Challenges in medical additive manufacturing

Maureen Antonia Johanna Maria van Eijnatten
The studies presented in this thesis were conducted at the Department of Oral and Maxillofacial Surgery / Oral Pathology and the 3D Innovation Lab of the VU University Medical Center, Amsterdam, The Netherlands.

The printing of this thesis has been financially supported by the department of Oral and Maxillofacial Surgery / Oral Pathology of the VU University Medical Center and the VU University Amsterdam.

Cover, layout & printing: Off Page, Amsterdam

ISBN: 978-94-6182-846-0

© Maureen van Eijnatten, 2017
Challenges in medical additive manufacturing

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof. dr. V. Subramaniam,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Geneeskunde
op vrijdag 8 december 2017 om 9.45 uur
in de aula van de universiteit,
De Boelelaan 1105

door

Maureen Antonia Johanna Maria van Eijnatten
geboren te Zwolle
promotor: prof.dr. T. Forouzanfar
copromotor: dr. J.E.H. Wolff
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>General introduction</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>A Novel Method of Orbital Floor Reconstruction Using Virtual Planning, 3-Dimensional Printing, and Autologous Bone</td>
<td>17</td>
</tr>
<tr>
<td>3</td>
<td>Influence of CT parameters on STL model accuracy</td>
<td>27</td>
</tr>
<tr>
<td>4</td>
<td>Impact of prone, supine and oblique patient positioning on CBCT image quality, contrast to noise ratio and figure of merit value in the maxillofacial region</td>
<td>43</td>
</tr>
<tr>
<td>5</td>
<td>CT image segmentation methods for bone used in medical additive manufacturing: A Literature Review</td>
<td>61</td>
</tr>
<tr>
<td>6</td>
<td>The impact of manual threshold selection in medical additive manufacturing</td>
<td>87</td>
</tr>
<tr>
<td>7</td>
<td>Accuracy of MDCT and CBCT in three-dimensional evaluation of the upper airway morphology</td>
<td>103</td>
</tr>
<tr>
<td>8</td>
<td>The accuracy of ultrashort echo time MRI sequences for medical additive manufacturing</td>
<td>117</td>
</tr>
<tr>
<td>9</td>
<td>General discussion</td>
<td>133</td>
</tr>
<tr>
<td>10</td>
<td>English summary &amp; Dutch summary</td>
<td>141</td>
</tr>
<tr>
<td>11</td>
<td>Acknowledgements</td>
<td>153</td>
</tr>
</tbody>
</table>
GENERAL INTRODUCTION
In the early 1980s, a revolutionary layer-by-layer manufacturing technique was developed to create industrial prototype parts from three-dimensional (3D) computational data. This novel approach was termed “rapid prototyping” (RP) and by the general public at large as “3D printing”. The first commercial photopolymer 3D printer was patented by Charles W. Hull in 1986 [1]. Since then, a plethora of different 3D printing technologies have been developed that include selective laser sintering (SLS), fused deposition modelling (FDM) and inkjet-based technologies [2]. Overtime, the term “rapid prototyping” has been replaced by “additive manufacturing” (AM), a term that currently encompasses a wide range of different technologies and applications [3]. AM technologies are currently being used in many fields, including the automotive, aerospace and consumer goods (customisation) industries and dental and medical science. According to the Wohlers Report 2017, the AM industry grew by 17.4% in 2016 with a global market share of $6.1 billion. Furthermore, total sales volume is expected to reach $25.6 billion in 2021 [4]. The major advantages of AM technologies over traditional subtractive manufacturing technologies are the possibilities to customise and create almost any complex geometry in a completely automated manner [5]. Another advantage is the speed with which a complex geometrical prototype or functional part can be manufactured when compared to conventional subtractive manufacturing technologies.

Over the past decades, advances in image processing and graphical processing power have extended the role of medical imaging far beyond traditional, two-dimensional visualisation. Novel image processing techniques can convert a stack of two-dimensional (2D) computed tomography (CT) or magnetic resonance (MR) images into a virtual 3D model. Such 3D models can be used as a blueprint to manufacture medical constructs using AM. Expiring patents and price cuts in medical 3D imaging and AM technologies have sparked an increase in the implementation of medical AM constructs in clinical settings [6].

Currently, there is a paradigm shift in medicine away from mass production towards iterations and from one-size-fits-all to personalised treatments. In this context, AM constructs have proven to be particularly valuable in the field of oral and maxillofacial surgery due to the plethora of complex bony geometries found in the skull area. Moreover, only minor dislocations in the facial area can have a profound impact on the function and aesthetic appearance of patients. In such cases, patient-specific AM anatomical models have been shown to improve communication between clinicians and patients [7]. In addition, AM models are being used for the preoperative planning and simulation of complex surgery in order to significantly improve the efficiency and outcome of the surgery [8]. Maxillofacial surgeons are also routinely using patient-specific AM drill guides [9] and saw guides [10] to translate a surgical plan into the operating room. To date, AM has been sporadically used to create patient-specific implants of the mandible [11], orbita [12]–[14], zygoma [15], cranium [16], tibia [17], and, although still in its infancy, scaffolds for tissue engineering [18]. Lastly, AM models of (rare) anatomical abnormalities can, in certain circumstances,
replace cadaveric materials in medical education and training [19], [20], thereby avoiding ethical controversies and reducing costs [21].

The current medical AM process comprises three different steps (Figure 1) [22]. The first step is 3D image acquisition using either CT or MRI technologies. In clinical settings, however, CT images are currently most often used due to their superior hard tissue contrast, spatial resolution and lower cost. The resulting imaging data is subsequently reconstructed in two-dimensional (2D) grey scale images that are saved in Digital Imaging and Communications in Medicine (DICOM) file formats. The second step is image processing, in which the tissue of interest (e.g., “bone”) is segmented and converted into a virtual 3D surface model by means of triangulation [23]. The resulting 3D surface model is saved as a Standard Tessellation Language (STL) model, which can then be used to design patient-specific constructs using dedicated computer-aided design (CAD) software. The third and final step in the medical AM process is the conversion of the designated STL model into a G-code that subsequently controls the path of the printing nozzle during AM.

One major challenge faced in the current medical AM workflow (Figure 1) is the lack of cohesive convergence between the medical imaging, computing and manufacturing technologies. This lack of convergence is mainly due to the fact that no single company provides all the necessary devices and services required for the medical AM process. The large variety of different CT/MRI scanners, software packages, 3D printing techniques and materials available on the market makes it difficult to select appropriate devices and software for medical AM. Moreover, such devices and software packages are often not certified for medical AM. As a result, concerns have subsequently been raised amongst policymakers and regulators responsible for product safety. Manufacturers of imaging devices, 3D printers and reverse engineering software companies should therefore agree on technical norms for the medical AM process. This is particularly important in the light of the new European Union medical device regulation (MDR) [24] that will require standardisation and CE certification for the whole medical AM process from 2020 onwards.

With regard to the first step in the AM workflow (Figure 1, step 1), there is general agreement amongst physicians and medical technicians that fan-beam CT technologies,
such as multi-detector row computed tomography (MDCT), offer the best images for medical AM [25]. However, it still remains unclear which CT technology offers the most accurate images for medical AM purposes. Moreover, to the best of my knowledge, no standardised scan protocols have been specifically developed for medical AM. [26] Currently, manufacturers are exploring new ways to improve CT image quality by using multiple photon energies (dual-energy CT) to enhance tissue characterisation. Furthermore, as opposed to conventional CT scanners, cone-beam computed tomography (CBCT) scanners are currently gaining popularity, especially in small dental and maxillofacial clinics. This surge in popularity is mainly due to the markedly lower purchasing costs, acceptable image quality and low radiation dose when compared with conventional CT scanners [27]. Recent studies have shown, however, that the partial CBCT gantry rotation produces a non-uniform dose distribution [28] that can in certain cases affect CBCT image quality. Therefore, CBCT scanning technology should not be routinely used to fabricate patient-specific implants. Interestingly, new research has shown that CBCT gantry and head positioning may have a positive impact on image quality in the maxillofacial region, which in turn may improve the overall accuracy of CBCT-derived AM constructs.

In the second step of the AM workflow (Figure 1, step 2), the major challenge is the conversion of grey scale voxel images (DICOM) into a triangulated 3D surface model (STL). This step is necessary since all CAD software packages approved by the United States Food and Drug Administration (FDA) require STL file formats for anatomical modelling and reverse engineering. The DICOM to STL conversion still requires multiple manual processing steps that are dependent on the spatial abilities and knowledge of anatomy of the technicians or physicians. The most important step in the DICOM to STL conversion process is image segmentation, which refers to the partitioning of images into regions of interest (ROIs) that correspond to specific anatomical structures. A wide range of different image segmentation methods have been developed over the past decades that include global thresholding, edge detection, region growing, statistical shape model (SSM)- and atlas-based methods, morphological snakes, active contours, random forests and artificial neural networks [29]. Of these methods, the most commonly used image segmentation method to date in medical AM is global thresholding. It should be noted that most software packages only offer a single, default threshold value for compact bone, soft tissue and cartilage. In addition, these default values are often not optimized for all types of MDCT, DECT and CBCT images and do not take into account the variations in grey values between the different scanners. Therefore, in most cases, manual threshold selection is necessary to acquire an optimal STL model. Threshold selection still remains, however, a subjective task that impedes the reproducibility and reliability of medical AM constructs. Moreover, global thresholding often causes data loss in the STL model. Such data loss can cause voids in the STL models that can subsequently lead to inaccuracies in 3D printed implants and lead to complications during surgery [30].

The aforementioned image segmentation process has been identified as one of the major causes of inaccuracies in medical AM constructs [31]. However, adequate
methods to assess the accuracy and reliability of medical AM constructs are still lacking. More specifically, novel methods are still being sought to validate the medical imaging and image processing steps in the medical AM workflow (Figure 1, step 1 and 2). One method of validating imaging devices and software packages is to use phantoms that provide ground truth measurements in clinical settings. Such phantoms are commonly composed of materials with realistic tissue (radio)densities. However, a major drawback with such phantoms is that they are manufactured in generic forms and sizes that do not resemble real patients. Thus, it is difficult to extrapolate the performance of an imaging system in phantoms to humans. Fortunately, 3D printing offers new possibilities to fabricate anthropomorphic phantoms that can be used to assess the accuracy of different CT scanners and software packages.

Currently, the number of patients treated with medical AM constructs is rapidly increasing. Accordingly, the number of CT and CBCT scans performed worldwide have subsequently increased. In this context, it should be noted that all X-ray based imaging modalities induce harmful ionizing radiation to the patient [32]. According to the ALARA (as low as reasonably achievable) principle, the cumulative lifetime radiation dose of patients should be minimised by reducing the overall number of examinations and by lowering the dose resulting from each individual examination [33]. Therefore, every CT scan that is used for medical AM needs foremost to be suitable for diagnostic purposes. Furthermore, novel ionizing radiation-free MRI modalities are being developed to image the musculoskeletal system [34]. MRI provides a noninvasive visualization of water microenvironments and other sources of protons in the human body, which provides excellent imaging possibilities for soft tissues. Cortical bone, however, has very short relaxation times [35], leading to fast decay of the MRI signal. As a result, cortical bone commonly appears as a signal void on MRI images obtained with routine clinical pulse sequences such as spin echo and gradient echo. With ultra-short echo time (UTE) sequences, an MR signal can be detected directly from bone [36], [37]. Therefore, UTE MRI sequences may be suitable to generate images for medical AM purposes. However, the accuracy of UTE-derived STL models still needs to be thoroughly evaluated. In addition, MRI remains an expensive imaging modality when compared to CT.

In conclusion, AM offers possibilities for fabricating different types of patient-specific constructs such as anatomical models, surgical guides and implants. However, despite recent advances and promising case studies involving medical AM, notable scientific, technological and regulatory challenges remain.

GENERAL AIM AND OUTLINE OF THE THESIS

The aim of this thesis is to provide an insight into the current scientific and technological challenges faced in medical additive manufacturing with respect to imaging and image processing. More specifically, the different parameters that influence the accuracy of patient-specific AM constructs will be identified.
REFERENCES


33. Implications of Commission Recommendations that Doses be kept as Low as Readily Achievable (1973) ICRP publication 22, Oxford, UK


A NOVEL METHOD OF ORBITAL FLOOR RECONSTRUCTION USING VIRTUAL PLANNING, 3-DIMENSIONAL PRINTING, AND AUTOLOGOUS BONE

Maarten Vehmeijer, Maureen van Eijnatten, Niels Liberton, Jan Wolff

doi: 10.1016/j.joms.2016.03.044
ABSTRACT

Fractures of the orbital floor are often a result of traffic accidents or interpersonal violence. To date, numerous materials and methods have been used to reconstruct the orbital floor. However, simple and cost-effective 3-dimensional (3D) printing technologies for the treatment of orbital floor fractures are still sought. This study describes a simple, precise, cost-effective method of treating orbital fractures using 3D printing technologies in combination with autologous bone. Enophthalmos and diplopia developed in a 64-year-old female patient with an orbital floor fracture. A virtual 3D model of the fracture site was generated from computed tomography images of the patient. The fracture was virtually closed using spline interpolation. Furthermore, a virtual individualized mould of the defect site was created, which was manufactured using an inkjet printer. The tangible mould was subsequently used during surgery to sculpture an individualized autologous orbital floor implant. Virtual reconstruction of the orbital floor and the resulting mould enhanced the overall accuracy and efficiency of the surgical procedure. The sculptured autologous orbital floor implant showed an excellent fit in vivo. The combination of virtual planning and 3D printing offers an accurate and cost-effective treatment method for orbital floor fractures.
INTRODUCTION

Blow-out fractures are often a result of motor vehicle accidents, interpersonal violence, or sports injuries. Such events can induce a blunt impact to the orbital area, which increases the intraorbital pressure and subsequently leads to a fracture of one or more orbital walls. This often results in a prolapse of the orbital content and, hence, displacement of the eye globe into the maxillary sinus cavity, which can lead to an enophthalmos (infra-position of the globe) and diplopia (double vision) of the eye.

To date, a wide range of different reconstruction materials have been used for the treatment of orbital floor fractures. Nevertheless, over the past few decades, autogenous bone grafts have been considered the gold standard [1]. However, a major drawback of autologous bone compared with prefabricated titanium meshes is that autologous bone is cumbersome to sculpture and can easily break if bent beyond its capacity. Furthermore, bone grafts can in some cases resorb unpredictably. However, orbital graft resorption can be minimized by proper fixation of the graft to the orbital floor [2].

