Dentomaxillofacial Radiology (2012) 41, 276–284

http://dmfr.birjournals.org

REVIEW ARTICLE

Three-dimensional segmentation of the upper airway using cone beam CT: a systematic review

NA Alsufyani*, C Flores-Mir and PW Major

Department of Dentistry, Faculty of Medicine and Dentistry, University of Alberta, Edmonton, AB, Canada

The objectives of this study were to systematically review the literature for studies that used cone beam CT (CBCT) to automatically or semi-automatically model the upper airway (including the pharyngeal, nasal and paranasal airways), and to assess their validity and reliability. Several electronic databases (MEDLINE®, MEDLINE In-Process & Other Non-Indexed Citations, all evidence-based medicine reviews including the Cochrane database, and Scopus) were searched. Abstracts that appeared to meet the initial selection criteria were selected by consensus. The original articles were then retrieved and their references were searched manually for potentially suitable articles that were missed during the electronic search. Final articles that met all the selection criteria were evaluated using a customized evaluation checklist. 16 articles were finally selected. From these, five scored more than 50% based on their methodology. Although eight articles reported the reliability of the airway model generated, only three used intraclass correlation (ICC). Two articles tested the accuracy/validity of airway models against the gold standard, manual segmentation, using volumetric measurements; however, neither used ICC. Only three articles properly tested the reliability of the three-dimensional (3D) upper airway model generated from CBCT and only one article had sufficiently sound methodology to test the airway model's accuracy/validity. The literature lacks proper scientific justification of a solid and optimized CBCT protocol for airway imaging. Owing to the limited number of adequate studies, it is difficult to generate a strong conclusion regarding the current validity and reliability of CBCT-generated 3D models.
Dentomaxillofacial Radiology (2012) **41,** 276–284. doi: 10.1259/dmfr/79433138

Keywords: image reconstruction; upper airway; cone beam computed tomography; review

Introduction

Obstruction of the upper airway often alters normal breathing, which can have a significant impact on the normal development of craniofacial structures.[1,2] Narrow maxillary arch, cross bites, clockwise mandibular growth rotation and mandibular retrognathia have been reported as being associated with chronic mouth breathing.[3] Many of these facial features have also been reported in subjects with sleep-disordered breathing such as obstructive sleep apnoea.[4,5] Such abnormalities require prompt attention, and an early diagnosis is imperative to ensure normal craniofacial development.[6]

Cone beam CT (CBCT) has become available for oral and maxillofacial imaging. It has been suggested that CBCT provides an accurate, efficient modality involving relatively less radiation than multidetector CT for improved understanding of airway anatomy, pathology and upper airway mechanics.[7,8]

Segmentation of the airway can be achieved manually or automatically. Manual segmentation seems to be the most accurate method and allows for the most operator control.[9] Accordingly, it is significantly time consuming because it requires the operator to outline the airway boundaries on each slice and then transform the data into a three-dimensional (3D) volume. Automatic segmentation, on the other hand, can drastically reduce segmentation time.[9]

Automatic or semi-automatic 3D segmentation of the upper airway can be very challenging, especially in the complex anatomy of the nasal airway. It has been

*Correspondence to: Dr Noura Alsufyani, Edmonton Clinic Health Academy, Department of Dentistry, 5th floor, 11405-87 AVE NW, Edmonton T6G 1C9, AB, Canada. E-mail: nourasuf@gmail.com
Received 19 September 2011; revised 8 November 2011; accepted 10 November 2011

noted that several studies assessing the use of CBCT scans to segment the airway did not provide validation of their proposed methods.[6,10–12] Any model of the upper airway reconstructed to accurately study the possible relationship between airway restriction and craniofacial growth using CBCT imaging must be both reliable and valid.

The purpose of this study is, therefore, to systematically review the medical and dental literature for studies using CBCT to automatically or semi-automatically model the upper airway (including the pharyngeal, nasal and paranasal airways) and to answer the following questions: Are 3D airway models automatically segmented from CBCT accurate and reliable? Can clinicians and surgeons use quantitative analysis based on these models?

Methods

A systematic search of multiple electronic databases was completed during the third week of May 2011. Databases searched were MEDLINE® (including MEDLINE In-Process & Other Non-Indexed Citations),[13] all evidence-based medicine (EBM) reviews (including the Cochrane library[14]) and Scopus.[15] Each database was searched with the following search terminologies (adapted to the requirements of the database): "airway OR upper OR nasal OR pharynx" and "segmentation OR reconstruction OR algorithm OR three dimensional imaging" and "cone beam computed tomography OR computed tomography". An example of the search terminology used in MEDLINE is summarised in Table 1. The search terminology for all EBM reviews and Scopus is provided in Appendix 1.

