in the clinical setting. However, the reliability displayed by the residents was significantly poorer. This is important because it is not uncommon for orthodontic residents to be the examiners in CBCT research as they are readily available in academic institutions.[92,112,114]

Overall, the results of the present study raise questions towards the value of quantitative assessments of the upper airway using CBCT imaging when using this common measurement protocol. While excellent reliability in the oropharyngeal region was found, the inter-examiner reliability of measurements of both volume of the nasopharynx and overall minimal cross sectional area was poor. As shown in Table 3.5, inter-observer differences can range upwards of 27% of the measured value, which should be taken in to consideration when changes in airway dimensions are being assessed. This has direct implications for associations with sleep disordered breathing as minimum cross-sectional area is a crucial measure of flow limitation and airway collapsibility,[127] and the nasopharynx is the most common area of obstruction with children with obstructive sleep apnea.[128]

A limitation of this study was that there could have been a greater number of scans for both the fast and slow scan protocols to allow for a more substantial assessment of reliability between these two imaging protocols. This can be difficult as more time would be required by the examiners, but this would be an important area of future research as the current literature does not address the topic of optimizing scan protocols to increase reliability while reducing the radiation dose to the patient. A further limitation of this study is that only one examiner of each level of experience or training were included, and that the sample was comprised of only adult patients. A future study with multiple examiners of each experience level and the inclusion of pediatric patients would provide a better understanding of how reliability is effected by these factors. Furthermore, the assessments of the CBCT scans were not performed on a greyscale monitor, and a future study which does would improve the available evidence. What is clear from the findings of this study is that in any future research assessing the upper airway using CBCT and reporting reliability, examiners must perform all steps in the assessment process manually as clinicians would in a clinical setting and this should be reported. Furthermore, any studies measuring changes in airway volume and/or minimum cross-sectional area should also report whether the differences found are above the range of errors introduced by the measurement protocols.

3.5 Conclusions

This is the first study to evaluate the reliability of upper pharyngeal airway assessment using CBCT where the examiners performed each step of the analysis manually, as would be conducted in a clinical setting. Selection of the threshold sensitivity value generally had poor reliability. Reliability improved with examiner experience, though was generally low for the hypopharynx and nasopharynx volumes and overall minimal cross sectional area. The oropharyngeal airway volume was the only parameter found to have generalized excellent intra-examiner and inter-examiner reliability.

Chapter 4: Should Dental CBCT Be Used Today For Quantitative Assessments of the Upper Pharyngeal Airway: Final Thoughts

This is the first study to evaluate the reliability of upper pharyngeal airway assessment using dental CBCT where the examiners performed each step of the analysis manually. Selection of the threshold sensitivity value generally had poor reliability. Reliability greatly improved with education and experience level of the examiner. Volumetric assessments demonstrated greater reliability than did minimum cross-sectional area, with oropharyngeal airway volume being the only parameter to have generalized excellent intra-examiner and inter-examiner reliability. The slow scan protocol generally showed greater reliability with a greater effect on inter-examiner reliability. However further research is necessary to make more definitive assertions about the effect of scan protocol on reliability.

Even once reliability is adequately established, this is not sufficient evidence to support the use of CBCT by clinicians to assess a patient's upper airway to diagnose OSA. Validity of CBCT to determine the true volumetric and cross-sectional area measurements of a patient's airway must then be evaluated, and this is fraught with confounding factors.

The primary confounding factor for CBCT studies assessing airway is head, body, and jaw position at the time of scan acquisition as they can have a large influence on the upper airway dimension. A non-randomized controlled trial study by Ono et al.[129] studied how changes in head/body position induce changes in upper-airway dimensions specifically related to three positions, supine, supine with the head rotated and lateral recumbent. They demonstrated a significant increase in volume in the retro-glossal region of oropharynx when subjects rotated their head to the left in the supine position and when changing from the supine to the lateral recumbent position.

Another non-randomized controlled trial study by Pirilä-Parkkinen et al.[130] compared the pharyngeal airway size in different cranio-cervical postures in children with sleep-disordered breathing (SDB) and asymptomatic control children who were age and gender matched. The upper airway in both groups were evaluated in neutral, extension, and flexion head positions. The hypopharyngeal airway in the SDB group increased by head extension compared to natural head position, and this increase was higher for the SDB group than in the asymptomatic group.

An additional non-randomized controlled trial study by Zhang et al.[131] investigated the effect of head and body positions on the oropharynx caliber in normal subjects when their jaw was protruded by using magnetic resonance imaging. Four different jaw, head and body positions were assessed: jaw protrusion, supine with jaw protrusion, supine-head rotation with jaw protrusion and lateral decubitus with jaw protrusion. The subjects in this study displayed no sign of breathing-related disorders. They found that jaw protrusion increased the volume of oropharynx at the level of the retro-palatal- and the retro-glossal regions compared with non-protruded positions.

Moreover, according to a systematic review on the effect of head and tongue posture on pharyngeal airway dimensions and morphology conducted by Gurani et al.[132], altered head, body and jaw position, respectively had a significant effect on the upper airway dimensions and volume at the time of image acquisition. The oropharyngeal airway and specifically the retro-palatal and retro-glossal regions of the oropharynx, were the most affected portions of the upper airway when evaluated in respect to head rotation, head extension, jaw protrusion and altered body position. Both volume and cross-sectional area showed an increase when evaluated in respect to head extension, head rotation, altered body position, and jaw protrusion. However, they stated that only limited and poor quality evidence was available since no validated method existed with regard to the position of head, jaw or body at the time of image acquisition. Therefore they concluded that higher levels of evidence was needed and future studies require a standardized method of head and tongue posture during image acquisition.

A study by Guijarro-Martinez and Swennen[94] states that other confounding factors for upper airway analysis with CBCT include respiratory phase and tongue posture during image acquisition, as they can qualitatively and quantitatively affect the size and shape of the oropharyngeal airway. To control these variables, it is suggested that the patient should be instructed to avoid swallowing and any other movement during the CBCT scan, breathe gently, and maintain the mandible in a reproducible position, either maximum intercuspation or centric relation.[94] As scanning technology improves and scan acquisition time decreases it will become much easier to control these variables.

In order to quantify the effect of patient body positioning during CBCT airway examination, Camacho et al.[133] conducted a retrospective study describing how total volume and cross-sectional area measurements change in OSA patients associated with a supine versus an upright position. They found that the airway was smaller when patients were in a supine compared with an upright position. Not only was a decrease seen in total airway volume but also a decrease in cross-sectional area was observed at the levels of the posterior nasal spine, uvula tip, retrolingual and tongue base. Minimum cross-sectional area of the overall airway was also decreased in the supine position compared to the upright position. Total airway volume decreased by 32.6% and cross-sectional area measurements decreased between 32.3% and 75.9% when patients were in a supine position. They concluded that the airway of OSA patients was significantly smaller when they were in a supine compared with an upright position. This can be problematic because in a clinical setting, CBCT assessments of the airway are generally taken with the patient in the upright position, potentially providing a false impression of the patient's airway dimensions while sleeping.

As the scans used in our study came from a bank of scans from UMN, it is unknown whether or not the above confounding factors were considered at the time of image acquisition. However, as the scans were selected at random for the assessment of examiner reliability, these confounding factors do not play a significant role in this study. The body mass index of the patients included in this study was not recorded by UMN and it is currently unknown how high levels of obesity as is often found in OSA patients may affect the reliability of measurements. However it will be imperative for future validity studies to take the above factors into account when establishing protocol and methodology. Unless the described issues are accounted for in future studies, quantitative assessments of patients' upper pharyngeal airway volume and minimum cross-sectional area using dental CBCT may indeed be meaningless.

There is a trend in orthodontics to use quantitative data of patients' airways pre and post-treatment to determine the effects of a particular intervention on the airway dimensions. Not only does this study indicate that these conclusions should not be made by clinicians based on dental CBCT imaging, but this also begs the question as to whether or not the airway volume and/or minimum cross-sectional area can be directly related to an individual's susceptibility to OSA. A group of studies, one by Barrera and another by Cheng, used MRI to determine how the airways of OSA versus healthy patients respectively behave.[134,135] What the combination of studies found was that in healthy patients, especially those with increased BMI, increased age, and smaller airways, they physiologically compensated for these anatomical risk factors for airway collapse by actively dilating their airways during inspiration via increased activity of the genioglossus muscle. In patients with OSA, this compensation did not occur. Therefore, quantitative airway dimensions may not play as significant role in the development of OSA compared to how the patients physiologically compensate for their anatomical risk factors. Ultimately the static dimensions of the airway as measured in an upright and awake patient in a CBCT scan may have little to no correlation with how the airway functions during sleep in any particular patient.

4.1 Conclusion

In conclusion, our data on reliability and the associated confounding factors with establishing validity of upper pharyngeal airway assessment suggests that CBCT might be reserved as a qualitative tool to evaluate the airway rather than a quantitative one. What is clear from this research is that further studies are required before CBCT can be advocated valid and reliable comparisons in upper airway dimensions either between patients or within an individual at different points in time.

Bibliography

1. MacDonald D. *Oral and Maxillofacial Radiology: A Diagnostic Approach*. West Sussex: Wiley-Blackwell; 2011.

2. Scarfe WC, Farman AG. What is Cone-Beam CT and How Does it Work? *Dent Clin North Am*. 2008;52(4):707-730. doi:10.1016/j.cden.2008.05.005.

3. Hatcher DC. Operational principles for cone-beam computed tomography. *J Am Dent Assoc*. 2010;141 Suppl(October):3S-6S. doi:10.14219/jada.archive.2010.0359.

4. Ballrick JW, Palomo JM, Ruch E, Amberman BD, Hans MG. Image distortion and spatial resolution of a commercially available cone-beam computed tomography machine. *Am J Orthod Dentofac Orthop*. 2008;134(4):573-582. doi:10.1016/j.ajodo.2007.11.025.

5. Razi T, Niknami M, Alavi Ghazani F. Relationship between Hounsfield Unit in CT Scan and Gray Scale in CBCT. *J Dent Res Dent Clin Dent Prospects*. 2014;8(2):107-110. doi:10.5681/joddd.2014.019.

6. De Vos W, Casselman J, Swennen GRJ. Cone-beam computerized tomography (CBCT) imaging of the oral and maxillofacial region: A systematic review of the literature. *Int J Oral Maxillofac Surg*. 2009;38(6):609-625. doi:10.1016/j.ijom.2009.02.028.