Recent advancements in head imaging and 3-dimensional (3D) printing have opened a wide range of new treatment possibilities [3]. Such technologies allow the creation of patient-specific medical constructs. Currently, medical 3D printing is most commonly used to manufacture medical models, drill guides, saw guides and individualized titanium reconstruction plates [4], [5].

To date, biocompatible 3D printable materials are still being sought. Therefore, we suggest combining currently available 3D printing technologies with autologous bone grafts for the reconstruction of the orbital floor. The aim of this study was to develop a simple and cost-effective method to produce individualized orbital implants using computed tomography (CT) imaging and virtual planning, 3D printed moulds, and autologous bone grafts.

MATERIALS AND METHODS

A swollen and painful left eye developed in a 64-year-old female patient after a bicycle accident. Clinical examination showed a left-sided periorbital hematoma, subconjunctival ecchymosis, mild enophthalmus, and diplopia. Furthermore, the patient’s left eye movement was restricted in the vertical plane. CT images showed a blowout fracture of the left orbit (Figure 1).

The CT images were acquired using a GE Discovery CT750 HD Multi-Slice CT scanner (GE Healthcare, Waukesha, WI). The following imaging parameters were used: tube voltage, 120 kV; tube current, 300 mA; and reconstruction kernel, soft. The slices were 0.6 mm thick, with 0.312-mm gaps between them. The resulting Digital Imaging and Communications in Medicine (DICOM) files were imported into OsiriX MD imaging software (Pixmeo, Geneva, Switzerland), and a virtual 3D model of the patient’s skull was created using a threshold value of 120 Hounsfield units.
The virtual 3D model was used to manufacture a 3D printed mould of the virtually reconstructed orbital floor. The virtual reconstruction was performed as follows: The boundary of the orbital floor defect was delineated in the virtual 3D model using 3-matic Medical 9.0 software (Materialise, Leuven, Belgium). In a second step, 7 spline interpolation curves were virtually drawn through the defect site to reconstruct the missing orbital floor bone (Figure 2). In a third step, a rectangular box (60 x 60 x 30 mm) was created around the reconstructed virtual 3D model of the orbital floor. The rectangular box was necessary to 3D print a stable tangible mould of the orbital floor (Figure 3), allowing the surgeon to apply force whilst sculpturing the autologous bone graft. The aforementioned virtual reconstruction of the orbital floor took approximately 45 minutes. After successful virtual reconstruction, one tangible mould was printed using a ZCorp inkjet printer (3D Systems, Rock Hill, SC). The total printing time required to fabricate the tangible mould was 84 minutes.

Figure 1. Computed tomography image (coronal view) showing a blowout fracture of the left orbit with protruding soft tissue into the maxillary sinus (arrow).
During surgery, bone was harvested from the anterior iliac crest. The resulting bone graft was thinned down to 1 mm in the outer area and to 2 to 3 mm in the central area. The previously printed mould of the orbital floor was wrapped in a sterile plastic bag and used as a template to mould and sculpture the graft into an individualized autologous orbital floor implant (Figure 3).

The individualized implant was meticulously placed into the orbital floor and fixated with two 3-hole miniplates (1.0 mm) and four 3-mm centre drive screws (KLS Martin, Tuttlingen, Germany) to the inferior orbital rim. The surgical procedure took 40 minutes. Postoperative CT images revealed a harmonious reconstruction of the orbital floor (Figure 4).

**Figure 2.** Virtual 3-dimensional model of the patient’s skull (grey), including the virtual mould of the orbital floor (A), the virtually reconstructed bone defect (B), and the spline interpolation curves (C).

**Figure 3.** Autologous bone graft (ie, implant) sculptured and moulded on a printed mould (A) with a reconstructed orbital floor (B).
RESULTS AND DISCUSSION

A wide variety of materials have been described for the treatment of orbital floor fractures, ranging from polymers to metals [6]. In the past decade, individualized, 3D printed titanium orbital implants have been used in maxillofacial surgery. However, such implants are difficult to fabricate, are expensive, and sometimes lack dimensional precision [7]. Furthermore, they are difficult to adapt or sculpture during surgery because of their mechanical properties. Medical titanium alloys typically have an elastic modulus (E-modulus or Young’s modulus) of 120 GPa [8], whilst the E-modulus of orbital bone ranges between 1.26 and 4.55 GPa [9]. Such differences can cause irritation and resorption of the bone underneath the titanium implant [10]. The major advantage of the moulded autologous bone implant used in this study is its biocompatibility, which subsequently causes less irritation to surrounding tissues.

Historically, mirroring of the contralateral healthy orbit has been commonly used in virtual planning and subsequent printing of orbital implants. However, this technique has certain drawbacks. The major drawback is that the human face is often asymmetric and therefore geometric variability between the two orbits is possible. Therefore,
we recommend using spline interpolation that offers the possibility of closing orbital defects using the surrounding healthy bony structures as a reference. However, creating an accurate spline-based mould requires technical and computational knowledge that may be challenging for most clinicians. This problem could be solved using large image databases combined with machine learning and pattern recognition algorithms that would automatically reconstruct complex anatomic geometries in trauma patients.

The orbital floor mould used in this study was printed in 84 minutes using a high-resolution 3D printer. During fabrication, the printer deposits a 0.1-mm-thick layer of powder over which the printing head moves and deposits a binder that solidifies the powder. The thin-layer increments subsequently offer very precise 3D printed constructs.

However, designing the very precise virtual model required for the fabrication of a mould is slightly more challenging since it is dictated by the correctness of the conversion of CT images into virtual 3D models (Figure 2). The accuracy of a virtual 3D model derived from CT images currently ranges between 0.4 and 0.7 mm [11]. Nevertheless, the autologous orbital floor implant used in this case report demonstrated an excellent fit in-situ and subsequently reduced the operation time by approximately 40% (Figure 4).

The total manufacturing costs of the orbital floor mould were low (€85) compared with other treatment options. To date, 3D printed titanium orbital implants have cost between €3,000 and €5,000 euros. The costs of prefabricated titanium orbital plates or meshes range between €250 euros (KLS Martin) and €600 and €800 (Synthes, West Chester, PA).

The treatment of bony defects with autologous bone still remains a challenge since the amount of donor bone available is anatomically limited. A major drawback of autogenous bone is, of course, the need for harvesting, which is accompanied by donor-site morbidity. In addition, the sculpturing of autologous bone to fit the recipient site can be time-consuming and manually cumbersome and subsequently result in inconsistent cosmetic outcomes. However, this problem can be minimized using the treatment method described in this study.

In conclusion, virtual preoperative planning combined with 3D printing offers a geometrically predictable and low-cost method of creating autologous individualized orbital floor implants.
REFERENCES


INFLUENCE OF CT PARAMETERS ON STL MODEL ACCURACY

Maureen van Eijnatten, Ferco Berger, Pim de Graaf, Juha Koivisto, Tymour Forouzanfar, Jan Wolff

ABSTRACT

Purpose
Additive manufactured (AM) skull models are increasingly used to plan complex surgical cases and design custom implants. The accuracy of such constructs depends on the standard tessellation language (STL) model, which is commonly obtained from computed tomography (CT) data. The aims of this study were to assess the image quality and the accuracy of STL models acquired using different CT scanners and acquisition parameters.

Design/methodology/approach
Images of three dry human skulls were acquired using two multi-detector row computed tomography (MDCT) scanners, a dual energy computed tomography (DECT) scanner and one cone beam computed tomography (CBCT) scanner. Different scanning protocols were used on each scanner. All images were ranked according to their image quality and converted into STL models. The STL models were compared to gold standard models.

Findings
Image quality differed between the MDCT, DECT and CBCT scanners. Images acquired using low-dose MDCT protocols were preferred over images acquired using routine protocols. All CT-based STL models demonstrated non-uniform geometrical deviations of up to +0.9 mm. The largest deviations were observed in CBCT-derived STL models.

Practical implications
While patient-specific AM constructs can be fabricated with great accuracy using AM technologies, their design is more challenging because it is dictated by the correctness of the STL model. Inaccurate STL models can lead to ill-fitting implants that can cause complications after surgery.

Originality/value
This paper suggests that CT imaging technologies and their acquisition parameters affect the accuracy of medical AM constructs.
INTRODUCTION

Additive manufacturing (AM), often referred to as three-dimensional (3D) printing, is becoming increasingly popular in the field of medicine. Recent developments in high-resolution imaging techniques and more affordable 3D printers have led to an increased use of different AM technologies in clinical settings [1]. AM medical models are becoming indispensable in medical education and communication [2], [3], treatment planning, and surgical guidance [4], and have been shown to improve surgical outcomes and reduce operating times [5]. Furthermore, AM drill and saw guides in combination with patient-specific constructs such as implants or reconstruction plates are being more readily used in maxillofacial surgery [6].

Medical AM differs from industrial AM in that it requires patient-specific 3D medical images. To date, such images are commonly acquired using computed tomography (CT) technology. The resulting CT datasets are archived as Digital Imaging and Communications in Medicine (DICOM) files that are subsequently converted into standard tessellation language (STL) models. Such models are necessary to design medical AM constructs using commercially available computer-aided design (CAD) software packages.

A wide range of different CT technologies are currently available on the market for diagnostic purposes. To date, multi-detector row computed tomography (MDCT) scanners are most commonly used in hospital environments. These scanners have multiple rows of detectors, hence short scanning times, and allow high-resolution imaging of bone [7]. In addition, dual energy computed tomography (DECT) scanners have recently been developed and offer enhanced tissue characterization by using two X-ray beams with different photon energy spectra [8]. However, the scanning protocols used with these advanced imaging technologies exhibit a wide variability in terms of image acquisition and exposure parameters. These parameters include the scan mode (e.g. axial or helical), tube potential (kV), exposure (mAs), rotation time, X-ray beam filter type and size, beam collimation, detector configuration, helical pitch (or table feed per gantry rotation), and field of view (FOV).

Cone beam computed tomography (CBCT) scanners are commonly used in private clinics due to their small size and low costs. CBCT technology uses a cone shape X-ray beam and a single flat panel detector to obtain 3D information of the head of a patient by data acquisition through a single 200°-360° gantry rotation [9]. Cone beam technology induces lower radiation doses compared with MDCT [10], [11] and has recently become more affordable. This has led to an increase in popularity amongst maxillofacial radiologists and surgeons [12]. Comparable to MDCT, modern CBCT scanners offer a wide range of image acquisition parameters that include scan time, tube potential (kV), exposure (mAs), rotation time, gantry rotation, FOV, and various metal artifact reduction algorithms.

The aforementioned CT technologies and resulting datasets are currently used for diagnostic purposes in clinical settings. Such diagnostic datasets are also pro tempore used to design medical constructs. However, it must be noted that imaging protocols
required for diagnostic purposes may differ markedly from those required by technicians to design patient-specific AM constructs [13]. Nevertheless, all clinicians are bound to the ALARA (as low as reasonably achievable) radiation safety principle. Therefore, all CT scans required for patient-specific constructs need to be foremost diagnostable. The aims of this study were to assess the image quality and the accuracy of STL models acquired using different MDCT and CBCT scanners and acquisition parameters.

**MATERIALS AND METHODS**

Three formalin-fixed human cadaver skulls of succumbed patients with intact bony structures were obtained from the Department of Anatomy of the VU University Medical Center, Amsterdam. All soft tissues were meticulously removed. The skulls were subsequently scanned using a GOM ATOS™ optical 3D scanner (GOM, Braunschweig, Germany). This device uses non-contact, blue-light technology to record a series of images of an object from different angles, and combines these images into a triangular mesh representing the surface of the object. The GOM ATOS™ optical 3D scanner offers a point accuracy of 0.05 mm and was therefore used to obtain a “gold standard” STL model of each skull (Figure 1).

![Diagram](image)

**Figure 1.** Outline of the study.
CT image acquisition

All three dry human skulls were subsequently scanned using two different MDCT scanners and a CBCT scanner (Figure 1): a Siemens Somatom Sensation 64-slice MDCT scanner (Siemens Medical Solutions, Erlangen, Germany), a GE Discovery CT750 HD 64-slice MDCT scanner (GE Healthcare, Little Chalfont, Buckinghamshire, UK), and a Vatech PaX Zenith 3D CBCT scanner (Vatech, Fort Lee, USA). In addition, the GE scanner offered a DECT image acquisition possibility using rapid switching X-ray beams with 140 kVp and 80 kVp voltage. All acquisition parameters are presented in Table 1.

Different scanning protocols were used on the aforementioned MDCT and CBCT scanners. Initially, routine scanning protocols recommended by the device manufacturers were used. In a second step, minimum and maximum tube potential (kV) and tube current (mA) were used (Table 2) in order to assess the influence of these parameters on image quality and accuracy of the STL models. Furthermore, the radiation output was calculated using the dose-length product (DLP) for all MDCT protocols, and the dose-area product (DAP) for all CBCT protocols.

All MDCT and DECT images were reconstructed using sharp reconstruction kernels (Siemens: H70h, GE: BONEPLUS for routine mode and DETAIL for dual energy mode). These kernels were selected after assessing all the different kernels available on the Siemens and GE scanners.