Two reviewers conducted the selection process independently. In case of disagreement, discussion between the reviewers was favoured to reach a consensus.

The first phase of the selection process involved reviewing the titles and abstracts of the potential articles according to the following inclusion criteria:

- upper airway assessment
- use of CBCT.

The second phase consisted of a detailed review of the entire retrieved article as selected in Phase 1. In addition to the initial selection criteria, two more were added at this stage:

- Only studies that involved an automated or semi-automated 3D/volumetric segmentation of the upper airway were selected.
- In studies involving a physical/geometric model, the design of the airway model must mimic the possible different diameters/shapes or angles of the human airway.

Finally, a manual search of potentially missing articles was completed using the references/bibliography of the articles identified at Phase 2. In addition, the authors of the selected studies were contacted to inquire about missing or incomplete data.

A customized systematic evaluation protocol (Table 2) was created to systematically assess the selected studies. For example, a study that included a randomized sample of human subjects (≥30), preferably included a test group with abnormal airways, used manual segmentation as a gold standard, analysed the entire upper airway with several types of measurements and executed proper statistical analyses would score higher and would be considered scientifically superior to another study that scored fewer points. Accordingly, any conclusions drawn from any of these articles had to be based on studies that scored more points, *i.e.* were superior in design and analysis. Because the accuracy of an airway model should be checked against a gold standard, ideally manual segmentation, by means of reliable measurements, more points were allocated to the criteria "study measurements" and "data analysis" (Table 2). "Study design" was taken into consideration but was allocated fewer points. No effort was made to validate this evaluation tool. The parameters of CBCT scan protocol used in the final selected articles were also collected.

Results

Database search

The search results and the number of articles at each phase from the various databases are provided in Table 3. Comparing the final results of the different databases, Scopus obtained the most articles (68.75%), whereas all EBM reviews originally obtained 15 potential studies, but none was deemed useful per our selection criteria. By the end of Phase 2, 10 studies were excluded owing to either duplication or selection criteria, and only 12 met the selection criteria. Finally, manual searching of

Table 1 Example of search terminology in MEDLINE®[13]

	Keywords	*Number of articles*
(1)	airway.mp.	94 883
(2)	exp Pharynx/or upper.mp.	39 104
(3)	nasal.mp. or exp Nasal Cavity/	80 344
(4)	(1) or (2) or (3)	194 220
(5)	exp Algorithms/or segmentation.mp.	
	exp Tomography, X-Ray Computed/or cone beam computed	159 925
(6)	tomography.mp. or exp Cone-Beam Computed Tomography/	243 627
(7)	(4) and (5) and (6)	138

Table 2 Evaluation checklist for the final selected studies

Parameters of evaluation			*Maximum score*
1. Study design	(a)	Randomized sample (✓)	1
	(b)	Sample size ≥30 (✓)	1
	(c)	Test group included (✓)	1
	(d)	Physical model (✓) Human (✓)	2
	(e)	Method of segmentation: Algorithm (✓) Commercial software (✓)	1
2. Study measurement	(f)	Validation/gold standard: Physical model (✓✓✓) Manual segmentation (✓✓✓✓)	4
	(g)	Part of airway: Oropharynx/nasopharynx (✓) Nasal cavity (✓) Paranasal sinuses (✓)	3
	(h)	Type of measurement: Linear (✓) Area (✓) Shape (✓) Volume (✓✓)	5
3. Data analysis	(i)	Reliability: Intra-examiner (✓) Inter-examiner (✓) Kappa or ICC (✓) Other statistical test: Appropriate (✓) *p*-value, r^2 reported (✓)	5
Total (✓)			23

the references from articles identified at Phase 2 obtained four additional studies. Therefore, the final number of articles deemed useful was 16.

Article scores and evaluation

The application of the customized evaluation tool is presented in Table 4. Results and conclusions of articles that scored more than or equal to 50% were given more weight because these studies presented more accurate methodology than other studies. Only five articles[9,12,16–18] scored more than 50%. The study by El and Palomo[9] presented the highest score (69.57%).

Detailed analysis of the 16 articles is summarised in Table 5. Although there were a few articles published in 2006 and 2007, the majority of the articles were recent (2009–2011). Most of the studies were not randomized, included subjects with normal/healthy airways and used software products for the segmentation process. Only two studies[9,12] used manual segmentation as a gold standard to validate their measurements and only eight[9,16–20,22,23] reported the reliability of their measurements. The majority of the articles included analysis of oronasopharyngeal airways and one or more measurements (linear, area or volume). In terms of data analysis, most studies did not score highly for reliability or appropriate statistical tests. Two studies[7,25] were computational fluid dynamic (CFD) studies and did not include such measurements or statistical analysis.