7. Grünheid T, Kolbeck Schieck JR, Pliska BT, Ahmad M, Larson BE. Dosimetry of a cone-beam computed tomography machine compared with a digital x-ray machine in orthodontic imaging. *Am J Orthod Dentofac Orthop*. 2012;141(4):436-443. doi:10.1016/j.ajodo.2011.10.024.

8. Ludlow JB, Davies-Ludlow LE, White SC. Patient Risk Related to Common Dental Radiographic Examinations. *J Am Dent Assoc*. 2008;139(9):1237-1243. doi:10.14219/jada.archive.2008.0339.

9. Ludlow JB, Ivanovic M. Comparative dosimetry of dental CBCT devices and 64-slice CT for oral and maxillofacial radiology. *Oral Surgery, Oral Med Oral Pathol Oral Radiol Endodontology*. 2008;106(1):930-938. doi:10.1016/j.tripleo.2008.03.018.

10. Ngan D. Comparison of radiation levels from computed tomography and conventional dental radiographs. *Aust Orthod J*. 2003;19(2):67-75.

11. Ludlow JB, Davies-Ludlow LE, Brooks SL, Howerton WB. Dosimetry of 3 CBCT devices for oral and maxillofacial radiology: CB Mercuray, NewTom 3G and i-CAT. *Dentomaxillofacial Radiol*. 2006;35(4):219-226. doi:10.1259/dmfr/14340323.

12. Palomo JM. Influence of CBCT exposure conditions on radiation dose. *Oral Surgery, Oral Med Oral Pathol Oral Radiol Endodontology*. 2008;105(6):773-782.

13. Guldner C. Potential of dosage reduction in cone-beam computed tomography (CBCT) for radiological diagnostics of the paranasal sinuses. *Eur Arch Otorhinolaryngol*. 2012;270(4):219=26.

14. Kim S, Choi Y, Hwang E, Chung K, Kook Y, Nelson G. Surgical positioning of orthodontic mini- implants with guides fabricated on models replicated with cone-beam

computed tomography. :82-89. doi:10.1016/j.ajodo.2006.01.027.

15. da Motta A. Superimposition of 3D cone-beam CT models in orthognathic surgery. *Dent Press J Orthod*. 2010;15(2):39-41. doi:10.1590/S2176-94512010000200005.Superimposition.

16. Katyal V, Pamula Y, Martin AJ, Daynes CN, Kennedy JD, Sampson WJ. Craniofacial and upper airway morphology in pediatric sleep-disordered breathing: Systematic review and meta-analysis. *Am J Orthod Dentofac Orthop*. 2013;143(1):20-30.e3. doi:10.1016/j.ajodo.2012.08.021.

17. Huynh NT, Desplats E, Almeida FR. Orthodontics treatments for managing obstructive sleep apnea syndrome in children: A systematic review and meta-analysis. *Sleep Med Rev*. 2016;25:84-94. doi:10.1016/j.smrv.2015.02.002.

18. Archbold KH, Pituch KJ, Panahi P, Chervin RD. Symptoms of sleep disturbances among children at two general pediatric clinics. *J Pediatr*. 2002;140(1):97-102. doi:10.1067/mpd.2002.119990.

19. Young T. Estimation of the clinically diagnosed proportion of sleep apnea syndrome in middle-aged men and women. *Sleep*. 1997;20(9):705-706.

20. McCrillis JM, Haskell J, Haskell BS, et al. Obstructive Sleep Apnea and the Use of Cone Beam Computed Tomography in Airway Imaging: A Review. *Semin Orthod*. 2009;15(1):63-69. doi:10.1053/j.sodo.2008.09.008.

21. Wiggins R V, Schmidt-nowara WW. Treatment of the Obstructive Sleep Apnea Syndrome. *West J Med*. 1987;(147):561-568.

22. Somers VK, Dyken ME, Clary MP, Abboud FM. Sympathetic neural mechanisms in obstructive sleep apnea. *J Clin Invest*. 1995;96(4):1897-1904. doi:10.1172/JCI118235.

23. Sauer C, Schlüter B, Hinz R, Gesch D. Childhood obstructive sleep apnea syndrome: an interdisciplinary approach: a prospective epidemiological study of 4,318 five-and-a-half-year-old children. *J Orofac Orthop*. 2012;73(5):342-358. doi:10.1007/s00056-012-0096-x.

24. Aasm. International Classification of Sleep Disorders. *Am Acad Sleep Med*. 2000.

25. Peppard P. Prospective Study of the Association Between Sleep-Disordered. *N Engl J Med*. 2000;342(19):1378-1384.

26. Young T, Peppard PE, Gottlieb DJ. Epidemiology of obstructive sleep apnea: A population health perspective. *Am J Respir Crit Care Med*. 2002;165(9):1217-1239. doi:10.1164/rccm.2109080.

27. Yaggi HK, Concato J, Kernan WN, Lichtman JH, Brass LM, Mohsenin V. Obstructive Sleep Apnea as a Risk Factor for Stroke and Death. *N Engl J Med*. 2005;353(19):2034-2041. doi:10.1056/NEJMoa043104.

28. Riley RW, Powell NB, Guilleminault C. Maxillary, mandibular, and hyoid advancement for treatment of obstructive sleep apnea: a review of 40 patients. *Yjoms*. 1990;48(1):20-26. http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=2294208&retm

ode=ref&cmd=prlinks%5Cnpapers2://publication/uuid/E9F3C58B-1CE1-4588-A540-B00E88E9513E.

29. Ioachimescu OC, Collop NA. Sleep-Disordered Breathing. *Neurol Clin*. 2012;30(4):1095-1136. doi:10.1016/j.ncl.2012.08.003.

30. Dempsey JA, Veasey SC, Morgan BJ OC. Pathophysiology of Sleep Apnea. *Physiol Rev*. 2010;90(1):47-112. doi:10.1152/physrev.00043.2008.

31. Valiathan M, El H, Hans MG, Palomo MJ. Effects of extraction versus non-extraction treatment on oropharyngeal airway volume. *Angle Orthod*. 2010;80(6):1068-1074. doi:10.2319/010810-19.1.

32. Auvenshine R. Anatomy of the Airway: An Overview. *Clin Sleep Med*. 2010;5(1):45-57.

33. Wang Q, Jia P, Anderson NK, Wang L, Lin J. Changes of pharyngeal airway size and hyoid bone position following orthodontic treatment of Class i bimaxillary protrusion. *Angle Orthod*. 2012;82(1):115-121. doi:10.2319/011011-13.1.

34. Pirilä-Parkkinena K, Löppönen H, Nieminen P, Tolonen U, Pääkkö E, Pirttiniemi P. Validity of upper airway assessment in children A clinical, cephalometric, and MRI study. *Angle Orthod*. 2011;81(3):433-439. doi:10.2319/063010-362.1.

35. Germec-Cakan D, Taner T, Akan S. Uvulo-glossopharyngeal dimensions in non-extraction, extraction with minimum anchorage, and extraction with maximum anchorage. *Eur J Orthod*. 2011;33(5):515-520. doi:10.1093/ejo/cjq109.

36. Tso HH, Lee JS, Huang JC, Maki K, Hatcher D, Miller AJ. Evaluation of the human airway using cone-beam computerized tomography. *Oral Surgery, Oral Med Oral Pathol Oral Radiol Endodontology*. 2009;108(5):768-776. doi:10.1016/j.tripleo.2009.05.026.

37. Schwab RJ, Pasirstein M, Pierson R, et al. Identification of upper airway anatomic risk factors for obstructive sleep apnea with volumetric magnetic resonance imaging. *Am J Respir Crit Care Med*. 2003;168(5):522-530. doi:10.1164/rccm.200208-866OC.

38. Schwab RJ, Goldberg AN. Upper Airway Assessment. *Otolaryngol Clin North Am*. 1998;31(6):931-968. doi:10.1016/S0030-6665(05)70100-6.

39. Johns MW. Daytime sleepiness, snoring, and obstructive sleep apnea; The Epworth Sleepiness Scale. *Chest*. 1993;103(1):30-36. doi:10.1378/chest.103.1.30.

40. George CFP. Sleep·5: Driving and automobile crashes in patients with obstructive sleep apnoea/hypopnoea syndrome. *Thorax*. 2004;59(9):804-807. doi:10.1136/thx.2003.007187.

41. Lee P-C, Hwang B, Soong W-J, Meng CCL. The specific characteristics in children with obstructive sleep apnea and cor pulmonale. *ScientificWorldJournal*. 2012;2012:757283. doi:10.1100/2012/757283.

42. Parish J. Obstructive sleep apnea and cardiovascular disease. *Mayo Clin Proc*. 2004;79(8):1036-1046.

43. Hagander L, Harlid R, Svanborg E. Quantitative sensory testing in the oropharynx: A means of showing nervous lesions in patients with obstructive sleep apnea and snoring.

Chest. 2009;136(2):481-489. doi:10.1378/chest.08-2747.

44. Gami AS, Howard DE, Olson EJ, Somers VK. Day-night pattern of sudden death in obstructive sleep apnea. *N Engl J Med*. 2005;352(12):1206-1214. doi:10.1056/NEJMoa041832.

45. Schafer K, Koehler U, Ewig S, Hasper E, Tasci S, Luderitz B. Obstructive sleep apnea as a risk marker in coronary artery disease. *Cardiology*. 1999;92(2):79-84.

46. Ambrosetti M. Is venous thromboembolism more frequent in patients with 1 obstructive sleep apnea syndrome ? *J Thromb Haemost*. 2004;2(10):1858-1860.

47. Sanner B. Pulmonary Hypertension in Patients with Obstructive Sleep Syndrome. *Arch Intern Med*. 1997;157(21):2483-2487.

48. Jean-Louis G, Zizi F, Clark LT, Brown CD, McFarlane SI. Obstructive sleep apnea and cardiovascular disease: role of the metabolic syndrome and its components. *J Clin Sleep Med*. 2008;4(3):261-272.

49. Zettergren-Wijk L, Forsberg C-M, Linder-Aronson S. Changes in dentofacial morphology after adeno-/tonsillectomy in young children with obstructive sleep apnoea - A 5-year follow-up study. *Eur J Orthod*. 2006;28(4):319-326. doi:10.1093/ejo/cji119.

50. Redline S, Tishler P V., Schluchter M, Aylor J, Clark K, Graham G. Risk factors for sleep-disordered breathing in children: Associations with obesity, race, and respiratory problems. *Am J Respir Crit Care Med*. 1999;159(5 I):1527-1532. doi:10.1164/ajrccm.159.5.9809079.