Image quality analysis

Following image acquisition and reconstruction, two radiologists with specific expertise in maxillofacial bone imaging visually assessed and ranked all 57 DICOM datasets hence

<table>
<thead>
<tr>
<th>Exposure parameters</th>
<th>Siemens Somatom Sensation 64-slice MDCT</th>
<th>GE Discovery CT750 HD 64-slice MDCT/DECT</th>
<th>Vatech PaX Zenith CBCT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tube voltage (kVp)</td>
<td>80-140</td>
<td>80-140</td>
<td>50-120</td>
</tr>
<tr>
<td>Tube current (mA)</td>
<td>150-650</td>
<td>150-650</td>
<td>5-10</td>
</tr>
<tr>
<td>Rotation time (s)</td>
<td>1</td>
<td>1</td>
<td>15-24</td>
</tr>
<tr>
<td>Q (mAs)</td>
<td>150-650</td>
<td>150-650</td>
<td>75-240</td>
</tr>
<tr>
<td>Slice thickness (mm)</td>
<td>0.6</td>
<td>0.625</td>
<td>n.a.</td>
</tr>
<tr>
<td>Slice increment (mm)</td>
<td>0.391</td>
<td>0.312</td>
<td>n.a.</td>
</tr>
<tr>
<td>Pitch (mm)</td>
<td>0.55-1</td>
<td>0.5-1</td>
<td>n.a.</td>
</tr>
<tr>
<td>CTDIvol (mGy)</td>
<td>5.94-113</td>
<td>17.7-227</td>
<td>n.a.</td>
</tr>
<tr>
<td>Voxel (mm)</td>
<td>0.391 x 0.391 x 0.6</td>
<td>0.312 x 0.312 x 0.625</td>
<td>0.2 x 0.2 x 0.2</td>
</tr>
<tr>
<td>Voxel number</td>
<td>512 x 512</td>
<td>512 x 512 x 512</td>
<td>1056 x 1056 x 1056</td>
</tr>
<tr>
<td>Scan angle</td>
<td>360°</td>
<td>360°</td>
<td>360°</td>
</tr>
<tr>
<td>Pulsed</td>
<td>No</td>
<td>Only for dual energy</td>
<td>Yes</td>
</tr>
<tr>
<td>Scan FOV (cm)</td>
<td>20</td>
<td>20</td>
<td>Free (~24x19)</td>
</tr>
</tbody>
</table>
The radiologists were blinded from the acquisition parameters and from the results of one another. A constant Window/Level setting was used to evaluate all images equally (W6000; L100). The following anatomical structures were visually assessed: 1) Laminae papyracea; 2) Infra-orbital canal; 3) Sphenoid sinus septum; and 4) Orbital apex (Figure 2). Each anatomical structure was given scores from 1 to 4 in ascending order; 1 = anatomic structure not identifiable due to poor image quality; 2 = structure identified but no details assessable and insufficient image quality; 3 = anatomic structure fully assessable in all parts and acceptable image quality; 4 = very good delineation of structure and excellent image quality.

Because of the small variations in the image quality of the different scanning protocols (Table 3), a second more specific ranking system was necessary. In this ranking system, all images were given scores from best to worst according to relative image quality (Table 4). All image assessments and hence rankings were repeated after a 6-month interval to minimize the radiologists’ learning effect.

### Table 2. MDCT, DECT and CBCT scanning protocols: tube potentials (kV) and exposures (mAs).

<table>
<thead>
<tr>
<th>Scanning protocol</th>
<th>MDCT (Siemens) kV / mAs</th>
<th>MDCT (GE) kV / mAs</th>
<th>DECT (GE) kV / mAs</th>
<th>CBCT (Vatech) kV / mAs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Routine</td>
<td>120 / 380</td>
<td>120 / 220</td>
<td>80 &amp; 140 / 375</td>
<td>115 / 144</td>
</tr>
<tr>
<td>2. Routine kV / max. mA</td>
<td>120 / 650</td>
<td>120 / 650</td>
<td>80 &amp; 140 / 630</td>
<td>115 / 240</td>
</tr>
<tr>
<td>3. Min. kV / max. mA</td>
<td>80 / 650</td>
<td>80 / 650</td>
<td>n.a.</td>
<td></td>
</tr>
<tr>
<td>4. Max. kV / max. mA</td>
<td>140 / 550</td>
<td>140 / 510</td>
<td>120 / 240</td>
<td></td>
</tr>
<tr>
<td>5. Min. kV / min. mA</td>
<td>80 / 150</td>
<td>80 / 150</td>
<td>50 / 75</td>
<td></td>
</tr>
<tr>
<td>6. Max. kV / Routine mA</td>
<td></td>
<td></td>
<td>120 / 144</td>
<td></td>
</tr>
<tr>
<td>7. Metal artifact reduction</td>
<td></td>
<td></td>
<td>115 / 144</td>
<td></td>
</tr>
<tr>
<td>8. Mediate kV / routine mA</td>
<td></td>
<td></td>
<td>80 / 144</td>
<td></td>
</tr>
</tbody>
</table>

Statistical analysis of the image quality

The mean values and standard deviations (SD) of both ranking systems (Table 3 and 4) were calculated using SPSS Statistics software (version 23.0 for Windows; SPSS Inc, Chicago, USA). Furthermore, all rankings were compared using a nonparametric Mann-Whitney U test with a 2-tailed p value. Statistical significance was established at $p \leq 0.05$.

Image processing

After image quality analysis, All MDCT, DECT, and CBCT DICOM datasets were subsequently converted into STL models using OsiriX MD software (Pixmeo, Geneva, Switzerland). Automatic thresholding was used with a uniform intensity value of -500 Hounsfield Units (HU).
Standard tessellation language deviation analysis

All STL files were imported into GOM Inspect software (GOM GmbH, Braunschweig, Germany), and were cleared of outliers and aligned with the gold standard STL model using a “Local Best-Fit” function. Geometrical differences between the STL models and the gold standard were calculated using the “Surface Comparison on Actual” function and visualized using colour maps. A detailed assessment of the orbital area was performed using the “Select on Surface” function. All distance values were exported to Excel software (Microsoft Office 2013, Redmond, WA, United States) and the mean and SD were calculated. Finally, histograms and kurtosis were calculated using an excel data analysis software plugin.

RESULTS

The image quality of the following four anatomical structures was visually assessed: laminae papyracea; infra-orbital canal; sphenoid sinus septum and orbital apex (Figure 2). Furthermore, the DLP of each MDCT examination and the DAP of each CBCT examination was calculated. The results are summarized in Table 3. Scanner-specific rankings of the different scanning protocols are presented in Table 4.

All visually examined images were subsequently converted into STL models. The resulting STL models were compared to a gold standard skull STL model acquired using an optical scanner. When compared with the gold standard STL models, all STL models obtained from MDCT, DECT, and CBCT datasets demonstrated large non-uniform deviations in the orbital wall, maxillary sinus, arcus zygomaticus, and nasal cavity.

Figure 2. Evaluated anatomical structures, 1) Laminae papyracea; 2) Infra-orbital canal; 3) Sphenoid sinus septum; and 4) Orbital apex. The images were acquired using a Siemens MDCT scanner, 80 kV, 150 mA, window/level: 6000/100.
Table 3. Image quality scores (scale 1 – 4) and dose-length product (DLP) / dose-area product (DAP) of all MDCT, DECT, and CBCT datasets a.

<table>
<thead>
<tr>
<th>Scanning protocol</th>
<th>MDCT (Siemens)</th>
<th>MDCT (GE)</th>
<th>DECT (GE)</th>
<th>CBCT (Vatech)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean ± SD</td>
<td>DLP (mGy-cm)</td>
<td>Mean ± SD</td>
<td>DLP (mGy-cm)</td>
</tr>
<tr>
<td>1. Routine</td>
<td>3 ± 0.3</td>
<td>1132</td>
<td>3 ± 0.5</td>
<td>1632</td>
</tr>
<tr>
<td>2. Routine kV / max. mA</td>
<td>3 ± 0.3</td>
<td>1865</td>
<td>3 ± 0.6</td>
<td>4676</td>
</tr>
<tr>
<td>3. Min. kV / max. mA</td>
<td>2 ± 0.6</td>
<td>544</td>
<td>3 ± 0.8</td>
<td>1676</td>
</tr>
<tr>
<td>4. Max. kV / max. mA</td>
<td>3 ± 0.0</td>
<td>2390</td>
<td>3 ± 0.8</td>
<td>5111</td>
</tr>
<tr>
<td>5. Min. kV / min. mA</td>
<td>3 ± 0.3</td>
<td>131</td>
<td>3 ± 0.7</td>
<td>398</td>
</tr>
<tr>
<td>6. Max. kV / Routine mA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Metal artifact reduction</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Mediate kV / routine mA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a p values indicate a comparison with the routine protocol (* = p < .05). All p values for MDCT and DECT scanning protocols are nonsignificant.

Table 4. Scanner-specific rankings of all MDCT, DECT, and CBCT datasets a.

<table>
<thead>
<tr>
<th>Scanning protocol</th>
<th>MDCT (Siemens)</th>
<th>MDCT (GE)</th>
<th>DECT (GE)</th>
<th>CBCT (Vatech)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean ± SD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Routine</td>
<td>3.2 ± 1.3</td>
<td>2.9 ± 1.1</td>
<td>1.9 ± 0.2</td>
<td>6.2 ± 0.9</td>
</tr>
<tr>
<td>2. Routine kV / max. mA</td>
<td>3.1 ± 1.1</td>
<td>.284</td>
<td>2.5 ± 1.3</td>
<td>1.1 ± 0.2</td>
</tr>
<tr>
<td>3. Min. kV / max. mA</td>
<td>1.6 ± 1.0</td>
<td>.000*</td>
<td>3.2 ± 1.2</td>
<td>.154</td>
</tr>
<tr>
<td>4. Max. kV / max. mA</td>
<td>3.5 ± 1.2</td>
<td>.179</td>
<td>1.9 ± 1.3</td>
<td>.000*</td>
</tr>
<tr>
<td>5. Min. kV / min. mA</td>
<td>3.7 ± 1.5</td>
<td>.074</td>
<td>4.3 ± 1.1</td>
<td>.000*</td>
</tr>
<tr>
<td>6. Max. kV / Routine mA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Metal artifact reduction</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Mediate kV / routine mA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a p values indicate a comparison with the routine protocol (* = p < .05).
The calculated deviations of one skull (no. 2) are presented in Figure 3. The other two skulls (no. 1, 3) presented similar deviations.

The STL models of skull 1, 2, and 3 acquired using the Siemens MDCT scanner showed deviations of 0.58 mm, 0.49 mm, and 0.77 mm and SDs of 0.91, 0.81, 0.99, respectively (Figure 3A). The GE MDCT STL deviations ranged from 0.56 mm, 0.45 mm, and 0.74 mm with SDs of 0.88, 0.76, and 0.96, respectively (Figure 3B). DECT STL models resulted in deviations of 0.44 mm, 0.45 mm, and 0.62 mm with SDs of 0.73, 0.71, and 0.84, respectively (Figure 3C). Furthermore, DECT STL models showed less detail in the orbital and maxillary sinus area when compared with the MDCT models. Finally, the CBCT scanner presented deviations of 0.63 mm, 0.63 mm, and 0.80 mm with SDs of 0.88, 0.88, and 0.95, respectively (Figure 3D).

The deviations between the gold standard STL models and STL models acquired from all three skulls using four different CT scanners generally ranged between -0.9 mm and

---

Figure 3. Geometric deviations mapped on the surface of the STL models acquired using a routine scanning protocol on four different CT scanners. (a) MDCT (Siemens) - 120 kV/380 mAs; (b) MDCT (GE) - 120 kV/220 mAs; (c) DECT (GE) - 80&140 kV/375 mAs; (d) CBCT (Vatech) - 115 kV/144 mAs
+0.8 mm and are summarized in a histogram (Figure 4). The MDCT STL models showed the smallest deviations, but the models were on average larger than the gold standard STL model. The kurtosis – the steepness of a curve compared to a normal curve – of the deviations in the MDCT STL models was 1.7 (Siemens) and 2.7 (GE), 0.08 in the DECT STL models and -0.9 in the CBCT STL models.

Detailed assessments of the orbital area in two Siemens MDCT STL models of skull 2 are presented in Figure 5. When compared with the routine scanning protocol (A), the min kV/min mA protocol (B) resulted in better anatomical details. The mean deviation in the orbits was 0.25 ± 0.41 mm (routine protocol) and 0.29 ± 0.44 mm (min mA/min kV protocol). In all MDCT STL models, the orbits were larger than the gold standard in 80 per cent of the surface area.

**Figure 4.** Frequency of the geometric deviations (in mm) of the STL models of all three skulls when compared with the gold standard.

**Figure 5.** An example of the deviations in the STL models attained using standard and min kV/min mA scanning protocols on the Siemens MDCT scanner.
DISCUSSION

In the present study, the image quality and the accuracy of skull-derived STL models acquired using different MDCT and CBCT scanners and acquisition parameters were assessed.

The image quality differed markedly between the MDCT, DECT and CBCT scanners (Table 3). The MDCT scanners offered the best image quality. Furthermore, the images acquired using min kV/min mA and hence low-dose MDCT scanning protocols were preferred over the images acquired using routine protocols (Table 4). Moreover, compared to the routine protocols, the low-dose protocols showed a significant reduction in the DLP of 88 and 76 per cent on the Siemens and GE scanner, respectively (Table 3). These findings have large clinical implications since the use of low-dose scanning protocols can markedly reduce the effective dose in patients who require patient-specific AM constructs without jeopardizing the overall diagnostic image quality.

After image quality assessment, the acquired DICOM datasets were converted into STL models and geometrically compared to the corresponding gold standard skull models. In this context, the MDCT scanners offered the most accurate STL models (Figure 3) with absolute mean deviations ranging between 0.45 mm and 0.77 mm when compared with the optical scan “gold standard” STL model. These results are in good agreement with the findings of Thorhauer et al. (2010), who found mean deviations of 0.512 mm between MDCT STL models and optical scan models of the knee [14]. Other studies reported mean deviations of 0.13 mm [15] and 0.45 mm [16] on long cadaver bones and 0.137 mm on dry human mandibles [17]. However, these studies only included large bony structures.

The STL models acquired using DECT scanning protocols demonstrated greater loss of bony details when compared to the MDCT protocols, especially in the thin bony structures of the orbit (Figure 3). Deviations ranging between 0.44 mm and 0.62 mm were found in the DECT STL models. These findings could be due to the rapid switching between photon energy spectra required for dual energy imaging on the GE scanner. Another explanation could be the lack of a bone reconstruction kernel for dual energy image reconstruction.

The largest geometrical deviations were observed in the CBCT STL models (Figure 3), with mean deviations ranging between 0.63 mm and 0.80 mm. Liang et al. (2010) reported mean deviations of between 0.165 and 0.386 mm for CBCT STL models [17], though deviations of up to 2-5 mm have also been reported [18]. The overall low accuracy of the CBCT STL models found in this study could be caused by the relatively high amount of noise, artefacts [19], and inhomogeneities in the CBCT X-ray beam [20].

The geometrical deviations that were observed in the orbital area of all STL models could have been caused by the thin (less than 1 mm) and unique anatomy of the orbits. Currently, most CT scanners have an in-plane spatial resolution of approximately 0.3 mm and a slice thickness of 0.6 mm and can therefore not detect such bony structures (smaller than 0.6 mm). Moreover, the complex, scaffold-like structure of the orbital area also comprises thick supportive bony structures. These structures attenuate the lower part
of the radiation energy spectrum and lead to beam hardening and photon starvation [21] (Figure 6). Hardened energy photons are less prone to attenuate in thin bony structures, causing a loss of contrast in these structures [22]. Interestingly, when compared with routine MDCT scanning protocols, min kW/min mA and hence low-dose MDCT protocols resulted in similar geometric deviations and an increase of detail in the orbital area (Figure 5). These results are in good agreement with a previous study by Oka et al. (2009), who found minor geometrical differences between bone models constructed using normal and low-dose CT parameters.