Cone beam CT scan protocol

CBCT scan parameters/protocol used for each study are presented in Table 6. The most common CBCT machines used were iCAT® (Imaging Sciences International, Hatfield, PA) and CB MercuRay™ (Hitachi Medical, Tokyo, Japan), in five articles each. The NewTom 3G (ImageWorks, Elmsford, NY) and Picasso Master 3D CBCT systems (VATECH, Seoul, Republic of Korea) were also used in one article each. The remaining four articles did not mention the CBCT machine used in their protocol. In studies that reported their scanning protocol, the field of view (FOV) ranged from 13 cm to 30.5 cm, the tube current from 2 mA to

Table 3 Number of articles per database

Database	*MEDLINE®[13]*	*Evidence-based medicine reviews*	*Scopus[15]*	*Manual search*	*Total articles*
Initial search	138	15	75	–	228
Phase I	4[a]	0	18	–	22
Phase II	1	0	18[b]	–	19
Final selection	1	0	11	4	16
Contribution of database to final selection (%)	6.25%	0%	68.75%	25%	

[a]Three articles excluded at next phase (duplicates).
[b]Seven articles excluded at next phase (violated selection criteria).

Table 4 Scores of the final 16 selected studies using checklist

	Parameters of scoring (x: maximum score)									
	Study design					*Study measurements*			*Data analysis*	
Studies evaluated	*(a) = 1*	*(b) = 1*	*(c) = 1*	*(d) = 2*	*(e) = 1*	*(f) = 4*	*(g) = 3*	*(h) = 5*	*(i) = 5*	*Total score, n (% out of 23)*
El and Palomo (2011)[9]	1	1	0	2	1	4	2	2	3	16 (69.57)
Shi et al (2006)[12]	1	0	0	2	1	4	1	4	0	13 (56.52)
Lenza et al (2010)[16]	0	1	0	2	1	0	1	4	4	13 (56.52)
Haskell et al (2009)[17]	0	0	0	2	1	0	1	5	4	13 (56.52)
Iwasaki et al (2009)[18]	0	1	0	2	1	0	2	5	2	13 (56.52)
Grauer et al (2009)[19]	0	1	0	2	1	0	1	3	3	11 (47.82)
Kim et al (2010)[20]	0	0	0	2	1	0	2	4	2	11 (47.82)
Tso et al (2009)[21]	1	0	0	2	1	0	1	4	2	11 (47.82)
Iannetti et al (2011)[22]	0	0	1	2	1	0	1	3	3	11 (47.82)
Schendel and Hatcher (2010)[23]	0	0	1	2	1	3	1	1	0	9 (39.13)
El et al (2011)[24]	0	0	1	2	1	0	1	4	0	9 (39.13)
Iwasaki et al (2011)[25,a]	0	1	1	2	1	0	3	0	0	8 (34.78)
Schendel et al (2011)[26]	0	0	0	2	1	0	1	3	0	7 (30.43)
Cheng et al (2007)[10]	0	0	0	2	1	0	1	3	0	7 (30.43)
Huynh et al (2009)[7,a]	0	0	0	2	1	0	1	0	0	4 (17.39)
Celenk et al (2009)[27]	0	0	0	2	1	0	1	0	0	4 (17.39)

[a]Computational fluid dynamics studies.

15 mA and the tube potential from 110 kVp to 120 kVp. The scanning time varied from 10 s to 40 s, the resolution from 0.25 mm to 0.6 mm and the voxel size from 3.527×3.527 pixels to 1024×1024 pixels.

Discussion

CBCT technology has introduced a paradigm shift in oral and maxillofacial imaging by transitioning from two-dimensional to 3D imaging. 3D segmentation of the upper airway using CBCT paved the way to studying the anatomy and function of narrowed airways in subjects with sleep disordered breathing, *e.g.* obstructive sleep apnoea (OSA), in ways that were unattainable before.[28] Most 3D airway models generated from CBCT have not been validated in the literature.[9]

Study design

Only a few of the studies[9,12,21] were randomized, and a sample size of more than 30 was found in only 5 studies.[9,16,18,19,25] Accordingly, these studies should have had less bias in their measurements. In three articles,[10,16,27] the authors attempted to use reconstruction algorithms instead of commercial software. Because the main purpose of these studies was to develop new or modify previous reconstruction methods using different algorithms, the sample size was smaller than for studies that used commercial software, and study measurements were limited if not absent. Most articles analysed subjects with healthy upper airways. Five articles[7,12,23,25,26] analysed constricted airways of subjects with OSA. Studies by Lenza et al,[16] Cheng et al[10] and Celenk et al[27] did not report whether the analysed airway was that of a healthy subject or an OSA patient. None of the 16 studies compared the accuracy of the 3D airway model between OSA patients and their healthy controls. If an OSA patient was an obligatory mouth breather or if the CBCT scan time was long, the patient would undergo multiple breathing cycles, thus causing some motion artefact that can affect the resolution of the airway boundaries. This technical difficulty was not addressed by any of the studies that included OSA patients.