51. Redline S. The genetics of sleep disorders. *Sleep Med Rev*. 2000;4(6):583-602. doi:10.1053/smrv.2000.0120.

52. Özbek MM, Toygar Memikoglu TU, Gögen H, Lowe AA, Baspinar E. Oropharyngeal airway dimensions and functional-orthopedic treatment in skeletal Class II cases. *Angle Orthod*. 1998;68(4):327-336. doi:10.1043/0003-3219(1998)068<0327:OADAFO>2.3.CO;2.

53. Lowe AA, Santamaria JD, Fleetham JA, Price C. Facial morphology and obstructive sleep apnea. *Am J Orthod Dentofacial Orthop*. 1986;90(6):484-491. doi:10.1016/0889-5406(86)90108-3.

54. Bender B. [Upper Airway Stimulation in OSA]. *Laryngorhinootologie*. 2016;95(11):795-807. doi:10.1055/s-0042-116895.

55. Isono S, Feroah TR, Hajduk EA, Brant R, Whitelaw WA, Remmers JE. Interaction of cross-sectional area, driving pressure, and airflow of passive velopharynx. *J Appl Physiol*. 1997;83(3):851-859. http://www.ncbi.nlm.nih.gov/pubmed/9292473.

56. Muñoz R, Durán-Cantolla J, Martinez-Vila E, et al. Central sleep apnea is associated with increased risk of ischemic stroke in the elderly. *Acta Neurol Scand*. 2012;126(3):183-188. doi:10.1111/j.1600-0404.2011.01625.x.

57. Liu Y, Zeng X, Fu M, Huang X, Lowe AA. Effects of a mandibular repositioner on

obstructive sleep apnea. *Am J Orthod Dentofac Orthop*. 2000;118(3):248-256. doi:10.1067/mod.2000.104831.

58. Nelson S, Hans M. Contribution of craniofacial risk factors in increasing apneic activity among obese and nonobese habitual snorers. *Chest*. 1997;111(1):154-162. doi:10.1378/chest.111.1.154.

59. Pahkala R, Puustinen R, Tuomilehto H, Ahlberg J, Seppä J. Risk factors for sleep-disordered breathing: the role of craniofacial structure. *Acta Odontol Scand*. 2011;69(3):137-143. doi:10.3109/00016357.2010.545033.

60. Hong JS, Oh KM, Kim BR, Kim YJ, Park YH. Three-dimensional analysis of pharyngeal airway volume in adults with anterior position of the mandible. *Am J Orthod Dentofac Orthop*. 2011;140(4):e161-e169. doi:10.1016/j.ajodo.2011.04.020.

61. El H, Palomo JM. Airway volume for different dentofacial skeletal patterns. *Am J Orthod Dentofac Orthop*. 2011;139(6):e511-e521. doi:10.1016/j.ajodo.2011.02.015.

62. Ferguson KA, Ono T, Lowe AA, Ryan CF, Fleetham JA. The relationship between obesity and craniofacial structure in obstructive sleep apnea. *Chest*. 1995;108(2):375-381. doi:10.1378/chest.108.2.375.

63. Yamaguchi K, Nanda RS. The effects of extraction and nonextraction treatment on the mandibular position. *Am J Orthod Dentofac Orthop*. 1991;100(5):443-452. doi:10.1016/0889-5406(91)70084-A.

64. Finkelstein Y, Wexler D, Horowitz E, et al. Frontal and lateral cephalometry in patients with sleep-disordered breathing. *Laryngoscope*. 2001;111(4 Pt 1):634-641. doi:10.1097/00005537-200104000-00014.

65. Bates CJ, McDonald JP. The relationship between severity of obstructive sleep apnoea/hypopnoea syndrome (OSAHS) and lateral cephalometric radiograph values: a clinical diagnostic tool. *Surg J R Coll Surg Edinburgh Irel*. 2005;3(5):338-346. doi:10.1016/S1479-666X(05)80113-1.

66. Lowe AA, Gionhaku N, Takeuchi K, Fleetham JA. Three-dimensional CT reconstructions of tongue and airway in adult subjects with obstructive sleep apnea. *Am J Orthod Dentofac Orthop*. 1986;90(5):364-374. doi:10.1016/0889-5406(86)90002-8.

67. Remmers JE, deGroot WJ, Sauerland EK, Anch a M. Pathogenesis of upper airway occlusion during sleep. *J Appl Physiol*. 1978;44(6):931-938.

68. Bradley T. Physiological determinants of nocturnal arterial oxygenation in patients with obstructive sleep apnea. *J Appl Physiol*. 1985;59(5):1364-1368.

69. Hancox RJ, Landhuis CE. Association between sleep duration and haemoglobin A1c in young adults. *J Epidemiol Community Health*. 2011;66(10):957-961. doi:10.1136/jech-2011-200217.

70. Choi JB, Loredo JS, Norman D, et al. Does obstructive sleep apnea increase hematocrit? *Sleep Breath*. 2006;10(3):155-160. doi:10.1007/s11325-006-0064-z.

71. Buchanan A, Cohen R, Looney S, Kalathingal S, De Rossi S. Cone-beam CT analysis of patients with obstructive sleep apnea compared to normal controls. *Imaging Sci Dent*. 2016;46(1):9-16. doi:10.5624/isd.2016.46.1.9.

72. Susarla SM, Thomas RJ, Abramson ZR, Kaban LB. Biomechanics of the upper airway: Changing concepts in the pathogenesis of obstructive sleep apnea. *Int J Oral Maxillofac Surg*. 2010;39(12):1149-1159. doi:10.1016/j.ijom.2010.09.007.

73. Tsuiki S, Isono S, Ishikawa T, Yamashiro Y, Tatsumi K, Nishino T. Anatomical balance of the upper airway and obstructive sleep apnea. *Anesthesiology*. 2008;108(6):1009-1015. doi:10.1097/ALN.0b013e318173f103.

74. Lenza MG, Lenza MMDO, Dalstra M, Melsen B, Cattaneo PM. An analysis of different approaches to the assessment of upper airway morphology: a CBCT study. *Orthod Craniofac Res*. 2010;13(2):96-105. doi:10.1111/j.1601-6343.2010.01482.x.

75. Zimmerman JN, Lee J, Pliska BT. Reliability of upper pharyngeal airway assessment using dental CBCT: a systematic review. *Eur J Orthod*. 2016:cjw079. doi:10.1093/ejo/cjw079.

76. Liberati A, Altman DG, Tetzlaff J, et al. *The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Health Care Interventions: Explanation and Elaboration*. Vol 62.; 2009. doi:10.1016/j.jclinepi.2009.06.006.

77. Beller E. PLOS Medicine: PRISMA for Abstracts: Reporting Systematic Reviews in Journal and Conference Abstracts. http://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1001419.

78. Alsufyani NA, Flores-Mir C, Major PW. Three-dimensional segmentation of the upper airway using cone beam CT: A systematic review. *Dentomaxillofacial Radiol*. 2012;41(4):276-284. doi:10.1259/dmfr/79433138.

79. Alves M, Franzotti ES, Baratieri C, Nunes LKF, Nojima LI, Ruellas ACO. Evaluation of pharyngeal airway space amongst different skeletal patterns. *Int J Oral Maxillofac Surg*. 2012;41(7):814-819. doi:10.1016/j.ijom.2012.01.015.

80. Alves M, Baratieri C, Nojima LI, Nojima MCG, Ruellas ACO. Three-dimensional assessment of pharyngeal airway in nasal- and mouth-breathing children. *Int J Pediatr Otorhinolaryngol*. 2011;75(9):1195-1199. doi:10.1016/j.ijporl.2011.06.019.

81. Bandeira AM, Oltramari-Navarro PVP, de Lima Navarro R, de Castro Ferreira Conti AC, de Almeida MR, Fernandes KBP. Three-dimensional upper-airway assessment in patients with bronchial asthma. *Angle Orthod*. 2014;84(2):254-259. doi:10.2319/030113-176.

82. Brunetto DP, Velasco L, Koerich L, Araújo MTDS. Prediction of 3-dimensional pharyngeal airway changes after orthognathic surgery: A preliminary study. *Am J Orthod Dentofac Orthop*. 2014;146(3):299-309. doi:10.1016/j.ajodo.2014.05.024.

83. Burkhard JPM, Dietrich AD, Jacobsen C, Roos M, L??bbers HT, Obwegeser JA. Cephalometric and three-dimensional assessment of the posterior airway space and imaging software reliability analysis before and after orthognathic surgery. *J Cranio-Maxillofacial Surg*. 2014;42(7):1428-1436. doi:10.1016/j.jcms.2014.04.005.

84. Celikoglu M, Ucar FI, Sekerci AE, Buyuk SK, Ersoz M, Sisman Y. Assessment of pharyngeal airway volume in adolescent patients affected by bilateral cleft lip and palate using cone beam computed tomography. *Angle Orthod*. 2014;84(6):995-1001. doi:10.2319/121913-930.1.

85. Chang Y, Koenig LJ, Pruszynski JE, Bradley TG, Bosio JA, Liu D. Dimensional changes of upper airway after rapid maxillary expansion: A prospective cone-beam computed tomography study. *Am J Orthod Dentofac Orthop*. 2013;143(4):462-470. doi:10.1016/j.ajodo.2012.11.019.

86. Cheung T, Oberoi S. Three Dimensional Assessment of the Pharyngeal Airway in Individuals with Non-Syndromic Cleft Lip and Palate. *PLoS One*. 2012;7(8):1-5. doi:10.1371/journal.pone.0043405.

87. De Souza K.R.S., Oltramari-Navarro P.V.P., Navarro R.L. CACCF, M.R. and de A. Reliability of a method to conduct upper airway analysis in cone-beam computed tomography. *Braz Oral Res*. 2013;27(1):48–54.

88. Di Carlo G, Polimeni A, Melsen B, Cattaneo PM. The relationship between upper airways and craniofacial morphology studied in 3D. A CBCT study. *Orthod Craniofacial Res*. 2015;18(1):1-11. doi:10.1111/ocr.12053.

89. El H, Palomo JM. Measuring the airway in 3 dimensions: A reliability and accuracy study. *Am J Orthod Dentofac Orthop*. 2010;137(4 SUPPL.):S50.e1-S50.e9. doi:10.1016/j.ajodo.2009.11.010.