Clinical implications
While patient-specific medical constructs can be fabricated with great accuracy (< 0.1 mm) using AM technologies, their design is more challenging since it is dictated by the correctness of the STL model. A recent study by Stoor et al. (2014) reported that 24 per cent of their additively manufactured orbital implants did not fit due to a poor representation of the orbital walls in the corresponding CT-based STL models. The consequence of ill-fitting implants is that increased movements can jeopardize wound healing and tissue repair and subsequently lead to complications after surgery [24]. Furthermore, geometrically inaccurate implants can cause nerve impingements in complex anatomical areas such as the orbit, with devastating outcomes. This opens new questions concerning the liability of AM medical constructs. Therefore, technicians working in the medical field need to understand the inaccuracies that can be introduced during the CT imaging and STL conversion process required for the fabrication of medical AM constructs.

Limitations of this study
The major limitation of this study was that skulls without soft tissues were used. The photon energy spectrum of an X-ray beam changes as it passes through water [25], which results in differences in absorption and scattering during clinical CT scanning. Furthermore, the intensity values of soft tissues are closer to the values of bone, hence the presence of soft tissues will make bone segmentation more difficult. Therefore,

![Image](https://via.placeholder.com/150)

**Figure 6.** Beam hardening and photon starvation in the orbital area can cause an enlargement of the STL model when compared to the actual bone.
the reported accuracies found in this study are not simply generalizable to clinical conditions. Nevertheless, dry skulls combined with optical scanning technology offered an highly accurate and detailed 3D geometric “gold standard” to evaluate the accuracy of CT-based STL models. The acquisition of a gold standard is not straightforward for complex anatomical structures, and even impossible to acquire in an vivo study.

CONCLUSION

CT imaging technologies and their acquisition parameters have an effect on the accuracy of medical AM constructs. MDCT, DECT and CBCT scanners differ in image quality and STL model accuracy in the maxillofacial area. Overall, MDCT scanners offered the best image quality and STL models. The use of low-kV and low-mA instead of routine MDCT protocols can markedly reduce the effective dose in patients who require patient-specific AM constructs.
REFERENCES


IMPACT OF PRONE, SUPINE AND OBLIQUE PATIENT POSITIONING ON CBCT IMAGE QUALITY, CONTRAST TO NOISE RATIO AND FIGURE OF MERIT VALUE IN THE MAXILLOFACIAL REGION

Juha Koivisto, Maureen van Eijnatten, Jorma Järnstedt, Kisri Holli-Helenius, Prasun Dastidar, Jan Wolff

doi: 10.1259/dmfr.20160418
ABSTRACT

Objective
To assess the impact of supine, prone and oblique patient imaging positions on the image quality, contrast-to-noise ratio and figure of merit value in the maxillofacial region using a cone beam computed tomography scanner. Furthermore, the CBCT supine images were compared to supine MSCT images.

Methods
One fresh frozen cadaver head was scanned in prone, supine and oblique imaging positions using a mobile CBCT scanner. MSCT images of the head were acquired in a supine position. Two radiologists graded the CBCT and MSCT images at ten different anatomical sites according to their image quality using a six-point scale. CNR and FOM values were calculated at two different anatomical sites on the CBCT and MSCT images.

Results
The best image quality was achieved in prone imaging position for sinus, mandible and maxilla, followed by supine and oblique imaging positions. 12-mA prone images presented high delineation scores for all anatomical landmarks, except for the ear region (carotid canal), which presented adequate to poor delineation scores for all studied head positions and exposure parameters. The MSCT scanner offered similar image qualities to the 7.5-mA supine images acquired using the mobile CBCT scanner. The prone imaging position offered the best CNR and FOM values on the mobile CBCT scanner.

Conclusions
Head positioning has an impact on CBCT image quality. The best CBCT image quality can be achieved using the prone and supine imaging positions. The oblique imaging position offers inadequate image quality except in the sinus region.
INTRODUCTION

Mobile cone beam computed tomography (CBCT) scanners were first introduced in the field of maxillofacial surgery in 2005 [1]. Since then, mobile CBCT scanners [2]–[8] have become increasingly popular as they offer comparable images to those acquired on traditional multislice computed tomography (MSCT) scanners [9]. The major advantages of CBCT scanners are their lower cost, size, weight and effective dose [10]–[12]. The lower effective dose is the result of the default exposure parameters and the partial gantry rotation used in CBCT scanners when compared to MSCT scanners [13], [14]. It has, however, been reported that the partial CBCT gantry rotation produces a non-uniform dose distribution [15] that can in some cases affect image quality. Nevertheless, the partial CBCT gantry rotation offers acceptable diagnostic image quality at a lower effective dose than MSCT scanners [16].

A recent study by Koivisto et al. [17] reported the positive effect of head positioning on the effective dose in CBCT scanners. Furthermore, a study by Yan et al. [18] reported a close relationship between image quality and effective dose when using low dose CBCT protocols. Taking the results of the previous studies into account, it can be hypothesized that CBCT images could be optimised by choosing the correct patient imaging position.

Currently, MSCT and CBCT image quality assessments are based on the visual grading of predefined anatomical landmarks [19]–[21], the quantitative analysis of physical properties such as modulation transfer function (MTF) [22] and contrast-to-noise ratio (CNR) [23]. The image contrast is affected by the x-ray mean photon energy that subsequently has an effect on the image quality [24]. Moreover, the relationship between the image quality and patient dose is commonly quantified using the traditional figure of merit (FOM) value that is defined as the \((\text{CNR})^2\) divided by the incident patient exposure [24]–[27]. The figure of merit value can provide useful information on the benefits and drawbacks of different imaging positions in CBCT and MSCT scanners. It also provides a means of comparing the ratio between the image quality and the resulting effective dose. Furthermore the FOM offers the possibility of calculating a numerical value that can be used to compare two inherently different imaging modalities. Moreover, the FOM value can be used to select an appropriate protocol depending on the diagnostic task. The combination of CNR and FOM calculations with image quality assessments offer an adequate method for comparing and grading MSCT and CBCT images.

The objectives of this study were to assess the impact of three different patient imaging positions (supine, prone, oblique) on CBCT image quality, contrast-to-noise ratio and the figure of merit value in the maxillofacial region. The second objective was to compare the CBCT supine images with MSCT images acquired in the same position.

MATERIALS AND METHODS

This study was performed according to the Ethical Principles for Medical Research Involving Human Subjects as defined by the World Medical Association (Helsinki Declaration) [28].
One fresh frozen human cadaver head was provided by the Department of Anatomy, Tampere University Hospital for image quality assessment.

**Scanners**

Cadaver images were obtained using a Planmed Verity mobile CBCT scanner (Planmed Oy, Helsinki, Finland) and a Philips Brilliance-64 (64 slice) MSCT scanner (Philips Medical Systems, Eindhoven, Netherlands). The mobile CBCT scanner uses a partial 210° gantry scan and 3.0 mm Al and 0.5 mm Cu filters. The Philips Brilliance 64 MSCT scanner uses a 360° rotation scan and combined filtration consisting of 2.5 mm aluminum and 1.2 mm titanium.

The mobile scanner was initially developed for extremity imaging and has recently been approved by the American Food and Drug Administration (FDA) for maxillofacial imaging. The scanner is equipped with a head positioning tray and a detachable forehead support. During oblique imaging, the cadaver head was placed on the default tray and attached to the forehead support. In order to acquire prone and supine images, the forehead support was removed and the gantry was tilted accordingly.

**Imaging positions**

A total of nine CBCT images were acquired in supine (Figure 1B), prone (Figure 1C) and oblique (Figure 1D) imaging positions. Furthermore, one MSCT image was acquired using the default supine imaging position (Figure 1A).

**Exposure protocols**

CBCT images were obtained using a constant 96 kV tube voltage. Three different tube currents (1): 4.8 mA, (2): 7.5 mA and (3): 12 mA were used to encompass the full dynamic range.
range of the mobile CBCT scanner. All CBCT exposures were performed using the smallest possible voxel size (0.2 mm). The MSCT exposures were performed using a default standard high-resolution imaging protocol (120 kV, 93 mA).

All effective doses required to calculate the figure of merit (FOM) value were scaled proportionally using the exposure values (mAs) attained in a previous study by Koivisto et al. [17] using an anthropomorphic phantom. All exposure protocols, parameters and calculated effective doses are presented in Table 1.

Standard axial, coronal and sagittal reconstructions were performed on a Philips Brilliance Extended Workstation V 4.5.2.40007 (Philips Healthcare, the Netherlands).

Table 1. Exposure parameters of the CBCT and MSCT exposures

<table>
<thead>
<tr>
<th>Planmed Verity</th>
<th>Philips Brilliance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CBCT</strong></td>
<td><strong>64 MSCT</strong></td>
</tr>
<tr>
<td><strong>Imaging position</strong></td>
<td><strong>Imaging protocol</strong></td>
</tr>
<tr>
<td>Supine/Prone/Oblique</td>
<td>Protocol 4.8 mA/7.5 mA/12 mA</td>
</tr>
<tr>
<td><strong>Imaging protocol</strong></td>
<td><strong>Parameter</strong></td>
</tr>
<tr>
<td><strong>Tube voltage (kVp)</strong></td>
<td>96/96/96</td>
</tr>
<tr>
<td><strong>Tube current (mA)</strong></td>
<td>4.8/7.5/12</td>
</tr>
<tr>
<td><strong>Exposure time (s)</strong></td>
<td>6/6/6</td>
</tr>
<tr>
<td><strong>Q (mAs)</strong></td>
<td>28.8/45/72</td>
</tr>
<tr>
<td><strong>Slice thickness (mm)</strong></td>
<td>0.2/0.2/0.2</td>
</tr>
<tr>
<td><strong>Slice increment (mm)</strong></td>
<td>0.2/0.2/0.2</td>
</tr>
<tr>
<td><strong>Pitch (mm)</strong></td>
<td>NA/NA/NA</td>
</tr>
<tr>
<td><strong>CTDIvol (mGy)</strong></td>
<td>-/-/-</td>
</tr>
<tr>
<td><strong>Voxel size (mm)</strong></td>
<td>0.2 x 0.2 x 0.2/0.2 x 0.2 x 0.2/0.2 x 0.2 x 0.2</td>
</tr>
<tr>
<td><strong>Scan angle</strong></td>
<td>210°/210°/210°</td>
</tr>
<tr>
<td><strong>Frame number</strong></td>
<td>300/300/300</td>
</tr>
<tr>
<td><strong>Radiation source</strong></td>
<td>Pulsed/Pulsed/Pulsed</td>
</tr>
<tr>
<td><strong>Field of view (cm)</strong></td>
<td>13 x 16/13 x 16/13 x 16</td>
</tr>
<tr>
<td><strong>Eff.dose, Supine (µSv)</strong></td>
<td>125/195/312</td>
</tr>
<tr>
<td><strong>Eff.dose, Prone (µSv)</strong></td>
<td>97/151/242</td>
</tr>
<tr>
<td><strong>Eff.dose, Oblique (µSv)</strong></td>
<td>29/45/72</td>
</tr>
</tbody>
</table>

NA, not applicable; CTDI, computed tomography dose index.
*Calculated from the effective doses reported by Koivisto et al. 2014
The reconstructed image data consisting of 474 MSCT and 652 CBCT DICOM (Digital Imaging and Communications in Medicine) files were subsequently exported to an Agfa Impax PACS system (Agfa Healthcare, Mortsel, Belgium) and randomly visually graded by two experienced radiologists using the same viewing conditions. All evaluations were repeated after four weeks to assess the intra-observer reliability.

### Image quality assessment

All image quality assessments were performed on ten different anatomical landmarks in four different anatomical regions: sinus (blue), ear (yellow), maxilla (green) and mandible (red) (Figure 2). The image quality was graded by two radiologists using the following scale: Level 5; excellent delineation of structures and excellent image quality, Level 4; clear delineation of structures and excellent image quality, Level 3; anatomic structures still fully assessable in all parts and good image quality, Level 2; structures identified and results in adequate image quality, Level 1; anatomic structures not identifiable due to poor image quality, Level 0; no diagnostic value.

### Statistical analysis of the image quality

The inter-rater agreement was evaluated by calculating the percentage of absolute agreement and two-way random effects intra-class correlation coefficients (ICC) with absolute agreement corresponding to 95% confidence interval (CI) [29]. The interclass agreement was evaluated by calculating the percentage of absolute agreement and two-way random effects intra-class correlation coefficients (ICC) with absolute agreement corresponding to 95% confidence interval (CI) [29].

<table>
<thead>
<tr>
<th>Organ nr.</th>
<th>Region</th>
<th>Anatomical landmarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Sinus</td>
<td>Fossa olfactorius</td>
</tr>
<tr>
<td>2</td>
<td>Sinus</td>
<td>Lamina papyracea</td>
</tr>
<tr>
<td>3</td>
<td>Sinus</td>
<td>Hiatus maxillaris</td>
</tr>
<tr>
<td>4</td>
<td>Sinus</td>
<td>Middle turbinate</td>
</tr>
<tr>
<td>5</td>
<td>Sinus</td>
<td>Ostium of the osteomeatal c.</td>
</tr>
<tr>
<td>6</td>
<td>Ear</td>
<td>Carotid canal</td>
</tr>
<tr>
<td>7</td>
<td>Maxilla</td>
<td>Foramen infraorbitale</td>
</tr>
<tr>
<td>8</td>
<td>Maxilla</td>
<td>Tooth 16 apex</td>
</tr>
<tr>
<td>9</td>
<td>Mandible</td>
<td>Tooth 43 apex</td>
</tr>
<tr>
<td>10</td>
<td>Mandible</td>
<td>Foramen mentale</td>
</tr>
</tbody>
</table>

Figure 2. Anatomical landmarks assessed on the reconstructed MSCT and CBCT images
correlation coefficient (ICC) (Two-Way Mixed) model (2,2) was used to determine inter-rater reliability [30].

The intra-rater reliability was calculated for four selected protocols (oblique position (4.8 mA), supine position (12 mA) and prone position (7.5 mA)) using ICC (Two-Way Mixed) model (3,1) [30]. ICCs range from 0 with no agreement to +1 with perfect agreement in the interpretation of the ICC. Landis et al. [31] has characterized the values of reliability coefficients as follows: less than 0.4, poor; 0.4 to 0.75, fair to good; and greater than 0.75, excellent. Although arbitrary, these divisions provide useful benchmarks [29].