Study measurements

Validity is defined in this systematic review as agreement in measurements between the software or segmentation algorithm and the gold standard/ground truth. Reliability or reproducibility is defined as the agreement between measurements for the same examiner (intra-examiner) or between different examiners (inter-examiner) using a commercial software or reconstruction algorithm.

Of the 16 articles, only 3[9,12,23] tested their measurements against a gold standard. Of these, El and Palomo[9] and Shi et al[12] used manual segmentation as a reference. El and Palomo[9] validated their measurements for the entire sample of 30, whereas Shi et al[12] validated their measurements with manual segmentation in only one case. Grauer et al[19] stated that the segmentation process/software they used was described and validated previously by Yushkevich et al[29] and was superior to manual segmentation. However, Yushkevich et al[29] validated InsightSNAP® version 1.4.0 (Cognitica, Philadelphia, PA) using MRI not CBCT. Clearly, MRI and CBCT are very different imaging modalities with different image resolutions that can affect the accuracy of segmentation significantly. 3D airway models generated from CBCT are being introduced as objective evaluation tools of surgical treatment of OSA patients, orthognathic surgeries and maxillary expansion and their impaction on airway dimensions. This necessitates a proper, scientific validation of the method used to generate this model, given that it serves as baseline for treatment.

Table 5 Analysis of study methodology for the selected 16 articles

	Study design	*Study measurements*	*Data analysis*
Study	*Sample* *Subjects* *Software/algorithm*[a]	*Gold standard* *Airway region* *Measurements*	*Reliability (IER)* *Other statistical test*
El and Palomo (2011)[9]	Randomized ($n = 30$)	Manual segmentation: OrthoSegment	ICC
	Normal airway	ONpharynx and part of nasal cavity	Not appropriate
	Dolphin3D®, InVivoDental®, OnDemand3D®	Volume	
Shi et al (2006)[12]	Randomized ($n = 20$)	Manual segmentation (for one case only)	NR
	NR	ONpharynx	Not appropriate
	Algorithm (Visual C++) VTK program language	Linear, area and volume	
Lenza et al (2010)[16]	Not randomized ($n = 34$)	None	Dahlberg's formula and ANOVA
	Normal airway	ONpharynx	Appropriate
	Mimics®	Linear, area and volume	
Haskell et al (2009)[17]	Not randomized ($n = 26$)	None	ICC
	OSA	ONpharynx-nasopharynx	Appropriate
	Dolphin3D, Image J®	Linear, area, shape and volume	
Iwasaki et al (2009)[18]	Not randomized ($n = 45$)	None	Paired t-test, correlation r, Dahlberg's formula
	Normal airway	ONpharynx	Not appropriate
	INTAGE Volume Editor®	Linear, area, shape and volume	
Grauer et al (2009)[19]	Not randomized ($n = 62$)	None	COV
	Normal airway	ONpharynx	Appropriate
	InsightSNAP®	Shape and volume	
Kim et al (2010)[20]	Not randomized ($n = 27$)	None	ICC
	Normal airway	ONpharynx and nasal cavity	Appropriate
	InVivoDental	Linear, shape and volume	
Tso et al (2009)[21]	Randomized ($n = 10$)	None	NR
	Normal airway	ONpharynx	Appropriate
	CBWorks®	Linear, area and volume	
Iannetti et al (2011)[22]	Not randomized ($n = 4$)	None	Wilcoxon signed rank test
	Craniofacial syndromic malformations	Nasal cavity	Not appropriate
	Dolphin3D	Area and volume	
Schendel and Hatcher (2010)[23]	Not randomized ($n = 1$)	Phantom	Mentioned (phantom), NR (OSA patient)
	OSA	ONpharynx	
	Airway phantom	Measurements:	Appropriate (phantom), NR (OSA patient)
	3dmDVultus®	Phantom; linear, area and volume OSA patient; area only	
El et al (2011)[24]	Not randomized ($n = 1$)	None	NR
	OSA	ONpharynx	NR
	Dolphin3D and OnDemand3D	Linear, area and volume	
Iwasaki et al (2011)[25,b]	Not randomized ($n = 40$)	None	NR
	Normal airway	ONpharynx nasal cavity and paranasal sinuses	NR
	INTAGE Volume Editor/refined by algorithm	None/CFD study	
Schendel et al (2011)[26]	Not randomized ($n = 1$)	None	NR
	OSA	ONpharynx	NR
	3dmdVultus®	Area and volume	
Cheng et al (2007)[10]	Not randomized ($n = 1$)	None	NR
	NR	ONpharynx	NR
	Algorithm (modified GVF snakes)	Area and volume	
Huynh et al (2009)[7,b]	Not randomized ($n = 4$)	None	NR
	OSA	ONpharynx	NR
	V-Works®, ImageJ, Pro/engineer®	None/CFD study	
Celenk et al (2009)[27]	Not randomized ($n = 1$)	None	NR
	NR	ONpharynx	NR
	Algorithm (3D Gaussian smoothing kernel, 3D PCA C^{++} programming)	None	