90. Enciso R, Nguyen M, Shigeta Y, Ogawa T, Clark GT. Comparison of cone-beam CT parameters and sleep questionnaires in sleep apnea patients and control subjects. *Oral Surgery, Oral Med Oral Pathol Oral Radiol Endodontology*. 2010;109(2):285-293. doi:10.1016/j.tripleo.2009.09.033.

91. Feng X, Li G, Qu Z, Liu L, Näsström K, Shi XQ. Comparative analysis of upper airway volume with lateral cephalograms and cone-beam computed tomography. *Am J Orthod Dentofac Orthop*. 2015;147(2):197-204. doi:10.1016/j.ajodo.2014.10.025.

92. Glupker L, Kula K, Parks E, Babler W, Stewart K, Ghoneima A. Three-dimensional computed tomography analysis of airway volume changes between open and closed jaw positions. *Am J Orthod Dentofac Orthop*. 2015;147(4):426-434. doi:10.1016/j.ajodo.2014.11.025.

93. Grauer D, Cevidanes LSH, Styner MA, Ackerman JL, Proffit WR. Pharyngeal airway volume and shape from cone-beam computed tomography: Relationship to facial morphology. *Am J Orthod Dentofac Orthop*. 2009;136(6):805-814. doi:10.1016/j.ajodo.2008.01.020.

94. Guijarro-Martínez R, Swennen GRJ. Three-dimensional cone beam computed tomography definition of the anatomical subregions of the upper airway: A validation study. *Int J Oral Maxillofac Surg*. 2013;42(9):1140-1149. doi:10.1016/j.ijom.2013.03.007.

95. Hart PS, McIntyre BP, Kadioglu O, et al. Postsurgical volumetric airway changes in 2-jaw orthognathic surgery patients. *Am J Orthod Dentofac Orthop*. 2015;147(5):536-546.

doi:10.1016/j.ajodo.2014.12.023.

96. Hong JS, Park YH, Kim YJ, Hong SM, Oh KM. Three-dimensional changes in pharyngeal airway in skeletal class III patients undergoing orthognathic surgery. *J Oral Maxillofac Surg*. 2011;69(11):401-408. doi:10.1016/j.joms.2011.02.011.

97. Iannetti G, Polimeni A, Pagnoni M, et al. Upper airway volume after Le Fort III advancement in subjects with craniofacial malformation. *JCraniofacSurg*. 2011;22(1536-3732 (Electronic)):351-355. doi:10.1097/SCS.0b013e3181f7e11b.

98. Iwasaki T, Saitoh I, Takemoto Y, et al. Tongue posture improvement and pharyngeal airway enlargement as secondary effects of rapid maxillary expansion: A cone-beam computed tomography study. *Am J Orthod Dentofac Orthop*. 2013;143(2):235-245. doi:10.1016/j.ajodo.2012.09.014.

99. Jiang Y-Y, Xu X, Su H-L, Liu D-X. Gender-related difference in the upper airway dimensions and hyoid bone position in Chinese Han children and adolescents aged 6-18 years using cone beam computed tomography. *Acta Odontol Scand*. 2015;73(5):391-400. doi:10.3109/00016357.2014.978366.

100. Kim MA, Kim BR, Choi JY, Youn JK, Kim YJR, Park YH. Three-dimensional changes of the hyoid bone and airway volumes related to its relationship with horizontal anatomic planes after bimaxillary surgery in skeletal Class III patients. *Angle Orthod*. 2013;83(4):623-629. doi:10.2319/083112-700.1.

101. Kim YJ, Hong JS, Hwang YI, Park YH. Three-dimensional analysis of pharyngeal airway in preadolescent children with different anteroposterior skeletal patterns. *Am J Orthod Dentofac Orthop*. 2010;137(3):306.e1-306.e11. doi:10.1016/j.ajodo.2009.10.025.

102. Kochel, J; Meyer-Marcotty, P; Sickel, F; Lindorf, H; Stellzig-Eisenhauer A. Short-term pharyngeal airway changes after mandibular advancement surgery in adult Class II-Patients-a three-dimensional retrospective study. *J Orofac Orthop Der Kieferorthopadie*. 2013;74(2):137-152.

103. Li L, Liu H, Cheng H, et al. CBCT Evaluation of the upper airway morphological changes in growing patients of class ii division 1 malocclusion with mandibular retrusion using twin block appliance: A comparative research. *PLoS One*. 2014;9(4):1-7. doi:10.1371/journal.pone.0094378.

104. Li Y-M, Liu J-L, Zhao J-L, et al. Morphological changes in the pharyngeal airway of female skeletal class III patients following bimaxillary surgery: A cone beam computed tomography evaluation. *Int J Oral Maxillofac Surg*. 2014;43(7):862-867. doi:10.1016/j.ijom.2014.03.009.

105. Mattos CT, Cruz CV, Da Matta TCS, et al. Reliability of upper airway linear, area, and volumetric measurements in cone-beam computed tomography. *Am J Orthod Dentofac Orthop*. 2014;145(2):188-197. doi:10.1016/j.ajodo.2013.10.013.

106. Oh KM, Kim MA, Youn JK, Cho HJ, Park YH. Three-dimensional evaluation of the relationship between nasopharyngeal airway shape and adenoid size in children. *Korean J Orthod*. 2013;43(4):160-167. doi:10.4041/kjod.2013.43.4.160.

107. Sears CR, Miller AJ, Chang MK, Huang JC, Lee JS. Comparison of pharyngeal airway changes on plain radiography and cone-beam computed tomography after orthognathic surgery. *J Oral Maxillofac Surg*. 2011;69(11):e385-e394. doi:10.1016/j.joms.2011.03.015.

108. Starbuck JM, Friel MT, Ghoneima A, Flores RL, Tholpady S, Kula K. Nasal airway and septal variation in unilateral and bilateral cleft lip and palate. *Clin Anat*. 2014;27(7):999-1008. doi:10.1002/ca.22428.

109. Stefanovic N, El H, Chenin DL, Glisic B, Palomo JM. Three-dimensional pharyngeal airway changes in orthodontic patients treated with and without extractions. *Orthod Craniofacial Res*. 2013;16(2):87-96. doi:10.1111/ocr.12009.

110. Valladares-Neto J, Silva MAG, Bumann A, Paiva JB, Rino-Neto J. Effects of mandibular advancement surgery combined with minimal maxillary displacement on the volume and most restricted cross-sectional area of the pharyngeal airway. *Int J Oral Maxillofac Surg*. 2013;42(11):1437-1445. doi:10.1016/j.ijom.2013.03.018.

111. Vizzotto MB, Liedke GS, Delamare EL, Silveira HD, Dutra V, Silveira HE. A comparative study of lateral cephalograms and cone-beam computed tomographic images in upper airway assessment. *Eur J Orthod*. 2012;34(3):390-393. doi:10.1093/ejo/cjr012.

112. Weissheimer A, Menezes LM De, Sameshima GT, Enciso R, Pham J, Grauer D. Imaging software accuracy for 3-dimensional analysis of the upper airway. *Am J Orthod Dentofac Orthop*. 2012;142(6):801-813. doi:10.1016/j.ajodo.2012.07.015.

113. Xu Y, Zhao S, Shi J, et al. 3-Dimensional Computed Tomographic Analysis of the Pharynx in Adult Patients With Unrepaired Isolated Cleft Palate. *J Oral Maxillofac Surg*. 2013;71(8):1424-1434. doi:10.1016/j.joms.2013.01.022.

114. Yoshihara M, Terajima M, Yanagita N, et al. Three-dimensional analysis of the pharyngeal airway morphology in growing Japanese girls with and without cleft lip and palate. *Am J Orthod Dentofac Orthop*. 2012;141(4 SUPPL):S92-S101. doi:10.1016/j.ajodo.2011.09.011.

115. Zhao Y, Nguyen M, Gohl E, Mah JK, Sameshima G, Enciso R. Oropharyngeal airway changes after rapid palatal expansion evaluated with cone-beam computed tomography. *Am J Orthod Dentofac Orthop*. 2010;137(4 SUPPL.):S71-S78. doi:10.1016/j.ajodo.2008.08.026.

116. Zheng ZH, Yamaguchi T, Kurihara A, Li HF, Maki K. Three-dimensional evaluation of upper airway in patients with different anteroposterior skeletal patterns. *Orthod Craniofacial Res*. 2014;17(1):38-48. doi:10.1111/ocr.12029.

117. Aboudara C, Nielsen I, Huang JC, Maki K, Miller AJ, Hatcher D. Comparison of airway space with conventional lateral headfilms and 3-dimensional reconstruction from cone-beam computed tomography. *Am J Orthod Dentofac Orthop*. 2009;135(4):468-479. doi:10.1016/j.ajodo.2007.04.043.

118. Haskell JA, McCrillis J, Haskell BS, Scheetz JP, Scarfe WC, Farman AG. Effects of Mandibular Advancement Device (MAD) on Airway Dimensions Assessed With Cone-Beam Computed Tomography. *Semin Orthod*. 2009;15(2):132-158.

doi:10.1053/j.sodo.2009.02.001.

119. Iwasaki T, Hayasaki H, Takemoto Y, Kanomi R, Yamasaki Y. Oropharyngeal airway in children with Class III malocclusion evaluated by cone-beam computed tomography. *Am J Orthod Dentofac Orthop*. 2009;136(3):318.e1-318.e9. doi:10.1016/j.ajodo.2009.02.017.

120. Ghoneima A, Imburgia A, Halum S, Van Dis M, Kula K. Three-dimensional airway analysis of trumpet players vs. non-trumpet players. *Oral Radiol*. 2015;31(2):105-113. doi:10.1007/s11282-014-0187-z.

121. Alves M, Baratieri C, Mattos CT, et al. Is the airway volume being correctly analyzed? *Am J Orthod Dentofac Orthop*. 2012;141(5):657-661. doi:10.1016/j.ajodo.2011.11.019.

122. Major MP, Witmans M, El-Hakim H, Major PW, Flores-Mir C. Agreement between cone-beam computed tomography and nasoendoscopy evaluations of adenoid hypertrophy. *Am J Orthod Dentofac Orthop*. 2014;146(4):451-459. doi:10.1016/j.ajodo.2014.06.013.

123. Commission E. Radiation Protection N°172. Cone Beam CT for Dental and Maxillofacial Radiology: Evidence Based Guideline. http://ec.europa.eu/energy/en/content/172.

124. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. *Stat Med*. 1998;17(1):101-110. doi:10.1002/(SICI)1097-0258(19980115)17:1<101::AID-SIM727>3.0.CO;2-E.