Contrast-to-noise ratio

The contrast-to-noise ratio (CNR) evaluation was performed in the fossa olfactorius (anatomical landmark no. 6) and the apical aspect of tooth 43 (D43, anatomical landmark no. 2). These anatomical sites were chosen since they represent commonly diagnosed regions in the head area. The locations of ROIs were verified by superimposing CBCT and MSCT images. During superimposition, all CBCT and corresponding MSCT images were manually aligned into a coordinate system using control points chosen by the experienced radiologist [32]–[34].

The CNR was calculated using the following equation:

\[
CNR = \frac{S_{ROI} - S_{BGR}}{\sigma_{mean}}
\]  

where, \(S_{ROI}\) is the mean signal at the anatomical landmark in grey value unit, \(S_{BGR}\) is the mean signal in the background (HU) and \(\sigma_{mean}\) is the average standard deviation in the anatomical landmark and background ROIs (Figure 3). Furthermore, the correlation between CNR and the subjective image quality was calculated to assess the association between the two image quality indicators.

![Figure 3](image)

**Figure 3.** Axial view of the tooth 43 (apex, D43) and background locations used for CNR calculation for MSCT supine position (A) and for CBCT using supine (B), prone (C) and oblique (D) positions.
IMPACT OF PRONE, SUPINE AND OBLIQUE PATIENT POSITIONING ON CBCT IMAGE QUALITY

Figure of Merit
In this study, the figure of merit (FOM) value was used to assess the diagnostic efficacy of the image quality versus the effective dose obtained using different imaging protocols. The FOM value was calculated using the following equation described by Ogden et al. [26]

\[
FOM = \frac{\text{CNR}^2}{E}
\]

where E is the effective dose. In the present study, the FOM value was calculated for two anatomical landmarks: the fossa olfactorius and the apex of tooth 43.

RESULTS

Image quality assessment
The average image quality results of nine different CBCT imaging protocols acquired in different imaging positions and one MSCT protocol acquired in one position are presented in Figure 4.

Statistical analysis of image quality
Statistical analysis was performed on all CBCT and MSCT image quality results. The absolute agreement percentages, inter-rater intraclass correlation coefficients (ICC) and 95% confidence intervals (CI) of all examination protocols are presented in Table 2. The intra-rater reliability results of the two observers are presented in Table 3.

![Figure 4](image)

Figure 4. Results of subjective image quality assessments undertaken by two radiologists using different CBCT and MSCT imaging protocols.
Contrast-to-noise ratio

The mean CBCT contrast-to-noise values calculated in different imaging positions were as follows: 3.6 in the prone imaging position, 1.8 in the supine imaging position and 1.5 in the oblique imaging position. The Philips Brilliance 64 MSCT scanner mean CNR was 2.6. Furthermore, the correlation between CNR and subjective image quality were 0.8217 for tooth 43 and 0.3581 for the fossa olfactorius. The CNR values and the subjective image qualities of the apex of tooth 43 and the fossa olfactorius are presented in Figure 5A and Figure 5B, respectively.

Figure of Merit

The mean FOM values of the Planmed Verity CBCT scanner in different imaging positions were as follows: 96.1 in the prone imaging position, 24.3 in the supine imaging position and 22.3 in the oblique imaging position. The mean FOM value of the Philips Brilliance 64 MSCT scanner was 10.6. The FOM values of the apex of tooth 43 and the fossa olfactorius obtained using one MSCT modality and multiple CBCT imaging modalities are presented in Figure 6A and Figure 6B, respectively.
**DISCUSSION**

In this study, the impact of supine, prone and oblique imaging positioning on image quality in four different skull regions was assessed. Furthermore, the contrast-to-noise ratio (CNR) and the figure of merit value were calculated and compared. Finally, the image quality of CBCT images acquired in the supine imaging position were compared to MSCT images acquired in the same imaging position.

The combined mobile extremity/maxillofacial CBCT scanner used in this study is unique in that it uses an adjustable gantry instead of a fixed gantry [35]. The adjustable gantry which moves in an enclosed housing, offers the possibility of acquiring patient images in seated and standing positions. Furthermore, supine, prone and oblique imaging positions are possible. The position of the head with respect to the X-ray beam and reciprocal sensor in the mobile CBCT scanner is almost identical to other dental CBCT scanners on the market in the supine (Figure 2B) and prone (Figure 2C) imaging positions. Therefore, there are no
major image quality differences between mobile CBCT scanners and other dental CBCT scanners. However, the results found in this study concerning the oblique imaging position are unique to the mobile CBCT scanner and therefore do not apply to other dental CBCT scanners.

**CBCT image quality**

*Sinus:* In the sinus region, the differences between the image qualities acquired using the supine and oblique imaging positions were negligible. However, the best CBCT mean image quality was observed in the prone imaging position. When compared to the prone and oblique imaging positions, the supine imaging position offered only poor image quality in the fossa olfactorius using 4.8 mA tube current. Furthermore, inadequate image quality was observed in the fossa olfactorius using the prone and supine imaging positions and 7.5 mA tube current. In the sinus region, only minor variations in the image quality were observed when using 4.8 mA and 7.5 mA. However, when using the highest 12 mA tube current, a significantly better image quality was observed in the sinus area. This phenomenon could be caused by the better signal-to-noise ratio attained using 12 mA tube current. Moreover, the better image quality observed in the prone imaging position could be due to the closer proximity of the cadaver head to the scanner sensor.

The low image quality achieved in the supine imaging position can be explained by the stronger attenuation and scattering resulting from the thick and complex bony structures in the anterior part of the skull [32], [33]. Furthermore, the low image quality observed in the oblique imaging position is likely to be caused by the longer radiation path length and subsequent attenuation caused by the parietal bones at the top of the cranium [22], [34]. These findings suggest that it could be advantageous to use the oblique imaging position instead of the supine imaging position in the sinus region. This hypothesis is in good agreement with a previous study by Heiland et al. who reported on the advantage of using the oblique imaging position to image the sinuses on a Siremobil Iso-C\textsuperscript{3D} CBCT scanner [1]. However it must be noted that an oblique imaging position is not attainable using a standard upright CBCT scanner.

*Ear:* In the ear region, the oblique imaging position resulted in poor image quality of the carotid canal. Furthermore, the supine and prone imaging positions offered only adequate image quality of the carotid canal. The low image quality observed in the oblique position could have been caused by the attenuation of the parietal bones situated on the top of the cranium.

The difficulty of acquiring adequate image quality in the carotid canal regardless of the tube current (4.8 mA to 12 mA) could be due to the attenuation caused by the dense bony structures of the temporal bone. These findings are in good agreement with a previous study by Pereira et al. who reported on the difficulty of obtaining adequate images in the TMJ region due to the dense bony structures in the petrous part of the temporal bone [36]. These findings are further supported by the results of Kwong et al. [37] and Sohaib
et al. [38] who concluded that no significant image quality difference occurred when tube currents were changed.

**Maxilla:** Using the mobile CBCT scanner, the prone and supine imaging positions offered better image quality in the maxilla than the oblique imaging position. The prone imaging position offered the best image quality of tooth 16 and the foramen infraorbitale using 12 mA. The impact of different tube currents on image quality in all imaging positions was moderate.

**Mandible:** In the mandible, the prone and supine imaging positions resulted in good image quality for all three investigated tube currents. Furthermore, the oblique imaging position demonstrated an adequate image quality of the apex of tooth 43 and poor image quality of the foramen mentale using 7.5 mA tube current. However, only poor image qualities were attainable in the mandible using the oblique imaging position and 4.8 mA tube current. The impact of different tube currents on the image quality in the prone imaging position was negligible and moderate in the oblique imaging position. However, the tube current did not have any effect on image quality in the supine imaging position.

In general, the low image quality observed in this study using the oblique imaging position is likely due to the anatomy of the head. Moreover, in the oblique position, the x-ray beam enters the head on the back side of the cranium, and therefore the radiation path length through the head becomes longer than in the prone position. This phenomenon subsequently causes stronger attenuation, which has a negative effect on the overall image quality [17].

**Comparison between MSCT and CBCT images acquired in the supine position**

The MSCT, CBCT inter-rater and intra-rater reliabilities were assessed according to the characterization method described by Landis et al. [31] The MSCT intra-rater reliability was excellent. On the other hand, the CBCT intra-rater reliability was excellent in the oblique and prone imaging positions and only good in the supine position.

In the supine position, the mobile CBCT scanner offered similar image quality to that of the MSCT scanner when using 7.5 mA tube current. However, the mobile CBCT scanner resulted in lower image quality and higher local variations in image quality when a 4.8 mA tube current was used. Furthermore, when using 12 mA tube current, the mobile CBCT scanner offered a good delineation of most anatomical structures and better image quality than those acquired on the MSCT scanner.

**Contrast-to-noise ratio**

In this study, the best CNR value was observed in the apex of tooth 43 in the prone imaging position. The lower CNR values observed in the oblique imaging position could have been caused by the longer radiation path length and the parietal bones on the top of the cranium. Interestingly, in the oblique imaging position, the delineation of the fossa
olfactorius resulted in a higher CNR value than in the supine imaging position. The lower CNR value observed in the fossa olfactorius using the supine imaging position can be explained by the stronger attenuation and scattering from the thick and complex bony structures in the anterior nasofrontal area of the skull. [39] Furthermore, the correlation between the CNR value and subjective image quality results was excellent in the apex of tooth 43. However, only a fair correlation between the CNR value and image quality was achieved in the fossa olfactorius. These results are in good agreement with a recent study by Choi et al. [22] who described a strong association between subjective image quality and CNR values.

**Figure of Merit**

The prone imaging position offered the best figure of merit value in the fossa olfactorius and the apex of tooth 43. The differences between the oblique and supine imaging positions were negligible. However, in the fossa olfactorius, the differences in the FOM values attained in the oblique and supine imaging positions were significant. This could have been caused by the highest CNR value attained at a moderate effective dose.

The MSCT scanner resulted in a higher FOM value than the mobile CBCT scanner in the supine imaging position. However, the figure of merit value obtained using the MSCT was lower than those acquired using the mobile CBCT scanner in prone and oblique imaging positions.

The lower figure of merit value obtained on the MSCT scanner can be explained by the manufacturer-recommended high-resolution imaging parameters that subsequently resulted in a higher effective dose than in the mobile CBCT scanner. Nevertheless,
the effective dose can be significantly reduced by using a lower mAs value that subsequently does not markedly affect the diagnostic image quality in the sinus region [38].

One weakness of this study is the use of a frozen cadaver head for image quality assessment. During freezing, residual water in the cadaver expands and subsequently causes anatomical alterations especially in the fine bony structures of the sino-nasal region. Furthermore, frozen fatty acids in the brain [40] decrease in volume and consequently increase the attenuation resulting in an increase of the HU- or the grey-values [40], [41]. The aforementioned factors subsequently reduce the image quality and contrast resolution [42]. Moreover, the figure of merit value was assessed based on the CNR value from a cadaver head while the effective dose used for calculating the FOM value was obtained using an anthropomorphic phantom. Another drawback of this study is that the image quality was assessed for only one anatomical landmark in the ear region while there is a growing interest for diagnostic imaging of temporal bones due to the cochlear implants [43]. Furthermore, an interpretation with the results obtained using the MSCT scanner should be carried out with care since the images were acquired using standard exposure parameters without image quality optimisation.

CONCLUSIONS
Head imaging positioning has an impact on the overall CBCT image quality. The best mean CBCT image quality, CNR value and FOM value were observed using the prone imaging position for all diagnostic regions. The oblique imaging position offers inadequate image quality except in the sinus region. The MSCT scanner offered similar mean image quality in the supine position to that acquired on the mobile CBCT scanner using 7.5 mA tube current.

CONFLICTS OF INTEREST
Juha Koivisto is an employee of Planmeca Oy.
REFERENCES


Impact of Prone, Supine and Oblique Patient Positioning on CBCT Image Quality


CT IMAGE SEGMENTATION METHODS FOR BONE USED IN MEDICAL ADDITIVE MANUFACTURING: A LITERATURE REVIEW

Maureen van Eijnatten, Roelof van Dijk, Johannes Dobbe, Geert Streekstra, Juha Koivisto, Jan Wolff

Medical Engineering & Physics [accepted Sept 2017]
ABSTRACT

Aim of the study
The accuracy of additive manufactured medical constructs is limited by errors introduced during image segmentation. The aim of this study was to review the existing literature on different image segmentation methods used in medical additive manufacturing.

Methods
Thirty-two publications that reported on the accuracy of bone segmentation based on computed tomography images were identified using PubMed, ScienceDirect, Scopus, and Google Scholar. The advantages and disadvantages of the different segmentation methods used in these studies were evaluated and reported accuracies were compared.

Results
The spread between the reported accuracies was large (0.04 mm to 1.9 mm). Global thresholding was the most commonly used segmentation method with accuracies under 0.6 mm. The disadvantage of this method is the extensive manual post-processing required. Advanced thresholding methods could improve the accuracy to under 0.38 mm. However, such methods are currently not included in commercial software packages. Statistical shape model methods resulted in accuracies from 0.25 mm to 1.9 mm but are only suitable for anatomical structures with moderate anatomical variations.

Conclusions
Thresholding remains the most widely used segmentation method in medical additive manufacturing. To improve the accuracy and reduce the costs of patient-specific additive manufactured constructs, more advanced segmentation methods are required.
INTRODUCTION

Additive manufacturing (AM), also referred to as three-dimensional (3D) printing, is becoming increasingly popular in medicine [1] since it offers the possibility to personalize patient care. [2] The use of AM anatomical models results in more precise treatment planning, better communication, [3,4] and improved training and education. [5,6] Furthermore, AM can be used for the fabrication of drill guides, [7] saw guides, [8] and patient-specific implants. [9] To date, medical AM is most commonly used in branches of surgery involving the musculoskeletal system, such as oral and maxillofacial surgery, traumatology, and orthopaedic surgery. However, it must be noted that the overall accuracy and repeatability of medical AM constructs used for bone reconstruction still need to be improved. [10]

In this context, a recent systematic review by Martelli et al. (2016) identified 34 studies (21.5%) that reported on the unsatisfactory accuracy of medical AM constructs. [11]

The current medical AM process used for the reconstruction of the musculoskeletal system can be divided into four basic steps: imaging (1); image processing (2), optionally followed by computer-aided design (3); and additive manufacturing (4). [12] Each of these steps can introduce geometric deviations that can cause distortions in the resulting medical AM constructs. [13] Recent studies, however, suggest that the majority of the inaccuracies are introduced during imaging (Figure 1: step 1) and image processing (Figure 1: step 2), rather than during the manufacturing, i.e., the 3D printing process, which is generally considered to be precise. [11,14,15]

Step 1: Imaging

CT scanners are best suited for imaging bony structures due to their superior hard tissue contrast and spatial resolution. [16] Today, a plethora of different CT technologies are available, ranging from single, helical CT to 128-slice dual-source CT configurations. Cone-beam computed tomography (CBCT) scanners are becoming increasingly popular in orthopaedic [17] and maxillofacial surgery [18] due to their lower radiation dose and costs. Raw CT data acquired during image acquisition is commonly reconstructed as a Digital Imaging and Communications in Medicine (DICOM) file.