3D, three-dimensional; ANOVA, analysis of variance; CFD, computational fluid dynamics; COV, coefficient of variation; GVF, gradient vector flow; ICC, intracorrelation coefficient; IER, inter-/intra-examiner reliability; NR, not reported; ONpharynx, oronasopharynx; OSA, obstructive sleep apnoea.
[a]Refer to original articles for manufacturer details of software products.
[b]CFD studies.

Table 6 CBCT scan protocol collected from the selected 16 articles

Study	*CBCT machine*[a]	*FOV*	*Tube current (mA)*	*Tube potential (kVp)*	*Time (s)*	*Resolution*
El and Palomo (2011)[9]	Hitachi CB Mercuray	12″	2	120	NR	1024 × 1024 pixels
Shi et al (2006)[12]	iCAT	22 cm	NR	NR	20 + 20	0.4 mm voxels
Lenza et al (2010)[16]	iCAT	NR	3–6	120	20	0.4 mm voxels
Haskell et al (2009)[17]	NewTom	12″	1–4	110	36/5.4 exposure	0.36 mm voxels
Iwasaki et al (2009)[18]	Hitachi CB Mercuray	NR	15	120	9.6	0.377 mm voxels
Grauer et al (2009)[19]	iCAT	Medium or full	NR	NR	NR	0.3 mm voxels
Kim et al (2010)[20]	Master 3D dental imaging system	12″	NR	NR	NR	NR
Tso et al (2009)[21]	Hitachi CB Mercuray	19 cm	10	120	10	0.6 mm voxels
Iannetti et al (2011)[22]	NR	NR	NR	NR	NR	NR
Schendel and Hatcher (2010)[23]	iCAT	13 cm	NR	NR	40	0.25 mm voxels
El et al (2011)[24]	Hitachi CB Mercuray	NR	15	120	9.6	0.377 mm voxels
Iwasaki et al (2011)[25,b]	Hitachi CB Mercuray	12″	15	120	NR	1024 × 1024 pixels
Schendel et al (2011)[26]	NR	NR	NR	NR	NR	NR
Cheng et al (2007)[10]	NR	NR	NR	NR	NR	3.527 × 3.527 pixels
Huynh et al (2009)[7,b]	iCAT	23 × 19 cm	NR	NR	NR	0.4 mm voxels
Celenk et al (2009)[27]	NR	NR	NR	NR	NR	NR

CBCT, cone beam CT; FOV, field of view; NR, not reported.
[a]Refer to original articles for manufacturer details of software products.
[b]CFD studies.

Schendel and Hatcher[23] used the measurements of an airway phantom as validation. However, the true complex anatomy of the human airway cannot be replicated and measured physically, hence the use of the airway phantom with uniform geometry. Therefore, manual segmentation, which by default should better represent ground truth, would be the ideal gold standard for segmentation, especially in the nasopharynx and nasal cavity.

Most authors analysed the pharyngeal airway with volumetric measurements. Only Iannetti et al,[22] Iwasaki et al[25] and El and Palomo[9] segmented the nasal cavity and/or maxillary sinuses. The shape of the oropharyngeal airway is similar to a tube and is completely hollow. This makes the process of segmentation straightforward. The anatomy of the nasal cavity is complicated, with the narrow and tortuous pathways of the conchae and meatuses; consequently, the segmentation process is extremely challenging, owing to difficulties encountered in defining the boundaries and grey level thresholding, especially with noisy CBCT images. Therefore, studies that focus only on the oropharyngeal airway will likely over-represent the true validity of the evaluated tools.

Three studies[7,25,27] failed to report any linear, area or volumetric measurements. Studies by Iwasaki et al[25] and Huynh et al[7] were CFD studies where the measurement of airflow, velocity, pressure and resistance were the parameters of concern. CFD studies simulate airflow in the airway to assess the functional changes in the airway rather than undertake anatomical and/or visual analysis of the airway. The main focus of the study by Celenk et al[27] was to develop a user-friendly method to detect and construct a 3D human airway using CBCT. While the use of a 3D Gaussian smoothing kernel seemed very promising, the authors could have attempted to validate their proposed method by comparing area and volumetric measurements of the airway against manual segmentation.