125. Aneja D, Vora S, Camci E, Shapiro L, Cox T. Automated Detection of 3D Landmarks for the Elimination of Non-Biological Variation in Geometric Morphometric Analyses. *Proc IEEE Int Symp Comput Based Med Syst*. 2015;18(0):78-83. doi:10.1016/j.drudis.2013.07.001.Phenotypic.

126. Bromiley PA, Schunke AC, Ragheb H, Thacker NA, Tautz D. Semi-automatic landmark point annotation for geometric morphometrics. *Front Zool*. 2014;11(1):61. doi:10.1186/s12983-014-0061-1.

127. Chen H, Aarab G, de Ruiter MHT, de Lange J, Lobbezoo F, van der Stelt PF. Three-dimensional imaging of the upper airway anatomy in obstructive sleep apnea: A systematic review. *Sleep Med*. 2016;21:19-27. doi:10.1016/j.sleep.2016.01.022.

128. Arens R, Marcus C. Pathophysiology of upper airway obstruction: a developmental perspective. *Sleep*. 2004;27(5):997-1019.

129. Ono T, Otsuka R, Kuroda T, Honda E ST. Effects of Head and Body Position on Two- and Three-dimensional Configurations of the Upper Airway. *J Dent Res*. 2000;79(11):1879-1884.

130. Pirilä-parkkinen K, Pirttiniemi P. Pharyngeal airway in children with sleep-disordered breathing in relation to head posture. *Sleep Breath*. 2012;16(3):737-746. doi:10.1007/s11325-011-0569-y.

131. Zhang W, Song X, Masumi I. Effects of head and body positions on 2- and 3-dimensional configuration of the oropharynx with jaw protruded : a magnetic resonance imaging study. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod*. 2011;111(6):778-784. doi:10.1016/j.tripleo.2011.02.019.

132. Gurani SF, Di Carlo G, Cattaneo PM, Thorn JJ, Pinholt EM. Effect of Head and Tongue Posture on the Pharyngeal Airway Dimensions and Morphology in Three-Dimensional Imaging: a Systematic Review. *J oral Maxillofac Res*. 2016;7(1):1-12. doi:10.5037/jomr.2016.7101.

133. Camacho M, Capasso R, Schendel S. Airway changes in obstructive sleep apnoea patients associated with a supine versus an upright position examined using cone beam computed tomography. *J Laryngol Otol*. 2014;128(August):824-830. doi:10.1017/S0022215114001686.

134. Barrera JE. Sleep Magnetic Resonance Imaging : Dynamic Characteristics of the Airway During Sleep in Obstructive Sleep Apnea Syndrome. *Laryngoscope*. 2011;121(June):1327-1335. doi:10.1002/lary.21810.

135. Cheng S, Brown EC, Hatt A, Butler JE, Gandevia SC, Bilston LE. Healthy humans with a narrow upper airway maintain patency during quiet breathing by dilating the airway during inspiration. *J Physiol*. 2014;592(21):4763-4774. doi:10.1113/jphysiol.2014.279240.

Appendix A

Table A.1 Characters of the included studies in the systematic review (N=42)

Study	Randomized	Sample Size Age Syndromic	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Alves et al.[79]	Reliability sample yes Original sample no	12 8-10 years old Non-syndromic	NP	NP	Upper airway	Total Airway Volume Minimum cross-sectional area	ICC >0.98 0.91	Dolphin Imaging® software, version 11.0	NP

Study	Randomized	Sample Size Age Syndromic	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Alves et al.[80]	Reliability sample yes Original sample no	16 8-10 years old Non-syndromic	NP	NP	Upper airway	Total Airway Volume Minimum cross-sectional area	ICC >0.98 0.91	Dolphin Imaging® software, version 11.0	NP

Study	Randomized	Sample Size Age Syndromic	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Bandiera et al.[81]	Only Asthmatic group	52 Asthma group mean 14.85 years old Asthmatic	No asthma Control mean 16.65 years old	NP	Upper airway	Total Airway Volume Nasopharyngeal minimum cross-sectional area Oropharyngeal minimum cross-sectional area	ICC 0.99 0.99 0.98	Dolphin Imaging® software, version 11.5	NP

Study	Randomized	Sample Size Age Syndromic	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Brunetto et al.[82]	No	20 18-30 years old Non-syndromic	NP	NP	Upper airway	Upper segment volume Lower segment volume Total upper airway volume Minimum cross-sectional area	ICC 0.941 0.934 0.948 0.902	Dolphin Imaging® software, version 11.5	NP

Study	Randomized	Sample Size Age Syndromic	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Burkhard et al.[83]	Yes	11 19-44 years old (mean 26 years old) Non-syndromic	NP	Automatic	Upper airway	Airway Diameter Total Airway Volume	Dahlberg formula Inter-observer 98.9% Inter-program 94.2% Inter-observer 99.2% Inter-program 96.1%	OsiriX®, Mimics® and BrainLab®	NP

Study	Randomized	Sample Size Age Syndromic	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Celikoglu et al.[84]	Reliability sample yes	10 Surgical group mean age 14.1 years old Bilateral cleft lip and palate	Non – syndromic Mean age 13.4 years old	NP	Upper airway	Total Airway Volume Nasopharyngeal airway volume Oropharyngeal airway volume	ICC >0.977 >0.977 >0.977	Mimics 15.01	NP

Study	Randomized	Sample Size Age Syndromic	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Chang et al.[85]	No	14 9-16 years old (mean 12.9 years old) Non-syndromic	NP	NP	Upper airway	Cross-sectional area	ICC 0.853	Dolphin Imaging® software, version 11.0	60 units

Study	Randomized	Sample Size Age Syndromic	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Cheung and Oberoi[86]	Reliability sample yes Original sample no	7 Mean age 10.6 years old Unilateral and bilateral cleft lip and palate	Matched non-syndromic	NP	Upper airway	Total Airway Volume Minimum cross-sectional area	Pearson correlation coefficient and Lin concordance 0.99 and 0.99 0.99 and 0.99	CB Works 3.0	NP

Study	Randomized	Sample Size Age Syndromic	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
De Souza et al.[87]	No	60 Mean 17.86 years old Non-syndromic	NP	NP	Upper airway	Total airway volume Nasopharyngeal minimum cross-sectional area	ICC Intra researcher 1: 0.99 Intra researcher 2: 0.99 Inter 0.95 Intra researcher 1: 0.98 Intra researcher 2: 0.93 Inter 0.88	Dolphin Imaging® software, version 11.5	NP

Study	Randomized	Sample Size Age	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
De Souza et al.[87] Continued						Oropharyngeal minimum cross-sectional area	Intra researcher 1: 0.99 Intra researcher 2: 0.98 Inter 0.98		
Di Carlo et al.[88]	Reliability sample yes Original sample no	7 13-43 years old Non-syndromic	NP	NP	Upper airway	Total volume Lower nasopharynx volume Velopharynx volume Oropharynx volume	ICC 0.9 0.9 0.7 0.9	Mimics 15.0	NP

Study	Randomized	Sample Size Age Syndromic	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
El and Palomo[89]	Yes	30	NP	OrthoSegment uses manual segmentation Other 3 programs use automatic segmentation	Upper airway	Oropharynx volume Nasopharynx volume	ICC Orthosegment: 0.99 Dolphin 3D: 0.99 InVivoDental: 0.99 OnDemand3D: 0.99 Orthosegment: 0.98 Dolphin 3D: 0.88 InVivoDental: 0.97	Dolphin Imaging® software, version 11.0 InVivoDental version 4.0.70 OnDemand3D version 1.0.1.8407 OrthoSegment	NP

Study	Randomized	Sample Size Age Syndromic	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
El and Palomo[89] Continued							OnDemand3D: 0.89		
Enciso et al.[90]	Reliability sample yes Original sample no	20 Mean age 57.5 years old OSA and snorers AHI>10	AHI<10 Mean age 50.8 years old	NP	Upper airway	Total volume Minimum cross-sectional area	ICC 0.965 0.979	vWorks 5.0	NP

Study	Randomized	Sample Size Age Syndromic	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Feng et al.[91]	Reliability sample yes Original sample no	10 9-43 years old Non-syndromic	NP	NP	Upper airway	Nasopharyngeal volume	ICC Intra researcher 1: 0.96 Intra researcher 2: 0.99 Inter measurement 1: 0.96 Inter measurement 2: 0.97	Dolphin Imaging® software, version 11.0	25, 30, 40, and 50

Study	Randomized	Sample Size Age Syndromic	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Glupker et al.[92]	Reliability sample yes Original sample no	10 Mean age 40.3 years old Non-syndromic	NP	NP	Upper airway	Nasopharyngeal volume Oropharyngeal volume Minimum constricted area	ICC >0.8 >0.8 >0.8	Dolphin Imaging® software, version 11.5	NP

Study	Randomized	Sample Size Age Syndromic	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Grauer et al.[93]	Reliability sample yes Original sample no	5 17-46 years old Non-syndromic	NP	Semiautomatic segmentation	Upper airway	Total Airway volume Superior component Inferior component	Mean coefficient of variation 1.90% NP NP	InsightSNAP software, version 1.4.0	NP

Study	Randomized	Sample Size Age Syndromic	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Guijarro-Martinez and Swennen[94]	No	35 23-35 years old Non-syndromic	NP	NP	Upper airway	Nasopharynx minimum cross-sectional area Nasopharyngeal volume	ICC Intra researcher 1: 0.848 Intra researcher 2: 0.937 Inter: 0.876 Intra researcher 1: 0.981 Intra researcher 2: 0.992 Inter: 0.986	Dolphin Imaging® software, version 11.0	Preliminary assessment of all scans using manual thresholding (range 48-81) Average threshold of preliminary scans was 70

Study	Randomized	Sample Size Age Syndromic	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Guijarro-Martinez and Swennen[94] Continued						Oropharyngeal minimum cross-sectional area Oropharyngeal volume	Intra researcher 1: 0.780 Intra researcher 2: 0.825 Inter: 0.837 Intra researcher 1: 0.997 Intra researcher 2: 0.999 Inter: 0.998		and this was the threshold value that was then used in the study

Study	Randomized	Sample Size Age Syndromic	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Guijarro-Martinez and Swennen[94] Continued						Hypopharyngeal minimum cross-sectional area	Intra researcher 1: 0.904 Intra researcher 2: 0.936 Inter: 0.839		
						Hypopharyngeal volume	Intra researcher 1: 0.994 Intra researcher 2: 0.996 Inter: 0.994		