One major challenge faced in medical AM is the large variety of different CT image acquisition and reconstruction parameters currently available (see Figure 1; step 1 A&B). To date, to the best of our knowledge, there are no standardized protocols available for medical AM. Image slice thickness and slice interval have been identified as the primary limiting factors for the overall accuracy of medical AM constructs, [19] especially when reconstructing thin bony structures from axial plane images, such as the orbital floor, [20] or where the imaging plane is nearly parallel to the bone surface to be reconstructed, such as in the tibial plateau. Moreover, imaging noise, beam hardening, patient motion, and metal artifacts can introduce inhomogeneities in the gray values of CT images.
Step 2: Image processing

The medical AM process always requires image processing: the conversion of CT images into 3D surface models (Figure 1, step 2). Such 3D surface models can be saved in a wide range of different file formats. [21] Currently, the most commonly used file format in medical AM is standard tessellation language (STL). Although rarely used, it is theoretically possible to convert CT images into slice formats that can subsequently be used for AM, thereby skipping the STL conversion process. However, it must be noted that the computer-aided design (CAD) software packages currently available on the market for medical AM still require STL files (Figure 1, step 3).

The conversion of DICOM datasets into 3D surface models has been identified as a major source of inaccuracies in medical AM. [14,22,23] The most important step in this conversion process is image segmentation (Figure 1, step 2A). [24] Segmentation refers to the partitioning of images into regions of interest (ROIs) that correspond to anatomical structures. At present, there are a plethora of different image segmentation methods available for bony structures. [25] It remains unclear, however, which segmentation method offers the most accurate 3D surface models. Therefore, the first aim of this study was to review the existing literature on the different CT image segmentation methods currently being used for bone segmentation in medical AM. The second aim was to evaluate the impact of the different image segmentation methods on the geometric accuracy of 3D surface models.

MATERIALS AND METHODS

Existing literature on the CT image segmentation of bony structures for medical AM applications was reviewed using the PubMed Medline literature database, ScienceDirect, Scopus, and Google Scholar. An initial database of 17,700 publications was generated using the search terms “medical AND X” + “medicine AND X”, with X representing the interchangeably used terms “additive manufacturing”, “rapid prototyping”, and “3D printing”. The acquired database was subsequently filtered using the search terms “bone” (7,630), “segmentation” (281), and “accuracy” (138). In order to acquire comparable results, the following inclusion and exclusion criteria were used to isolate the publications of interest to this study:

Inclusion criteria:
• Evaluates the geometric accuracy of 3D surface models of bone derived from CT images
• Reports accuracy using a mean spatial distance-based metric in millimeters
Exclusion criteria:

- Uses non-biological materials or artificial phantoms

- Reports accuracy using non-spatial distance-based metrics, such as relative distance or volume difference (%), maximum distance, 95%-value of all distances (95\textsuperscript{th} percentile), overlap-based metrics, and contour-based metrics.

- Uses a non-CT-based imaging technology, e.g., magnetic resonance imaging (MRI)

- Only reports on the accuracy of the final AM model and not on the preceding 3D surface model

* This review focuses on the different CT image segmentation methods used in medical AM.

Figure 1. Overview of the parameters that can influence the accuracy of medical AM constructs.
Using the aforementioned inclusion and exclusion criteria, 10 publications were identified that contained information on the geometric accuracy of 3D surface models of bone acquired from CT images that were subsequently used for medical AM. From a parallel search, 22 additional publications were identified that reported on the accuracy of 3D surface models but did not focus on AM applications. Additional information was collected from four review articles focusing on medical AM. [19,26–28]

All relevant details of the publications included in this study are summarized in Table 1. The publications were subsequently sorted based on the type of image segmentation method used in the study. The reported accuracies and standard deviations (if available) are summarized in Table 2. In cases where multiple image segmentation methods, software packages, anatomical structures, or CT modalities were evaluated, their recorded accuracies are presented separately in Table 2. When multiple measurements were performed using the same image segmentation method or anatomical structure, the mean value and standard deviation of the reported accuracies are presented as one entry in Table 2.

RESULTS

A total of 32 publications were included in this study (Table 1). These publications were published between 2002 and 2017. To the best of our knowledge, no papers on the accuracy of medical AM constructs were published before 2002. [29] However, it must be noted that the first publications on medical AM date back as far as the early 1990s. [30]

CT image segmentation methods used in medical AM

The reviewed publications reported that the following three semi-automated CT image segmentation methods are often used in medical AM: global thresholding, [31] edge detection, [32] and region growing. [33]

Of the image segmentation methods mentioned above, global thresholding is the most commonly used bone segmentation method in medical AM. In global thresholding, a single threshold value \( t \) for bone is manually [34] or automatically [31,35] selected. All voxels with a gray value equal or greater than \( t \) are included in a segmented volume using a binary mask \( M_{x,y} \) (Equation 1):

\[
M_{x,y} = \begin{cases} 
1 & I_{x,y} \geq t \\
0 & I_{x,y} < t 
\end{cases}
\]  

where \( I_{x,y} \) denotes the grey value at coordinates \( x \) and \( y \) in a CT image slice.

Global thresholding has certain drawbacks, namely, voxels reside on tissue boundaries that contain more than one tissue type and induce a blurring of gray values across boundaries. This phenomenon is referred to as the partial volume effect (PVE). As
a consequence of the PVE, precise delineation of tissue boundaries using only a single threshold value remains difficult and can result in an over- or underestimation of the region of interest. Furthermore, a single threshold value does not take artifacts, image noise, photon starvation, or variations in gray values between different CT scanners and protocols into account, which can lead to inconsistent segmentation results. These drawbacks have led to the development of more sophisticated thresholding methods such as local thresholding.

Local thresholding, also referred to as multi-level thresholding, divides an image into multiple ROIs for which an individual thresholding bandwidth \( t_k \) to \( t_{k+1} \) can be selected. [36] All voxels with a grey value between \( t_k \) and \( t_{k+1} \) are included in \( k \) segmented volumes using a binary mask \( M_{x,y} \) (Equation 2):

\[
M_{x,y} = \begin{cases} 
0 & \text{if } l_{x,y} \leq t_1 \\
1 & \text{if } t_1 < l_{x,y} \leq t_2 \\
2 & \text{if } t_2 < l_{x,y} \leq t_3 \\
\cdot & \text{if } \ldots \\
k & \text{if } t_k < l_{x,y} \leq t_{k+1} \\
0 & \text{if } l_{x,y} > t_{k+1}
\end{cases}
\]  

(2)

where \( k \) denotes the \( k^{th} \) grey value band, of which \( t_1 \) is the lower limit and \( t_{k+1} \) the upper limit. Furthermore, 3D adaptive thresholding algorithms have recently been developed that offer the possibility to iteratively update the voxel classification in an additional second step after global thresholding. [37]

Edge detection, on the other hand, identifies local edges on CT images by calculating gray value gradients (derivatives). The gradients with a magnitude that is higher than a chosen threshold value are defined as edges. A range of edge detection operators are currently available that includes Sobel, Laplacian (2\text{nd} derivative) and Canny. [25] Canny edge detection remains one of the most commonly used, fastest and most accurate operators. [36] Edge detection is well suited for the segmentation of structures with different contrast in different regions, such as long bones. However, local gray value changes induced by noise and (metal) artifacts are often mistakenly identified as edges. It should also be noted that, without further processing, edge detection methods do not necessarily segment all bone voxels in the image and thus need to be combined with other methods, such as region growing.

In region growing, a specific voxel is manually selected as a seed point for a specific tissue type. Subsequently, the gray values of the neighboring voxels are compared to the gray value of the seed point. Voxels that meet the predefined homogeneity criteria are labeled and grouped together, thereby creating a region \( R_i \) of connected voxels in the CT image \( I \):
In practice, region growing is seldom the only segmentation method used, but it is commonly combined with other methods such as (global) thresholding. An advantage of region growing is that it discards voxels that are not connected to the anatomical structure of interest, resulting in a shorter 3D printing time and material savings. A disadvantage of region growing is that each separate bony structure requires an individual, manually placed seed point. Furthermore, noise and the partial volume effect can cause voids or erroneously connected structures in a segmented image.

In recent years, a number of advanced CT image segmentation methods have been developed. [38] The most established and validated advanced image segmentation method to date is based on statistical shape models (SSM). SSM-based methods generate a statistical shape model of an anatomical ROI from a training dataset. This training dataset comprises a large number of fully segmented CT images, which are averaged to form a mean anatomical shape model, along with common modes of variation and their probabilities. The model given by:

$$\bar{S}(\hat{b}) = \hat{v} + \sum_k b_k p_k,$$

describes every possible shape $\bar{S}$ as a linear combination of $k$ eigenmodes $\hat{p}$ with weights $\hat{b}$, where $\hat{v}$ denotes the mean shape model. This model can be fitted to a new, unsegmented image dataset, and the segmentation in terms of weights $\hat{b}$ then forms an approximation of the segmented shape.[39]

The major challenge faced in SSM-based segmentation is the large number of uniform training datasets required. Another drawback with SSM-based methods is that they are prone to failure if the datasets differ strongly from the training shapes. In clinical settings, such differences can occur, for example, in fractured bony structures that no longer correspond to the initial anatomical shape. Another example is that an SSM trained on a European population may not account for some of the morphological differences present in an African population.

Accuracy of CT image segmentation methods
In all the reviewed publications, the accuracy was assessed by comparing the CT-derived 3D surface models to a "ground truth" (Table 1). The ground truth was obtained either by performing manual segmentation by an expert, by obtaining a laser surface scan or optical 3D surface scan from the dry or dissected bone, or by performing linear measurements on the dry or dissected bone using calipers or a coordinate measurement machine (CMM). In all publications that focused on SSM-based segmentation methods, the accuracy was
Accuracy of CT image segmentation methods was evaluated using leave-one-out cross validation, in which the distances were measured from the omitted training shape to the closest match of the “ground-truth” shape model trained on all other available shapes.

In total, 22 of the reviewed publications (67%) evaluated the accuracy by calculating the mean absolute distance between a CT-derived 3D surface model and a geometrically aligned ground truth 3D surface model. Let us define the distance $d(p,q)$ as the distance between a point $p$ belonging to the surface of a 3D surface model $P$ and its closest point $q$ on the ground truth model $Q$. The mean surface distance (MSD) between the two surfaces $P$ and $Q$ can then be defined as follows:

$$MSD = \frac{1}{|P|} \iint_{p \in P} d(p, Q) \, dP$$  \hspace{1cm} (5)$$

where $|P|$ denotes the surface area of $P$. \[40\] Note that some variations in the definition of $d(p,q)$ were found in the reviewed publications. For example, Smith et al. (2013) used perpendicular projections of the distance of each point $p$ along the triangle normal vector, \[14\] whilst Gassman et al. (2008) created a distance map based on the Euclidean distance, or the shortest distance from each point $p$ to the target surface $Q$. \[41\] Finally, Zhang et al. \[42\] and Rajamani et al. \[43\] used the Hausdorff distance (HD), which can be defined as follows:

$$HD(P, Q) = \max_{p \in P} \{ \min_{q \in Q} \{ d(p, q) \} \}$$  \hspace{1cm} (6)$$

Two publications \[44,45\] used the root-mean-square distance (RMSD) that can be derived from Equation 5:

$$RMSD = \sqrt{\frac{1}{|P|} \iint_{p \in P} d(p, Q)^2 \, dP}$$  \hspace{1cm} (7)$$

In four publications, \[29,46–48\] the mean absolute distance (MAD) was calculated by performing a limited number of linear measurements $n$, which can be written as follows:

$$MAD = \frac{1}{n} \sum_{n} |D_n - x_n|$$  \hspace{1cm} (8)$$
where $D$ is a linear measurement performed on the ground truth, and $x$ is a measurement performed on the CT-derived 3D surface model. Additionally, in four publications, [49–52] the mean distance (MD) was calculated as the signed mean of $n$ measurements between different anatomical landmarks, [49,52] reference points, [51] or photographs of a cryosectioned cadaver, [50] which can be described as follows:

$$MD = \frac{1}{n} \sum_{n} (D_n - x_n)$$

Table 2 summarizes the geometric accuracies reported in the publications. All 3D surface model accuracies generally varied between 0.04 mm and 1.9 mm. Manual segmentation resulted in an accuracy of 0.20 mm ± 0.15 mm to 0.48 mm ± 0.51 mm. [53,54] In global thresholding, accuracies varied between 0.04 mm ± 0.59 mm and 0.62 mm ± 0.76 mm. [14,29,36,46,48,49,52,55–59] However, it must be noted that most authors reported that additional manual editing was performed. Advanced thresholding methods resulted in accuracies below 0.38 mm ± 0.14 mm. [36,37,42,45,46] SSM-based segmentation methods achieved accuracies ranging from 0.25 mm ± 0.20 mm to 1.9 mm ± 0.20 mm. [44,60–68] Finally, the publications on more experimental segmentation methods (“other” in Table 2) reported geometric accuracies ranging between 0.28 mm ± 0.04 mm and 1.27 mm ± 0.92 mm. [41,47,50,51,59,69,70]

DISCUSSION

To date, the most common CT image segmentation method used for bone segmentation in medical AM is global thresholding. All publications reviewed in this study that used global thresholding reported geometric accuracies under 0.62 mm ± 0.76 mm [59] (Table 2). However, the reported accuracies generally included additional extensive manual post-processing, which is often very time consuming. Furthermore, the reproducibility of manual post-processing remains a challenge, especially when suboptimal CT images with thick slices, noise, metal artifacts or motion artifacts are used. In this context, Fasel et al. [56] reported that a manual clean-up time of 10 hours was required to achieve an accuracy of 0.25 mm ± 0.23 mm. Such extensive post-processing times are currently a major reason for the high costs related to medical AM. [27]

In order to overcome this time-consuming manual task, more advanced thresholding methods can be applied. In this context, an algorithm using multiple local thresholds, automatically determined using edge detection, achieved an accuracy of 0.18 mm ± 0.02 mm without the need for any supplementary manual editing. [36] Furthermore, another advanced thresholding method that iteratively reclassifies boundary voxels according to their surroundings resulted in accuracies of 0.08 mm (calcaneus) and 0.20 mm (spine). [37,42] However, to the best of our knowledge, such advanced thresholding methods
are currently not implemented in commercially available medical AM software packages. In this context, Kang et al. [49] compared four different commercial software packages and reported that global thresholding was the segmentation method of choice in all four software packages.