Data analysis

The quality of statistical analysis used in the majority of the articles was poor. To measure the validity, the intraclass correlation coefficient (ICC) is the most appropriate statistical tool. The ICC is a general measurement of agreement or consensus. It is an improvement over Pearson's r and Spearman's ρ, because it takes into account the differences in ratings along with the correlation between raters.[30,31]

El and Palomo[9] validated their human airway model by means of volumetric measurement against manual segmentation; however, they did not use the ICC. Instead, they used linear regression analysis, the paired *t*-test and Pearson's *r* to validate their measurements. Shi et al[12] validated their human airway model by means of linear measurement against manual measurement; however, they did not use the ICC. The authors used the paired *t*-test to report the differences in linear measurements. Linear regression analysis provides information about the linear relationship or correlation between two random variables, not the agreement.[32]

To measure reliability or reproducibility, the ICC is the most appropriate test tool. Only three studies[9,17,20] used the ICC for intraexaminer agreement; however, none reported the 95% confidence interval (CI) of the ICC. The lower limit of the CI of the ICC indicates how small the examiner agreement might be. For example, if the ICC yielded less than or equal to 0.80 and the lower bound of the CI was 0.60, this does not necessarily imply good agreement. Lenza et al[16] and Iwasaki et al[18] used Dahlberg's formula to detect errors between measurements. Springate[33] examined the use of Dahlberg's formula to estimate errors and found that it can under- or overestimate the true value of the random error.

Grauer et al[19] used coefficient of variation (COV) and Iannetti et al[22] used the Wilcoxon signed-rank test to measure the reliability of volumetric measurements. The COV is a measure of dispersion and the Wilcoxon signed-rank test detects differences in the means; however, none of these tests measures "agreement".

In terms of statistical tests used to analyse the possible relationship and/or correlation between the different airway dimensions and craniofacial parameters, most authors mistakenly used univariate statistical tests (*e.g.* the *t*-test) for each variable instead of multivariate analysis for all the variables tested. In doing so, the alpha (Type II error) was inflated and possible intercorrelation between the variables was ignored. Haskell et al[17] used multiple linear regression to analyse 7 predictors and over 12 outcome variables, and had a sample size of only 26. This may have affected the power of the regression model.

To summarise, El and Palomo[9] were the only authors to test the accuracy of airway models against manual segmentation. In their study, they reconstructed the pharyngeal airway and part of the nasal cavity separately, and used volumetric measurements in 30 CBCT image sets. They concluded that the volumetric measurements of the three software products tested [Dolphin3D version 11 (Dolphin Imaging & Management Solutions, Chatsworth, CA), InVivoDental version 4.0.70 (Anatomage, San Jose, CA) and OnDemand3D™ version 1.0.1.8407 (CyberMed, Seoul, Republic of Korea)] were reproducible and had high correlation with the measurements of manual segmentation, but were not valid. In other words, the software consistently over- or underestimated the true, manual volumetric measurements—hence the high correlation—which were therefore not accurate, suggesting "systematic errors". However, it should be noted that the authors measured "linear" correlation (by using linear regression) instead of ICC to measure the validity. The largest difference was found between the OnDemand3D system and manual segmentation in the oropharyngeal airway volume: $-2163.25\,\text{mm}^3$ (95% CI = $-2945.69\,\text{mm}^3$, $-1380.80\,\text{mm}^3$). Although this difference was found to be statistically significant, it is uncertain whether it would be of clinical significance. It is unclear whether linear and volumetric measurements are sufficient parameters that can be used to validate 3D airway models or be accurate indicators/predictors of surgical outcomes. Perhaps it would be more meaningful if airway models were analysed based not only on measurements but also on geometrical assessment using shape analysis.

Cone beam CT protocol

When reported, most of the studies used CBCT machines that required the patient to be seated. Lenza et al[16] used the NewTom system, which required the patient to be supine. It has been shown that the dimension of the airway changes when patients move from a sitting to a supine position, mostly owing to the relaxation of the soft palate and the tongue, and the change in hyoid bone position.[34] Given that patients are awake during the CBCT scan and sleeping conditions are not simulated, the airway should, in our opinion, be imaged while patients are seated.