Study	Randomized	Sample Size Age Syndromic	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Hart et al.[95]	No	71 Mean age 18.8 years old Non-syndromic	NP	NP	Upper airway	Total airway volume Nasopharyngeal volume Oropharyngeal volume Minimum cross-sectional area	ICC All values ranged from 0.77-0.99	Invivo5	-1000 and -604.3 Hounsfield units

Study	Randomized	Sample Size Age Syndromic	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Hong et al.[96]	Reliability sample yes Original sample no Reliability sample yes	10 18-30 years old (mean age 20.6 years old) Non-syndromic	NP	NP	Upper airway	Total airway volume Minimum cross-sectional area	Mean coefficient of variation 1.94% for all measurements	InVivoDental	-1024 to -300 Hounsfield units

Study	Randomized	Sample Size Age Syndromic	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Iannetti et al.[97]	No	4 5-9 years old Aperts or Crouzon syndromes	NP	NP	Upper airway	Total Airway volume	Intra: Wilcoxon signed rank test Z = -0.770, P = 0.441 Mean difference 11.8 mm^3 Inter: Mann-Whitney test 4.34 Mean difference 12.7 mm^3	Dolphin Imaging® software, version 11.0	NP

Study	Randomized	Sample Size Age Syndromic	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Iwasaki et al.[98]	Reliability sample yes Original sample no	10 Mean age 9.96 years old Non-syndromic requiring RME	Non-syndromic not requiring RME Mean age 9.68 years old Age, sex, and dentition matched	Threshold segmentation	Upper airway	Intraoral airway volume Retropalatal airway volume Oropharyngeal airway volume Total airway volume	ICC All measurements ranged from 0.965-0.998	INTAGE Volume Editor	NP

Study	Randomized	Sample Size Age Syndromic	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Jiang et al.[99]	Reliability sample yes Original sample no	20 6-18 years old Non-syndromic	NP	NP	Upper airway	Total Airway volume Minimum cross-sectional area	ICC >0.98 for all measurements	Mimics 16.01	NP

Study	Randomized	Sample Size Age Syndromic	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Kim et al.[100]	Reliability sample yes Original sample no	10 17-48 years old (mean age 30.04 years old) Non-syndromic	NP	NP	Upper airway	Total airway volume Nasopharyngeal airway volume Oropharyngeal airway volume Hypopharyngeal airway volume	Dahlberg formula Varied from 1054.47 to 1418.88 mm^3 for the volumetric measurements	InVivoDental	-1024 to -300 Hounsfield units

Study	Randomized	Sample Size Age Syndromic	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Kim et al.[101]	Reliability sample yes Original sample no	15 Mean age 11.19 years old Non-syndromic	NP	NP	Upper airway	Superior pharyngeal airway volume Middle pharyngeal airway volume Inferior pharyngeal airway volume Total airway volume Minimum cross-sectional area	Dahlberg formula Varied from 57.36 to 91.37 mm^3 for the volumetric measurements Varied from 11.33 to 36.12 mm^2	InVivoDental	NP

Study	Randomized	Sample Size Age Syndromic	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Kochel et al.[102]	Reliability sample yes Original sample no	20 Mean age 31.8 years old Non-syndromic	NP	NP	Upper airway	Total airway volume Upper pharyngeal airway volume Middle pharyngeal airway volume Lower pharyngeal airway volume Upper minimum CSA	Dahlberg formula 78.0 mm^3 90.3 mm^3 125.1 mm^3 66.4 mm^3 10.1 mm^2	Mimics® Innovation Suite 14.1	NP

Study	Randomized	Sample Size Age Syndromic	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Kochel et al.[102] Continued						Middle pharyngeal minimum cross-sectional area Lower pharyngeal minimum cross-sectional area Smallest pharyngeal cross-sectional area	3.5 mm^2 2.8 mm^2 5.2 mm^2		

Study	Randomized	Sample Size Age Syndromic	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Lenza et al.[74]	Reliability sample yes Original sample no	5 Mean age 18 years old Non-syndromic	NP	NP	Upper airway	Lower nasopharyngeal airway volume Upper velopharyngeal airway volume Lower velopharyngeal airway volume Upper OAV	Dahlberg formula 145.42 mm^3 249.68 mm^3 168.32 mm^3 283.86 mm^3	Mimics® Innovation Suite 12.13	NP

Study	Randomized	Sample Size Age Syndromic	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Lenza et al.[74] Continued						Lower Oropharyngeal airway volume Total airway volume Lower nasopharyngeal minimum cross-sectional area Upper velopharygeal minimum cross-sectional area Lower VCSA	$364.43\ mm^3$ $475.58\ mm^3$ $19.08\ mm^2$ $31.95\ mm^2$ $10.40\ mm^2$		

Study	Randomized	Sample Size Age Syndromic	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Lenza et al.[74] Continued						Upper Oropharyngeal minimum cross-sectional area Lower oropharyngeal minimum cross sectional area Smallest cross-sectional area	19.07 mm^2 14.34 mm^2 22.93 mm^2		

Study	Randomized	Sample Size Age Syndromic	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Li, L. et al.[103]	No	60 Mean age 11.57 years old Non-syndromic Retrusive mandible	Normal mandibular length Mean age 11.72 years old Matched for age, sex, and development condition	NP	Upper airway	Minimum cross-sectional area	Method error Varied from 5.76-7.85 mm^2	Mimics® Innovation Suite 16.0	NP

Study	Randomized	Sample Size Age Syndromic	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Li, YM. et al.[104]	No	29 18-35 years old (mean age 23.6 years old) Class III skeletal non-syndromic	NP	NP	Upper airway	Nasopharyngeal airway volume Oropharyngeal airway volume Total airway volume	r Interobserver >0.9 for all measurements	Mimics® Innovation Suite 10.01	NP

Study	Randomized	Sample Size Age Syndromic	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Mattos et al.[105]	Yes	12 NP Non-syndromic	NP	NP	Upper airway	Palatal plane minimum cross-sectional area Soft palate level minimum cross-sectional area	ICC Undergrad 0.993 Ortho 0.993 Radio 0.993 Inter 0.988 Undergrad 0.975 Ortho 0.984 Radio 0.996 Inter 0.974	Dolphin Imaging® software, version 11.5	NP

Study	Randomized	Sample Size Age Syndromic	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Mattos et al.[105] Continued						Tongue level MCSA cross-sectional area Vallecula level minimum cross-sectional area Sagittal area	Undergrad 0.935 Ortho 0.974 Radio 0.987 Inter 0.960 Undergrad 0.993 Ortho 0.984 Radio 0.989 Inter 0.696 Undergrad 0.983 Ortho 0.979 Radio 0.985 Inter 0.977		

Study	Randomized	Sample Size Age	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Mattos et al.[105] Continued						Minimum axial area	Undergrad 0.999 Ortho 0.869 Radio 0.999 Inter 0.932		
						Total airway volume	Undergrad 0.995 Ortho 0.987 Radio 0.994 Inter 0.992		

Study	Randomized	Sample Size Age	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Oh et al.[106]	No	64 8-13 years old (mean age 11.03 years old) Non-syndromic	NP	NP	Upper airway	Nasopharyngeal volume and minimum cross-sectional area	ICC Ranged from 0.969 to 0.998 for all measurements	InVivoDental	-1,024 to -300 Hounsfield units

Study	Randomized	Sample Size Age	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Sears et al.[107]	Reliability sample yes Original sample no	8 Mean age 23.85 years old Non-syndromic	NP	NP	Upper airway	Nasopharyngeal volume Oropharyngeal volume Hypopharyngeal volume	Pearson correlation 0.88 0.97 0.79	CB Works 2.1	NP
Starbuck et al.[108]	Reliability sample yes Original sample no	10 7-18 years old Cleft lip and palate	NP	Semiautomatic segmentation	Upper airway	Nasal airway volume	ICC 0.98	Dolphin Imaging® software, version 11.5	NP

Study	Randomized	Sample Size Age	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Stefanovic et al.[109]	NP	62 Mean age 12.97 years old Non-syndromic extraction group	Age and gender matched non-extraction group Mean age 12.86 years old	NP	Upper airway	Nasopharyngeal volume Oropharyngeal volume Minimum cross-sectional area	ICC >0.98 for all measurements	Dolphin Imaging® software, version 11.0	NP

Study	Randomized	Sample Size Age	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Valladares-Neto et al.[110]	Original sample no Reliability sample yes	13 Mean age 35.5 years old Non-syndromic	NP	NP	Upper airway	Upper volume Lower volume Minimum cross-sectional area	Dahlberg formula −0.41 to 0.56 ml −0.41 to 0.56 ml −22.30 mm^2	InVivoDental software (version 5.0)	NP

Study	Randomized	Sample Size Age	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Vizzotto et al.[111]	No	NP Mean age 17.5 years old Non-syndromic	NP	NP	Upper airway	Nasopharyngeal axial area Oropharyngeal axial area	ICC 0.81-0.95 for all measurements	Image Tool software version 3.0	NP

Study	Randomized	Sample Size Age	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Weissheimer et al.[112]	No	33 7.2-14.5 years old (mean age 10.7 years old) Non-syndromic	NP	Acrylic phantom Semiautomatic segmentation	Upper airway	Oropharyngeal volume	ICC ITK-Snap 0.99 Mimics 0.99 OsiriX 0.99 Dolphin 3D 0.99 InVivoDental 0.99 OnDemand3D 0.94 Mimics FT 1.00	Mimics 14.12, Dolphin Imaging® software, version 11.7, Ondemand3D version 1.0.9.1451, OsiriX version 4.0, ITK-Snap version 2.2.0	FT indicated fixed thresholding was used at −1000 to −587 grey levels

Study	Randomized	Sample Size Age	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Weissheimer et al.[112] Continued							ITK-Snap FT 1.00 OsiriX FT 1.00 OnDemand3D FT 1.00		
Xu et al.[113]	No	62 22-27 years old (mean age 25.8 years old) Cleft lip and palate	23-27 years old (mean age 25.1 years old) Non-syndromic	NP	Upper airway	Total airway volume Minimum cross-sectional area	Pearson correlation coefficient Intra 0.999 Inter 0.999 Intra 0.997 Inter 0.992	Mimics 10.01	NP