Currently, only two commercial software packages for medical AM have been approved by the United States Food and Drug Administration (FDA): Mimics™ (Materialise®, Belgium), and Osirix® MD (Pixmeo, Switzerland). [71] Furthermore D2P™ (DICOM TO PRINT) (3D Systems, Rock Hill, USA) has very recently (Jan 9, 2017) received 510k clearance and is currently pending CE approval. These software packages are generally equipped with global thresholding, region growing, dynamic region growing, local editing by a local threshold, and multiple manual mask editing tools (e.g., delete, draw). More advanced semi-automated image segmentation algorithms are currently available in Amira® 3D analysis software for Life Sciences (FEI Visualization Sciences Group, Hillsboro, OR, USA) that allow for interactive segmentation using local image histograms and edge detection algorithms. However, this software is not presently cleared for medical use, although this will be a requirement from 2020 onwards under the new European Union medical device regulation (MDR) for patient-specific constructs. [72] In this context, standardization and CE certification will be required for the whole medical AM process (Figure 1), including the evaluation of the repeatability and accuracy of the image segmentation process.

The previously discussed drawbacks of global and advanced thresholding have led to the development of novel SSM-based CT image segmentation approaches. However, the SSM publications assessed in this study demonstrated a large variability in segmentation accuracies (Table 2). Moreover, only two out of eleven publications (18%) reported accuracies equal to or less than 0.5 mm. [60,61] The major challenge in SSM-based segmentation methods is the quality and number of available training datasets, which can affect the generalization ability and specificity of an SSM. [39] The generalization ability quantifies to which extent the variability of the anatomical shapes in the patient population are incorporated in the SSM. For example, when too few training datasets are used, the SSM may fail to represent new shapes. The specificity of an SSM refers to its ability to only generate segmented shapes that are similar to those incorporated in the training datasets. Therefore, SSM-based segmentation methods are only applicable for CT datasets with moderate anatomical variations. Note that 3D statistical deformable models have also been used to segment bony structures, such as the knee joint, on magnetic resonance images. [73]

Clinical implications

The majority of the accuracies that are depicted in Table 2 may be considered sufficient for the creation of anatomical models, [2] of which hundreds are being successfully manufactured and used in clinical settings every year. [74,75] However, the use of inaccurate 3D surface models to create patient-specific AM implants or guides can lead to ill-fitting
Table 1. Publications included in this review.

<table>
<thead>
<tr>
<th>Ref</th>
<th>Year</th>
<th>Author</th>
<th>Segmentation method</th>
<th>Software</th>
<th>Anatomy (#)</th>
</tr>
</thead>
<tbody>
<tr>
<td>55</td>
<td>2016</td>
<td>Szymor</td>
<td>Global thresholding</td>
<td>Slicer</td>
<td>Skull (1)</td>
</tr>
<tr>
<td>61</td>
<td>2015</td>
<td>Chu</td>
<td>Random forest &amp; multi-atlas</td>
<td>Custom &amp; Amira</td>
<td>Femur (60) + acet. (60)</td>
</tr>
<tr>
<td>48</td>
<td>2015</td>
<td>Poleti</td>
<td>Global thresholding</td>
<td>Dolphin (A)/ InVesalius (B)</td>
<td>Mandible (10)</td>
</tr>
<tr>
<td>64</td>
<td>2014</td>
<td>Balestra</td>
<td>Articulated SSM</td>
<td>Custom &amp; Amira</td>
<td>Femur (26) + pelvis (26)</td>
</tr>
<tr>
<td>56</td>
<td>2014</td>
<td>Fasel</td>
<td>Global thresholding</td>
<td>n.a. (3 institutes)</td>
<td>Skull (3)</td>
</tr>
<tr>
<td>49</td>
<td>2014</td>
<td>Kang</td>
<td>Global thresholding</td>
<td>InVivoDental (A) / Mimics (B) / OnDemand3D (C) / OsiriX (D)</td>
<td>Skull (3)</td>
</tr>
<tr>
<td>46</td>
<td>2014</td>
<td>Santolaria</td>
<td>Global thresholding (A) / Advanced thresholding (B)</td>
<td>Mimics</td>
<td>Mandible (1)</td>
</tr>
<tr>
<td>57</td>
<td>2014</td>
<td>Van den Broeck</td>
<td>Global thresholding</td>
<td>Mimics</td>
<td>Tibia (10)</td>
</tr>
<tr>
<td>44</td>
<td>2014</td>
<td>Zhang</td>
<td>Region-based SSM</td>
<td>Custom &amp; Stradwin</td>
<td>Femur (41)</td>
</tr>
<tr>
<td>60</td>
<td>2013</td>
<td>Chang</td>
<td>Base invariant wavelet active shape model</td>
<td>Custom</td>
<td>Maxilla (19)</td>
</tr>
<tr>
<td>47</td>
<td>2013</td>
<td>Engelbrecht</td>
<td>“Commercial company” (A) / “doctor” (B)</td>
<td>Mimics &amp; SimPlant Ortho Pro</td>
<td>Mandible (7)</td>
</tr>
<tr>
<td>14</td>
<td>2013</td>
<td>Smith</td>
<td>Global thresholding</td>
<td>Mimics</td>
<td>Hip (6) + shoulder (6) joint</td>
</tr>
<tr>
<td>69</td>
<td>2013</td>
<td>Yokota</td>
<td>SSM</td>
<td>Custom</td>
<td>Femur (200) + acet. (200)</td>
</tr>
<tr>
<td>45</td>
<td>2013</td>
<td>Zhou</td>
<td>3D adaptive thresholding</td>
<td>MATLAB</td>
<td>Hip joint (70)</td>
</tr>
<tr>
<td>69</td>
<td>2012</td>
<td>Oneill</td>
<td>Semi-automatic morphological snakes</td>
<td>Custom</td>
<td>Femur (6)</td>
</tr>
<tr>
<td>58</td>
<td>2011</td>
<td>Anstey</td>
<td>Global thresholding</td>
<td>Mimics</td>
<td>Femur (1) + acet. (1)</td>
</tr>
<tr>
<td>36</td>
<td>2011</td>
<td>Rathnayaka</td>
<td>Global thresholding (A) / multi-thresholding (B) / edge detection (C)</td>
<td>MATLAB &amp; Amira</td>
<td>Femur (5)</td>
</tr>
<tr>
<td>54</td>
<td>2010</td>
<td>Thorhauer</td>
<td>Manual</td>
<td>Mimics</td>
<td>Knee (2)</td>
</tr>
<tr>
<td>63</td>
<td>2009</td>
<td>Kainmueller</td>
<td>SSM</td>
<td>Custom</td>
<td>Femur (30) + acet. (50)</td>
</tr>
</tbody>
</table>
## Table 1. Publications included in this review.

<table>
<thead>
<tr>
<th>Ref</th>
<th>Year</th>
<th>Author</th>
<th>Segmentation method</th>
<th>Software</th>
<th>Anatomy (#)</th>
<th>Material</th>
<th>Material (manufacturer/# slices)</th>
<th>CT tube voltage / tube current</th>
<th>CT slice thickness / voxel size</th>
<th>Ground truth</th>
</tr>
</thead>
<tbody>
<tr>
<td>55</td>
<td>2016</td>
<td>Szymor</td>
<td>Global thresholding</td>
<td>Slicer</td>
<td>Skull (1)</td>
<td>Dry bone</td>
<td>CBCT (Accuitomo)</td>
<td>90 kV / 5 mA</td>
<td>0.5 mm / 0.25 mm</td>
<td>Optical 3D scan</td>
</tr>
<tr>
<td>61</td>
<td>2015</td>
<td>Chu</td>
<td>Random forest &amp; multi-atlas &amp; articulated SSM</td>
<td>Custom &amp; Amira</td>
<td>Femur (60) + acet. (60)</td>
<td>In-vivo</td>
<td>MDCT</td>
<td>1.6 mm / 0.6 mm</td>
<td>1.6 mm / 18.5 mAs</td>
<td>Manual segmentation</td>
</tr>
<tr>
<td>48</td>
<td>2015</td>
<td>Poleti</td>
<td>Global thresholding</td>
<td>Dolphin (A)/ InVesalius (B)</td>
<td>Mandible (10)</td>
<td>Dry bone</td>
<td>CBCT (i-CAT)</td>
<td>120 kV / 18.5 mAs</td>
<td>- / 0.3 mm</td>
<td>Digital calipers</td>
</tr>
<tr>
<td>56</td>
<td>2014</td>
<td>Balestra</td>
<td>Articulated SSM</td>
<td>Custom &amp; Amira</td>
<td>Femur (26) + pelvis (26)</td>
<td>In-vivo</td>
<td>MDCT</td>
<td>1.6 mm / 0.6 mm</td>
<td>Manual segmentation</td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>2014</td>
<td>Fasel</td>
<td>Global thresholding</td>
<td>n.a. (3 institutes)</td>
<td>Skull (3)</td>
<td>Cadaver</td>
<td>MDCT (Siemens Sensation/64)</td>
<td>0.6 mm / 0.44 mm</td>
<td>Optical 3D scan</td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>2014</td>
<td>Kang</td>
<td>Global thresholding</td>
<td>InVivoDental (A) / Mimics (B) / OnDemand3D (C) / OsiriX (D)</td>
<td>Skull (3)</td>
<td>Dry bone</td>
<td>MDCT (Siemens Sensation/64)</td>
<td>0.6 mm / 0.44 mm</td>
<td>Optical 3D scan</td>
<td></td>
</tr>
<tr>
<td>46</td>
<td>2014</td>
<td>Santolaria</td>
<td>Global thresholding (A) / Advanced thresholding (B)</td>
<td>Mimics</td>
<td>Mandible (1)</td>
<td>Dry bone</td>
<td>CBCT (Kodak 9500)</td>
<td>- / 0.2 mm - 0.3 mm (CMM)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>57</td>
<td>2014</td>
<td>Van den Broeck</td>
<td>Global thresholding</td>
<td>Mimics</td>
<td>Tibia (10)</td>
<td>Cadaver</td>
<td>MDCT (GE LightSpeed VCT)</td>
<td>120 kVp / 160 mAs</td>
<td>0.6 mm / 0.39 mm</td>
<td>Optical 3D scan</td>
</tr>
<tr>
<td>47</td>
<td>2014</td>
<td>Zhang</td>
<td>Region-based SSM</td>
<td>Custom &amp; Stradwin</td>
<td>Femur (41)</td>
<td>In-vivo</td>
<td>MDCT</td>
<td>1.6 mm / 1.0 mm</td>
<td>Manual segmentation</td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>2013</td>
<td>Chang</td>
<td>Base invariant wavelet active shape model</td>
<td>Custom</td>
<td>Maxilla (19)</td>
<td>In-vivo</td>
<td>CBCT (Sirona)</td>
<td>- / 0.29 mm</td>
<td>Manual segmentation</td>
<td></td>
</tr>
<tr>
<td>69</td>
<td>2013</td>
<td>Engelbrecht</td>
<td>&quot;Commercial company&quot; (A) / &quot;doctor&quot; (B)</td>
<td>Mimics &amp; SimPlant Ortho Pro</td>
<td>Mandible (7)</td>
<td>Cadaver</td>
<td>CBCT (KaVo 3D exam)</td>
<td>- / 0.3 mm</td>
<td>Laser surface scan</td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>2013</td>
<td>Smith</td>
<td>Global thresholding</td>
<td>Mimics</td>
<td>Hip (6) + shoulder (6)</td>
<td>Cadaver</td>
<td>MDCT (GE Lightspeed+XCR/16)</td>
<td>1.25 mm / 0.68 mm</td>
<td>Manual segmentation</td>
<td></td>
</tr>
<tr>
<td>69</td>
<td>2012</td>
<td>Yokota</td>
<td>SSM</td>
<td>Custom</td>
<td>Femur (200) + acet. (200)</td>
<td>In-vivo</td>
<td>MDCT</td>
<td>1.5 mm / 0.73 mm</td>
<td>Manual segmentation</td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>2011</td>
<td>Zhou</td>
<td>3D adaptive thresholding</td>
<td>MATLAB</td>
<td>Hip joint (70)</td>
<td>In-vivo</td>
<td>MDCT (GE Toshiba)</td>
<td>1.5 mm / 0.73 mm</td>
<td>Manual segmentation</td>
<td></td>
</tr>
<tr>
<td>58</td>
<td>2011</td>
<td>Anstey</td>
<td>Global thresholding</td>
<td>Mimics</td>
<td>Femur (1) + acet. (1)</td>
<td>Cadaver</td>
<td>MDCT (GE LightSpeed+XCR/16)</td>
<td>0.63 mm / -</td>
<td>Laser surface scan</td>
<td></td>
</tr>
<tr>
<td>54</td>
<td>2010</td>
<td>Rathnayaka</td>
<td>Global thresholding (A) / multi-thresholding (B) / edge detection (C)</td>
<td>MATLAB &amp; Amira</td>
<td>Femur (5)</td>
<td>Cadaver</td>
<td>MDCT (Philips Brilliance/64)</td>
<td>140 kVp / -</td>
<td>3D contact scanner</td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>2011</td>
<td>O'neill</td>
<td>Semi-automatic morphological snakes</td>
<td>Custom</td>
<td>Femur (6)</td>
<td>In-vivo</td>
<td>MDCT (GE Lightspeed Plus)</td>
<td>1.5 mm / 0.73 mm</td>
<td>Manual segmentation</td>
<td></td>
</tr>
<tr>
<td>53</td>
<td>2010</td>
<td>Thorhauer</td>
<td>Manual</td>
<td>Mimics</td>
<td>Knee (2)</td>
<td>Cadaver</td>
<td>MDCT (GE Lightspeed/64)</td>
<td>1 mm / 0.9 mm</td>
<td>Laser surface scan</td>
<td></td>
</tr>
<tr>
<td>42</td>
<td>2009</td>
<td>Kainmueller</td>
<td>SSM</td>
<td>Custom</td>
<td>Femur (30) + acet. (30)</td>
<td>In-vivo</td>
<td>MDCT</td>
<td>0.5 mm - 5 mm / 0.5 mm - 0.9 mm</td>
<td>Manual segmentation</td>
<td></td>
</tr>
</tbody>
</table>
Table 1. (continued)