Studies that reported the FOV of the CBCT used FOVs ranging from 13 cm to 30.5 cm. A 13 cm FOV is acceptable to image one part of the upper airway (oropharynx or nasal cavity); larger dimensions are satisfactory to image the entire upper airway (superior limits of nasal cavity to epiglottis inferiorly). A tube potential of 120 kVp was used in most studies, and tube current ranged from 1 mA to 15 mA. Whether the tube potential or tube current was fixed or adjustable depended on the CBCT machine used. In the cases in which these parameters were adjustable, none of the authors explained why they selected these specific scanning parameters. The scan time varied greatly between studies (9.6–40 s) and the voxel size ranged from 0.25 mm to 0.6 mm. This was also dependant on the CBCT machine in addition to the operator's selection, and was not explained or justified by any of the authors. Conceptually, increasing the tube potential, tube current and scan time and reducing the voxel size would gain the highest resolution for optimum segmentation of the airway, albeit at the expense of a radiation dose to the patient.[35] This was not addressed in any of the articles included.

In conclusion, only 3 articles of the total 16 adequately tested the reliability of 3D upper airway models generated from CBCT, and only 1 article had a sound methodology to test their accuracy. The literature lacks scientific justification of a solid and optimized CBCT protocol for airway imaging. Owing to the limited number of adequate studies, it is difficult to generate a strong conclusion regarding the validity and reliability of CBCT-generated 3D models.

References

1. Abramson Z, Susarla SM, Lawler M, Bouchard C, Troulis M, Kaban LB. Three-dimensional computed tomographic airway analysis of patients with obstructive sleep apnea treated by maxillomandibular advancement. *J Oral Maxillofac Surg* 2011; **69**: 677–686.
2. McCrillis JM, Haskell J, Haskell BS, Brammer M, Chenin D, Scarfe WC, et al. Obstructive sleep apnea and the use of cone beam computed tomography in airway imaging: A review. *Semin Orthod* 2009; **15**: 63–69.
3. Aboudara CA, Hatcher D, Nielsen IL, Miller A. A three-dimensional evaluation of the upper airway in adolescents. *Orthod Craniofac Res* 2003; **6** Suppl 1:173–175.
4. Miles PG, Vig PS, Weyant RJ, Forrest TD, Rockette HE Jr. Craniofacial structure and obstructive sleep apnea syndrome-a qualitative analysis and meta-analysis of the literature. *Am J Orthod Dentofacial Orthop* 1996; **109**: 163–172.
5. Yu X, Fujimoto K, Urushibata K, Matsuzawa Y, Kubo K. Cephalometric analysis in obese and nonobese patients with obstructive sleep apnea syndrome. *Chest* 2003; **124**: 212–218.
6. 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 Dentofacial Orthop* 2009; **135**: 468–479.
7. Huynh J, Kim D, McQuilling M. Pharyngeal airflow analysis in obstructive sleep apnea patients pre- and post-maxillomandibular advancement surgery. *J Fluids Eng* 2009; **131**: 091101–091111.
8. Osorio F, Perilla M, Doyle DJ, Palomo JM. Cone beam computed tomography: An innovative tool for airway assessment. *Anesth Analg* 2008; **106**: 1803–1807.
9. El H, Palomo JM. Measuring the airway in 3 dimensions: A reliability and accuracy study. *Am J of Orthod Dentofacial Orthop* 2010; **137**: S50.e1–S50.e9.
10. Cheng I, Nilufar S, Flores-Mir C, Basu A. Airway segmentation and measurement in CT images. In: 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; 23–26 August, 2007; Lyon, France. Lyon, France: IEEE EMBC; 2007.
11. Ogawa T, Enciso R, Shintaku WH, Clark GT. Evaluation of cross-section airway configuration of obstructive sleep apnea. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 2007; **103**: 102–108.
12. Shi H, Scarfe W, Farman A. Upper airway segmentation and dimensions estimation from cone-beam CT image datasets. *Int J CARS* 2006; **1**: 177–186.
13. nlm.nih.gov [homepage on the Internet]. Bethesda, MD: United States National Library of Medicine. Available at: http://www.ncbi.nlm.nih.gov/pubmed/
14. thecochranelibrary.com [homepage on the Internet]. Oxford, UK: Cochrane Collaboration; 1993. Available from: http://www.thecochranelibrary.com/view/0/index.html
15. scopus.com [homepage on the Internet]. Amsterdam, Netherlands: Elsevier; 2004. Available from: http://www.scopus.com/home.url
16. Lenza MG, de O. Lenza MM, 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**: 96–105.
17. 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**: 132–158.
18. 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 Dentofacial Orthop* 2009; **136**: 318.e1–318.e9.
19. 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 Dentofacial Orthop* 2009; **136**: 805–814.
20. 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 Dentofacial Orthop* 2010; **137**: 306.e1–306.e11.
21. Tso HH, Lee JS, Huang JC, Maki K, Hatcher D, Miller AJ. Evaluation of the human airway using cone-beam computerized tomography. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 2009; **108**: 768–776.
22. Iannetti G, Polimeni A, Pagnoni M, Fadda MT, Ramieri V, Tecco S, et al. Upper airway volume after le fort III advancement in subjects with craniofacial malformation. *J Craniofac Surg* 2011; **22**: 351–355.
23. Schendel SA, Hatcher D. Automated 3-dimensional airway analysis from cone-beam computed tomography data. *J Oral Maxillofac Surg* 2010; **68**: 696–701.
24. El AS, El H, Palomo JM, Baur DA. A 3-dimensional airway analysis of an obstructive sleep apnea surgical correction with cone beam computed tomography. *J Oral Maxillofac Surg* 2011; **69**: 2424–2436.
25. Iwasaki T, Saitoh I, Takemoto Y, Inada E, Kanomi R, Hayasaki H, et al. Evaluation of upper airway obstruction in class II children with fluid-mechanical simulation. *Am J Orthod Dentofacial Orthop* 2011; **139**: e135–e145.
26. Schendel S, Powell N, Jacobson R. Maxillary, mandibular, and chin advancement: Treatment planning based on airway anatomy in obstructive sleep apnea. *J Oral Maxillofac Surg* 2011; **69**: 663–676.
27. Celenk M, Farrell ML, Eren H, Kumar K, Singh GD, Lozanoff S. Upper airway detection and visualization from cone beam image slices. *J Xray Sci Technol* 2010; **18**: 121–35.
28. Hatcher DC. Cone beam computed tomography: Craniofacial and airway analysis. *Sleep Med Clin* 2010; **5**: 59–70.
29. Yushkevich PA, Piven J, Hazlett HC, Smith RG, Ho S, Gee JC, et al. User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *Neuroimage.* 2006; **31**: 1116–1128.
30. Koch, Gary G. Intraclass correlation coefficient. In: Kotz S, Johnson NL. Encyclopedia of Statistical Sciences. 4th edn. Hoboken, NJ: John Wiley & Sons; 1982:213–217.
31. Müller R, Büttner P. A critical discussion of intraclass correlation coefficients. *Stat Med* 1994; **13**: 2465–2476.
32. Zou KH, Tuncali K, Silverman SG. Correlation and simple linear regression. *Radiology* 2003; **227**: 617–622.
33. Springate SD. The effect of sample size and bias on the reliability of estimates of error: A comparative study of Dahlberg's formula. *Eur J Orthod*; in press. Epub ahead of print March 2011.
34. Battagel JM, Johal A, Smith AM, Kotecha B. Postural variation in oropharyngeal dimensions in subjects with sleep disordered breathing: A cephalometric study. *Eur J Orthod* 2002; **24**: 263–276.
35. Palomo JM, Rao PS, Hans MG. Influence of CBCT exposure conditions on radiation dose. *Oral Surg, Oral Med, Oral Pathol, Oral Radiol Endod* 2008; **105**: 773–782.