Study	Randomized	Sample Size Age	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Yoshihara et al.[114]	Yes	10 Mean age 10.6-14.7 years old Cleft lip and palate	Mean age 10.9-15.4 years old Non-syndromic	NP	Upper airway	Superior oropharyngeal volume Inferior oropharyngeal volume Total airway volume Minimum cross-sectional area	Dahlberg formula Varied from 62.44 to 101.13 mm^3 for volumetric measurements Varied from 3.01 to 5.16 mm^2	3-D Rugle	NP

Study	Randomized	Sample Size Age	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Zhao et al.[115]	No	48 8.9–15.1 years old (mean age 12.8 years old) Non-syndromic requiring RME	8.6–15.8 years old (mean age 12.8 years old) Age and sex matched Non-syndromic not requiring RME	NP	Upper airway	Oropharyngeal airway volume Retropalatal airway volume Retroglossal airway volume Minimum cross-sectional area	ICC Subjects 0.990 Controls 0.991	Vwork version 5.0	NP

Study	Randomized	Sample Size Age	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Zheng et al.[116]	Original sample no Reliability sample yes	15 Mean age 15.65 years old Non-syndromic	NP	NP	Upper airway	Nasopharyngeal volume Oropharyngeal volume Hypopharyngeal Total airway volume Minimum cross-sectional area	Dahlberg formula Ranged from 91.53–152.82 mm^3 for volume measurements Ranged from 9.16 to 33.28 mm^2	CBWorks 2.1	-1024 and -318 Hounsfield units

Study	Randomized	Sample Size Age	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Aboudara et al.[117]	Original sample no Reliability sample yes	10 6-17 years old (mean age 14 years old) Non-syndromic	NP	NP	Upper airway	 Nasopharyngeal volume Nasopharyngeal area	Pearson product correlation, mean percentage error, and mean absolute error >0.9 1.60% 48.7 ± 41.1 mm^3 >0.9 2.00% 6.7 ± 7.6 mm^2	3-D Doctor	NP

Study	Randomized	Sample Size Age	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Haskell et al.[118]	No	26 NP Non-syndromic OSA patients with appliance	Non-syndromic OSA patients without appliance	NP	Upper airway	Total airway volume Minimum cross-sectional area	ICC 0.995 with appliance 0.999 without appliance 0.990 with appliance 0.995 without appliance	Dolphin Imaging® software, version 11.0	NP

Study	Randomized	Sample Size Age	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
Iwasaki et al.[119]	Original sample no Reliability sample yes	10 Mean age 8.8 years old Non-syndromic class III malocclusion	Mean age 8.4 years old Non-syndromic class I malocclusion	NP	Upper airway	Upper airway volume Nasopharyngeal minimum cross-sectional area Oropharyngeal minimum cross-sectional area	ICC and Dahlberg formula 0.975-0.999 162.48 mm^3 0.975-0.999 1.37 mm^2 0.975-0.999 1.69 mm^2	INTAGE Volume Editor	–1024 to –300

Study	Randomized	Sample Size Age	Control	Gold Standard/ Segmentation	Airway Region	Measurements (ie/ volume, minimum cross-sectional area)	Reliability Test Used and Statistics	Imaging Software	Threshold Value(s)
NP, Not Provided.									

Table A.2 CBCT machine settings of the included studies in the systematic review (N=42)

Study	CBCT Machine	Field of View	Tube Current (mA)	Tube Potential (kVp)	Exposure Time (sec)	Resolution/Voxel size (mm)
Alves et al.[79]	i-CAT, Imaging Sciences International	13 × 17 cm	5	120	20	0.4
Alves et al.[80]	i-CAT, Imaging Sciences International	13 × 17 cm	5	120	20	0.4
Bandiera et al.[81]	i-CAT, Imaging Sciences International	13 × 23 cm	36.9	120	40	0.4
Brunetto et al.[82]	i-CAT, Imaging Sciences International	13 × 17 cm	5	120	20	0.4
Burkhard et al.[83]	KaVo 3D Exam, KaVo Dental GmbH	NP	NP	NP	NP	NP
Celikoglua et al.[84]	NewTom 5G	13 cm	NP	NP	14-18	0.3
Chang et al.[85]	Scanora 3D	14.5 × 13.0 cm	NP	125	20	0.35
Cheung and Oberoi[86]	Hitachi MercuRay, Hitachi Medical Corporation	8 x 8 inch	NP	NP	NP	0.4
De Souza et al.[87]	i-CAT, Imaging Sciences International	13 × 23 cm	36.9	120	40	0.4
Di Carlo et al.[88]	NewTom 3G	12 inch	NP	NP	NP	0.36
El and Palomo[89]	Hitachi CB MercuRay, Hitachi Medical Systems America	12 inch	2	120	9.6	0.377
Enciso et al.[90]	Newtom QR 3G	NP	NP	68	NP	NP
Feng et al.[91]	3D eXam, KaVo	NP varied	5	120	14.7	0.2
Glupker et al.[92]	NP	13.3 inches	NP	NP	8.9	0.3
Grauer et al.[93]	i-CAT, Imaging Sciences International	NP	NP	NP	20-38	0.3

Study	CBCT Machine	Field of View	Tube Current (mA)	Tube Potential (kVp)	Exposure Time (sec)	Resolution/Voxel size (mm)
Guijarro-Martinez and Swennen[94]	i-CAT, Imaging Sciences International	17 × 22 cm	48	120	20	0.4
Hart et al.[95]	Iluma Ultra, IMTEC or	19 × 22 cm	3.8	120	40	0.3
	ProMax 3D, Planmeca	17 × 20 cm	14-Jan	90	27	0.2
Hong et al.[96]	Master 3D, Vatech	20 × 19 cm	3.6	90	15	0.3
Iannetti et al.[97]	NP	NP	NP	NP	NP	NP
Iwasaki et al.[98]	CB MercuRay, Hitachi Medical	512 × 512 matrix	15	120	9.6	0.377
Jiang et al.[99]	Galileos, Sirona		7	85	14	0.15
Kim et al.[100]	Master 3D, Vatech	19 × 20 cm	NP	NP	NP	0.3
Kim et al.[101]	Master 3D, Vatech	12 inches	NP	NP	NP	0.3
Kochel et al.[102]	KaVo 3D eXam®, KaVo Dental	23 × 17 cm	3−8	90−120	8.5	0.4
Lenza et al.[74]	Newtom QR 3G	12 inches	NP	NP	NP	0.36
Li, L. et al.[103]	KaVo 3D Exam, KaVo Dental GmbH	NP	5	120	8.9	0.3
Li, YM. et al.[104]	Galileos, Sirona	NP	07-May	85	NP	NP
Mattos et al.[105]	i-CAT, Imaging Sciences International	13 × 17 cm	5	120	20	0.25
Oh et al.[106]	Master 3D, Vatech	20 × 19 cm	3.6	90	15	0.3

Study	CBCT Machine	Field of View	Tube Current (mA)	Tube Potential (kVp)	Exposure Time (sec)	Resolution/Voxel size (mm)
Sears et al.[107]	Hitachi CB MercuRay, Hitachi Medical Systems America	12 inches	10	100	9.6	NP
Starbuck et al.[108]	i-CAT, Imaging Sciences International	13 cm	NP	NP	8.9	0.3 or 0.4
Stefanovic et al.[109]	Hitachi MercuRay, Hitachi Medical Corporation	12 inches	2	120	9.6	0.377
Valladares-Neto et al.[110]	i-CAT, Imaging Sciences International	12 inches	47.7	120	40	0.4
Vizzotto et al.[111]	i-CAT, Imaging Sciences International	13 cm	08-Mar	120	NP	0.25
Weissheimer et al.[112]	i-CAT, Imaging Sciences International	NP	8	120	40	0.3
Xu et al.[113]	3D Accuitomo 170 XYZ slice view tomograph, J Morita Mfg Corp	17 × 12 cm	4.5	85	NP	NP
Yoshihara et al.[114]	CB MercuRay, Hitachi Medical	192.5 mm	15	120	9.6	0.377
Zhao et al.[115]	NewTom 3G	NP	NP	NP	36	NP
Zheng et al.[116]	CB MercuRay, Hitachi Medical	NP	10	110	10	NP
Aboudara et al.[117]	NewTom-9000, Quantitative Radiology	9 X 9 cm	15	110	18	0.3
Haskell et al.[118]	i-CAT, Imaging Sciences International	22 cm	NP	NP	20	0.4
Iwasaki et al.[119]	CB MercuRay, Hitachi Medical	NP	15	120	9.6	0.377
NP, Not Provided.						

Table A.3 Examination characteristics of the included studies in the systematic review (N=42)

Study	Number of Examiners	Number of Times Repeated	Time Period Between Repeated Measurements	Qualifications of Examiners
Alves et al.[79]	1	1	1 week	NP
Alves et al.[80]	1	1	1 week	NP
Bandiera et al.[81]	1	1	30 days	NP
Brunetto et al.[82]	1	1	2 weeks	NP
Burkhard et al.[83]	2	0	0	NP
Celikoglua et al.[84]	1	1	2 weeks	NP
Chang et al.[85]	1	3	1 week	Orthodontist
Cheung and Oberoi[86]	1	1	NP	NP
De Souza et al.[87]	2	1	3 weeks	NP
Di Carlo et al.[88]	1	1	NP	Dentist
El and Palomo[89]	1	1	2 weeks	Orthodontist
Enciso et al.[90]	1	1	60 days	NP
Feng et al.[91]	2	1	NP	NP
Glupker et al.[92]	1	1	2 weeks	Orthodontic resident
Grauer et al.[93]	1	3	NP	NP
Guijarro-Martinez and Swennen[94]	2	1	4 weeks	Oral maxillofacial surgeon (one examiner)
Hart et al.[95]	1	1	NP	Dentist
Hong et al.[96]	1	3	NP	NP
Iannetti et al.[97]	2	2	NP	Dentist and physician
Iwasaki et al.[98]	1	1	1 week	Orthodontist
Jiang et al.[99]	1	1	2 weeks	NP
Kim et al.[100]	1	1	NP	NP
Kim et al.[101]	1	1	2 weeks	NP
Kochel et al.[102]	1	1	2 weeks	NP

Study	Number of Examiners	Number of Times Repeated	Time Period Between Repeated Measurements	Qualifications of Examiners
Lenza et al.[74]	2	1	NP	Orthodontists
Li, L. et al.[103]	1	1	1 month	NP
Li, YM. et al.[104]	2	1	1 week	Orthodontists
Mattos et al.[105]	3	1	2 weeks	An undergraduate student, an orthodontist, and a dental radiologist
Oh et al.[106]	1	2	1 week	Orthodontist
Sears et al.[107]	1	1	2 weeks	NP
Starbuck et al.[108]	1	1	2 weeks	Orthodontist
Stefanovic et al.[109]	1	1	2 weeks	NP
Valladares-Neto et al.[110]	1	1	10 days	Orthodontist
Vizzotto et al.[111]	1	1	15 days	Dental radiologist
Weissheimer et al.[112]	1	1	2 weeks	Orthodontic resident
Xu et al.[113]	2	1	1 month	NP
Yoshihara et al.[114]	1	1	2 weeks	Orthodontic resident
Zhao et al.[115]	1	1	NP	Orthodontist
Zheng et al.[116]	1	1	1 week	NP
Aboudara et al.[117]	1	1	NP	Orthodontist
Haskell et al.[118]	1	2	NP	NP
Iwasaki et al.[119]	1	1	1 week	NP
NP, Not Provided.				

Table A.4 Raw data for threshold value

Scan #	1	2	3	4	5	6	7	8	9	10
A1	45	55	56	55	56	60	55	59	52	44
A2	56	55	55	56	58	61	55	58	53	54
B1	48	54	56	52	60	66	53	60	60	54
B2	50	58	60	57	60	62	59	58	58	54
C1	48	52	51	55	55	55	47	58	56	53
C2	48	56	53	64	60	58	58	55	58	53
D1	55	55	53	56	55	59	50	59	52	50
D2	57	58	56	51	55	55	53	56	51	50
E1	64	62	66	65	64	68	66	66	62	63
E2	64	67	65	67	65	65	66	66	64	62
F1	82	60	56	72	57	67	49	79	55	64
F2	75	63	60	70	60	70	54	66	65	62

Table A.5 Raw data for measured minimum cross-sectional area in mm^2

Scan #	1	2	3	4	5	6	7	8	9	10
A1	131.6	160.9	223.3	360.6	285.4	241.0	514.6	119.9	291.7	334.0
A2	139.6	153.5	217.9	269.9	290.5	138.8	484.1	120.3	296.5	355.1
B1	131.2	158.2	218	357.6	293.6	134.2	437.4	126.5	304.1	363.4
B2	135.1	164.5	235.6	354.4	294.8	143.7	499.9	117.5	284.7	359.6
C1	134.1	128.6	204.3	332	283.7	121.3	397.8	120.3	288.5	156.1
C2	134.6	67.9	214.9	363.7	293.8	137.3	464.8	114	301.1	356.5
D1	142.3	156.5	219.8	333.8	283.6	126.7	396.8	121.5	281.2	340
D2	144.4	172.2	232	354.7	280.9	237	500	116	298.3	342.6
E1	149.8	181.5	250.6	378.9	1960.3	245.9	524.9	134.4	336.9	396.6
E2	149.8	189.6	241.5	385.3	309.7	235.7	561	132.6	330	387.6
F1	175.7	84.6	221.5	388.1	288.5	137.8	411.3	166	276.7	442.9
F2	156.5	72.6	93.8	174.3	130	68.8	89.2	134.8	146.8	134.3

Table A.6 Raw data for measured total upper pharyngeal airway volume in mm^3

Scan #	1	2	3	4	5	6	7	8	9	10
A1	18142.5	21518.0	26567.3	31449.5	35365.6	20576.0	48384.6	21322.8	27739.6	31061.2
A2	20803.9	21392.8	26475.9	30935.3	35772.6	20184.6	48304.2	21613.3	28788.3	32910.1
B1	17433.7	20253.4	25638.3	27699.9	34712.1	20025.7	45887.2	20594.7	27814.5	34471.9
B2	17870.4	20770.3	23517.2	29463.1	33749.7	20040.1	48609.2	19780.7	29345.7	33167.1
C1	18949.3	19580	23051.4	35840.6	34885.8	22252.7	45739.5	22586.2	31990.2	35415.5
C2	17796.9	19104.4	27112.6	34893.2	38568.5	21496.4	50818	21075.3	30585.8	36175.2
D1	20176.6	20063.4	24741.9	36663.4	36711.6	22701	47964.3	23236	30973.1	34383
D2	20391	21034.7	24461.5	33618.9	36548.8	22340.2	48932	22614.3	29390.4	34852.9
E1	21151.8	18377.4	26401.7	30991.8	217481.9	24961.7	52516.5	24859.6	33116.4	38339.9
E2	21608	21867.9	30758.6	39394.6	41750.4	23305.7	54142.5	25156.8	33130.9	38242.5
F1	33252.9	18328.6	21029.4	29329.8	37555.7	23800	47284.8	34484	31430.8	36325.5
F2	31231.9	21425.5	29630.6	35591	37740.4	25109.4	48732.2	24846.5	35079.1	33699.2

Table A.7 Raw data for measured nasopharyngeal airway volume in mm^3

Scan #	1	2	3	4	5	6	7	8	9	10
A1	4917.6	6457.7	3996.1	5254.9	7532.8	9063.2	5134.2	6503.7	6647.7	5099.6
A2	5514.0	5790.6	4269.9	5861.2	7693.3	9099.4	5334.2	7078.9	6681.3	4496.2
B1	4232.1	3947.1	3480.6	5867.9	7120.4	7401.6	6046.7	8316.8	6014.9	5177.9
B2	4333.2	4030.1	3737.8	6284.5	7147.1	6889.3	6276.4	8369.8	5352.8	5171.3
C1	4578.5	3692.5	3216.5	5830.4	6801.6	8504.6	5289.8	7498.5	6018.7	4916.6
C2	4860.6	4697.7	3510.8	6374.1	6971.1	9606.6	7084.2	8723.1	6812.8	5283.6
D1	5122.4	4842.1	3349.2	5584.7	6690.4	9850	5775.4	7673	4871.1	3908
D2	5141.9	4000.4	3945.7	5404.9	6958.1	9608.2	5962.7	7487.2	5389.2	4170.7
E1	6051.6	4155.2	4275.2	4738.1	7507.3	9591.4	5949.9	7787.4	5283.6	5525.9
E2	5574.3	4570.6	4226.1	6055.3	7111.6	6472.3	4711.1	6365.7	4501.6	5196.7
F1	17922.6	5870.1	4107.6	5862.1	7863.8	8104.1	6093.1	11767.2	5974.4	7626.2
F2	8573.8	5674.9	6415.3	6852.6	8010.5	7497.7	6708.7	6634.5	6185.5	6157.6

Table A.8 Raw data for measured oropharyngeal airway volume in mm^3

Scan #	1	2	3	4	5	6	7	8	9	10
A1	9339.0	11019.5	16245.8	20037.2	20871.9	9123.3	35746.4	9454.3	16804.8	20865.2
A2	9231.5	12143.2	16877.8	19769.8	20972.5	9255.7	35031.7	9305.4	16742.3	22756.0
B1	9556.6	13287.3	17046.6	18584.3	21601.1	11108.4	32380.6	7793.2	18484.4	22740.4
B2	9736	13033.8	15185.5	19563.6	22039.7	11062.8	34046.2	7653.1	18685.4	23079.9
C1	9283.4	12588.1	14302.2	19677.9	21494.8	8969.3	33486.4	8834.9	18561.9	23139.3
C2	9140.7	12113.3	17507.9	20423.2	23406.6	9060	35557.5	6985.2	19475.2	23347.1
D1	10048.7	12593.5	15529.6	20378.1	21104.3	8607.6	34362.4	9382.6	18718.4	23823.8
D2	10101	13295.2	14349.3	19231.5	21364.1	8055.6	34652.8	8387.8	12302.2	22933.6
E1	11032.3	13143.1	17579.6	21940.6	23931.8	10440.7	39290.7	11734.5	20749.6	26464.1
E2	11601.7	13813.7	18417.2	22987.2	27276.4	12609.9	40242.3	11967.2	21046.3	25701.5
F1	14988.2	12642.8	17931.8	23823.8	21755.8	11500.1	35360.4	16555.2	19372.1	26140.5
F2	17274.8	12803.9	19143.7	22598.3	24686.8	15793.7	36371.7	11143.1	22288.5	26624.1

Table A.9 Raw data for measured hypopharyngeal airway volume in mm^3

Scan #	1	2	3	4	5	6	7	8	9	10
A1	3999.0	3877.3	6002.6	6417.6	8300.0	2268.8	7349.8	5649.1	5081.2	5789.0
A2	6087.0	3722.9	5199.4	5234.2	6685.7	2279.1	7826.4	5570.2	6089.5	6022.1
B1	4069.8	3642.6	5615.3	2705.2	6184.1	2027.7	7501.5	4189.3	3987.5	6988.3
B2	3996.9	3655	5395.9	4281	4349.2	1949.8	7831.4	4215.1	5680.9	5651.1
C1	5068.3	3081.6	5694.4	10248.5	7787	4763.1	7869.6	6473.8	6941.5	6773.6
C2	3933.4	2097.7	4511.4	7796.6	6866.2	3401.7	8030.4	5624.7	3684.8	7434.2
D1	4829.4	3551.3	3880.1	11497.7	8184	2339.6	8518	7337.6	6839.1	7590.7
D2	5249.1	3523.7	4572.5	9319.7	8272.2	4655.6	8488.8	6766.9	11835	7298.8
E1	6028.1	2828.2	3517.8	4601.1	7753.9	4975.8	9419.7	7148.1	7131.1	8888.4
E2	5555.9	3828.8	5100.4	9233.4	7948.9	4479.6	8726	7510.7	8167.4	9281.8
F1	3979.2	4673	3382	7265.1	8597.2	3785.8	8758.3	7336.3	7105.6	5398.9
F2	4705.8	3476	6828.4	9957.8	9045.2	4282.7	9299.9	5958.1	6942.3	7850.3

Figure A.1 Landmarks used for hard tissue orientation of the CBCT scans

Opisthion	On the occipital bone, the midpoint on the posterior margin of the foramen magnum
Incisive Foramen	The opening in the hard palate immediately behind the maxillary incisor teeth
Porion	The point on the cranium located at the upper margin of each ear canal (external auditory meatus)
Orbitale	A point midway between the lowest point on the inferior margin of the two orbits

Figure A.2 Examiner data collection form

Scan #	1	2	3	4	5	6	7	8	9	10
Threshold value										
Minimum cross-sectional area (mm^2)										
Total airway volume (mm^3)										
Nasopharyngeal airway volume (mm^3)										
Oropharyngeal airway volume (mm^3)										
Hypopharyngeal airway volume (mm^3)										