Appendix 1

Search terminology used for other databases

Keyword	*Hits*
All EBM reviews	
1. airway.mp. [mp = ti, ot, ab, tx, kw, ct, sh, hw]	8976
2. Upper.mp. [mp = ti, ot, ab, tx, kw, ct, sh, hw]	12 465
3. Nasal.mp. [mp = ti, ot, ab, tx, kw, ct, sh, hw]	7409
4. pharyn*.mp. [mp = ti, ot, ab, tx, kw, ct, sh, hw]	2353
5. 1 or 2 or 3 or 4	28 333
6. segmentation.mp. [mp = ti, ot, ab, tx, kw, ct, sh, hw]	142
7. reconstruction.mp. [mp = ti, ot, ab, tx, kw, ct, sh, hw]	2171
8. algorithm.mp. [mp = ti, ot, ab, tx, kw, ct, sh, hw]	1850
9. three dimensional imaging.mp. [mp = ti, ot, ab, tx, kw, ct, sh, hw]	46
10. 6 or 7 or 8 or 9	4153
11. cone beam computed tomography.mp. [mp = ti, ot, ab, tx, kw, ct, sh, hw] or Computed tomography.mp. [mp = ti, ot, ab, tx, kw, ct, sh, hw]	243 627
12. 5 and 10 and 11	15
Scopus	
1. TITLE-ABS-KEY(airway OR upper OR nasal OR pharynx)	790 315
2. TITLE-ABS-KEY(segmentation OR reconstruction OR algorithm OR three dimensional imaging)	91 788
3. TITLE-ABS-KEY(cone beam computed tomography OR computed tomography)	3135
4. 1 and 2 and 3	69