<table>
<thead>
<tr>
<th>Ref</th>
<th>Year</th>
<th>Author</th>
<th>Segmentation method</th>
<th>Software</th>
<th>Anatomy (#)</th>
</tr>
</thead>
<tbody>
<tr>
<td>70</td>
<td>2009</td>
<td>Ramme</td>
<td>Expectation maximization</td>
<td>BRAINS2</td>
<td>Phalanges (15)</td>
</tr>
<tr>
<td>42</td>
<td>2009</td>
<td>Zhang</td>
<td>3D adaptive thresholding</td>
<td>Custom</td>
<td>Calcaneus (1) + spine (2)</td>
</tr>
<tr>
<td>53</td>
<td>2008</td>
<td>DeVries</td>
<td>Manual</td>
<td>BRAINS2</td>
<td>Phalanges (15)</td>
</tr>
<tr>
<td>41</td>
<td>2008</td>
<td>Gassman</td>
<td>Artificial neural network</td>
<td>BRAINS2</td>
<td>Phalanges (15)</td>
</tr>
<tr>
<td>59</td>
<td>2008</td>
<td>Gelaude</td>
<td>Global thresholding (A) / self-developed contour-based algorithm (B)</td>
<td>Mimics</td>
<td>Femur (15)</td>
</tr>
<tr>
<td>62</td>
<td>2008</td>
<td>Seim</td>
<td>SSM</td>
<td>Custom</td>
<td>Pelvis (50)</td>
</tr>
<tr>
<td>51</td>
<td>2006</td>
<td>Gelaude</td>
<td>Adaptive contouring</td>
<td>Mimics &amp; MATLAB</td>
<td>Pelvis (1) + humerus (1)</td>
</tr>
<tr>
<td>52</td>
<td>2006</td>
<td>Loubele</td>
<td>Global thresholding</td>
<td>Custom</td>
<td>Skull (1)</td>
</tr>
<tr>
<td>66</td>
<td>2006</td>
<td>Rueda</td>
<td>Active appearance model</td>
<td>Custom</td>
<td>Mandible (215)</td>
</tr>
<tr>
<td>60</td>
<td>2005</td>
<td>Schmidt</td>
<td>Global thresholding / Self-developed border tracing algorithm *</td>
<td>Mimics / Custom</td>
<td>Femur (1) + acet. (1)</td>
</tr>
<tr>
<td>67</td>
<td>2004</td>
<td>Lamecker</td>
<td>SSM</td>
<td>Custom</td>
<td>Pelvis (23)</td>
</tr>
<tr>
<td>43</td>
<td>2004</td>
<td>Rajamani</td>
<td>SSM</td>
<td>Custom</td>
<td>Femur (14)</td>
</tr>
<tr>
<td>29</td>
<td>2002</td>
<td>Choi</td>
<td>Global thresholding</td>
<td>V-works</td>
<td>Skull (1)</td>
</tr>
</tbody>
</table>

Acet. = acetabulum; CMM = coordinate measurement machine; MDCT = multi-detector row computed tomography; C/ methods used.
<table>
<thead>
<tr>
<th>Material</th>
<th>CT scanner</th>
<th>CT tube voltage / tube current</th>
<th>CT slice thickness / voxel size</th>
<th>Ground truth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cadaver</td>
<td>MDCT (Siemens Sensation/64) 120 kVp / 105 mA</td>
<td>0.4 mm / 0.34 mm</td>
<td>Laser surface scan</td>
<td></td>
</tr>
<tr>
<td>In-vivo</td>
<td>MDCT (Toshiba HR / Siemens)</td>
<td>0.7 mm - 0.4 mm / 0.8 mm - 0.49 mm</td>
<td>Manual segmentation scan</td>
<td></td>
</tr>
<tr>
<td>Cadaver</td>
<td>MDCT (Siemens Sensation/64) 120 kVp / 105 mA</td>
<td>0.4 mm / 0.34 mm</td>
<td>Laser surface scan</td>
<td></td>
</tr>
<tr>
<td>Cadaver</td>
<td>MDCT (Siemens Sensation/64) 120 kVp / 105 mA</td>
<td>0.4 mm / 0.34 mm</td>
<td>Laser surface scan</td>
<td></td>
</tr>
<tr>
<td>Cadaver</td>
<td>MDCT (Siemens Sensation)</td>
<td>- / 1 mm</td>
<td>Optical 3D scan</td>
<td></td>
</tr>
<tr>
<td>In-vivo</td>
<td>MDCT</td>
<td>5 mm / 0.9 mm</td>
<td>Manual segmentation scan</td>
<td></td>
</tr>
<tr>
<td>Cadaver</td>
<td>MDCT</td>
<td>2 mm / -</td>
<td>Manual segmentation scan</td>
<td></td>
</tr>
<tr>
<td>Skull phantom (A)</td>
<td>MDCT (Siemens Sensation/16)</td>
<td>120 kV / 120 mA – 200 mA</td>
<td>Laser segmentation scan</td>
<td></td>
</tr>
<tr>
<td>In-vivo</td>
<td>MDCT</td>
<td>1 mm / -</td>
<td>Manual segmentation scan</td>
<td></td>
</tr>
<tr>
<td>Cryosectioned cadaver</td>
<td>MDCT</td>
<td>5 mm / 1.4 mm</td>
<td>Manual segmentation scan</td>
<td></td>
</tr>
<tr>
<td>In-vivo</td>
<td>MDCT</td>
<td>5 mm / 1.4 mm</td>
<td>Manual segmentation scan</td>
<td></td>
</tr>
<tr>
<td>Dry bone</td>
<td>MDCT (Siemens Plus/4) 120 kVp / 200 mA</td>
<td>1 mm / -</td>
<td>Anatomical landmarks</td>
<td></td>
</tr>
</tbody>
</table>

γγ; CBCT = cone-beam computed tomography. * Schmidt et al. (2005) did not disclose the identity of the segmentation methods used.
Table 2. Mean and standard deviation of the reported accuracies of different CT image segmentation methods. Cad = cadaver, MSD = mean surface distance; RMSD = root-mean-square distance; MAD = mean absolute difference; MD = mean difference.
implants and complications during surgery. Moreover, recent advances in computer-assisted surgery, such as surgical navigation, augmented reality, and robot-guided surgery, will require highly accurate 3D surface models (<1 mm). In this regard, it should be noted that accuracy also depends on the resolution of the CT scanner. Currently, routine diagnostic scan protocols typically have slice thicknesses ranging between 0.6 mm and 2 mm and voxel sizes ranging from 0.2 mm to 0.6 mm, which subsequently limits the accuracy that can be achieved in clinical settings.

Suggestions for improvement
The limitations of the thresholding and SSM-based CT image segmentation methods commonly used in medical AM are currently impeding the cost-effectiveness of patient-specific AM constructs due to the overall time required for the segmentation and manual post-processing. Therefore, more advanced methods are necessary that can accurately segment different anatomical structures and their morphological variations. Therefore, the authors of this study suggest that algorithms such as morphological snakes, artificial neural networks, and expectation maximization could markedly optimize CT image segmentation in medical AM. In particular, deep convolutional neural networks (CNNs) can be taught to obtain mid- and high-level abstractions from CT images.

Based on the reviewed publications, the authors of this study suggest that future research should focus on the clinical translation of advanced segmentation methods and intelligent user interfaces. Such a novel interface would facilitate local segmentation using multiple threshold values derived from local image histograms, edge detection algorithms, and CNNs. Furthermore, an engineer should be able to interact with the segmentation process in real time based on a preview of the 3D surface model. Such an approach will, however, only be possible with advancements in graphical computing power.

Limitations of the reviewed studies
A major limitation of the reviewed publications was the incomplete specifications of the CT data. A total of eleven publications (34%) did not mention the CT slice thickness and/or the voxel size (see Table 1). Furthermore, in twenty-three publications (72%) the CT tube voltage and tube current were not mentioned, and in eleven publications (34%) the manufacturer of the CT scanner was not specified.

Another challenge faced when evaluating the publications was the diversity of metrics used to assess the accuracies. A sizable body of work reported accuracies using relative distance-based metrics such as volume percentages. These publications were not included in this review. Only publications that used mean distance-based accuracy metrics in millimeters (mean surface distance, root-mean-square distance, mean absolute difference and mean difference) were included. However, it should be noted that the root-mean-square distance is frequently misinterpreted in practice, and that the mean absolute difference and mean difference are commonly only based on a limited number of linear
measurements. We therefore recommend that the geometric accuracy of bone 3D surface models be evaluated using the mean surface distance. Nevertheless, when only the mean surface distance is reported, large deviations in local anatomical regions can either remain unnoticed or can strongly influence the mean surface distance. [67] Therefore, the authors of this study recommend that an additional surface comparison is provided using colors to depict anatomy-specific deviations (a “heat map”). Such heat maps were used in eleven publications (34%).

Another limitation of the reviewed studies was the lack of information on the amount of manual editing that was required in order to achieve the reported accuracies. Manual editing inherently introduces intra- and inter-operator variability. Therefore, we suggest that the experience of the operator, the exact nature of manual editing, and the time spent should always be specified in detail.

Finally, numerous studies included in this review acquired high-resolution CT scans using dry bones (5 publications) or cadaver specimens (12 publications), which resulted in mean accuracies of 0.30 mm ± 0.30 mm and 0.46 mm ± 0.35 mm, respectively. In-vivo CT images (13 publications) resulted in a mean accuracy of 0.73 mm ± 0.36 mm. One explanation for these differences is that the photon energy spectrum of an X-ray beam changes as it passes through soft tissues, [81] which results in differences in absorption and scattering during CT imaging. Moreover, the gray values of soft tissues are closer to those of bone than air. Consequently, soft tissues make semi-automated bone segmentation less accurate. [82] Therefore, the reported accuracies found in those studies using dry bones are not translatable to clinical conditions. Furthermore, musculoskeletal bones in humans differ markedly in morphology, which causes differences in the accuracy and manual clean-up time. For example, segmentation of small scaffold-like orbital bones is more challenging than that of large femoral bones.

Taking the above into account, we hypothesize that the best method to evaluate the accuracy of medical AM constructs is to use multiple life-like cadaveric specimens and to subsequently remove all soft tissues from the bone after CT imaging. This offers the possibility to acquire an optical 3D surface scan of the bone that can be used as an accurate ground truth (<0.05 mm). Note that boiling or chemical treatments should be avoided since these treatments can introduce geometric deviations to the bone of up to 0.49 mm [59] and 0.43 mm, [57] respectively.

CONCLUSION

This literature review revealed that a plethora of different CT image segmentation methods are currently used for bone segmentation. The accuracy of these image segmentation methods differed markedly. Global thresholding remains the most widely used CT image segmentation method in medical AM but often requires extensive manual post-processing. Advanced thresholding approaches could improve the accuracy of global thresholding, but such methods are currently not implemented in commercially available
software packages. The development of fully automatic and adaptive image segmentation algorithms could improve the accuracy and reduce the costs of patient-specific AM constructs. More specifically, future research should focus on the development of image segmentation methods using CNNs.

COMPETING INTERESTS
None declared

FUNDING
None

ETHICAL APPROVAL
Not required
REFERENCES


42. Zhang J (2009) Model-Based Segmentation and Registration of Multimodal Medical Images, National University of Singapore


THE IMPACT OF MANUAL THRESHOLD SELECTION IN MEDICAL ADDITIVE MANUFACTURING

Maureen van Eijnatten, Juha Koivisto, Kalle Karhu, Tymour Forouzanfar, Jan Wolff

doi: 10.1007/s11548-016-1490-4
ABSTRACT

Purpose

Medical additive manufacturing requires standard tessellation language (STL) models. Such models are commonly derived from computed tomography (CT) images using thresholding. Threshold selection can be performed manually or automatically. The aim of this study was to assess the impact of manual and default threshold selection on the reliability and accuracy of skull STL models using different CT technologies.

Method

One female and one male human cadaver head were imaged using multi-detector row CT, dual-energy CT, and two cone-beam CT scanners. Four medical engineers manually thresholded the bony structures on all CT images. The lowest and highest selected mean threshold values and the default threshold value were used to generate skull STL models. Geometric variations between all manually thresholded STL models were calculated. Furthermore, in order to calculate the accuracy of the manually and default thresholded STL models, all STL models were superimposed on an optical scan of the dry female and male skulls (“gold standard”).

Results

The intra- and inter-observer variability of the manual threshold selection was good (intra-class correlation coefficients >0.9). All engineers selected grey values closer to soft tissue to compensate for bone voids. Geometric variations between the manually thresholded STL models were 0.13 mm (multi-detector row CT), 0.59 mm (dual-energy CT), and 0.55 mm (cone-beam CT). All STL models demonstrated inaccuracies ranging from -0.8 mm to +1.1 mm (multi-detector row CT), -0.7 mm to +2.0 mm (dual-energy CT), and -2.3 mm to +4.8 mm (cone-beam CT).

Conclusions

This study demonstrates that manual threshold selection results in better STL models than default thresholding. The use of dual-energy CT and cone-beam CT technology in its present form does not deliver reliable or accurate STL models for medical additive manufacturing. New approaches are required that are based on pattern recognition and machine learning algorithms.
INTRODUCTION

Additive manufacturing (AM), also known as three-dimensional (3D) printing, refers to a process where a series of successive layers are laid down to create a 3D construct. AM combined with advanced medical imaging technologies such as computed tomography (CT) and magnetic resonance imaging (MRI) has resulted in a paradigm shift in medicine from traditional serial production to patient-specific constructs. This combination of technologies offers new possibilities for the fabrication of implants, saw guides and drill guides that are designed to meet the specific anatomical needs of patients [1].

The three-step medical AM process begins with image acquisition (Figure 1, Step 1), which is commonly performed using a multi-detector row computed tomography (MDCT) scanner. However, dual-energy computed tomography (DECT), which offers the possibility of acquiring CT images using two different X-ray spectra, is becoming more common in hospital environments [2]. Furthermore, cone-beam computed tomography (CBCT) is being increasingly used in dentistry and maxillofacial surgery due to its low costs and reduced radiation dose when compared with MDCT scanners [3].

Images acquired using CT technologies are commonly saved as Digital Imaging and Communications in Medicine (DICOM) files. These files contain voxels with grey values that are proportional to the attenuation coefficient in the corresponding volume of the patient. In MDCT, these grey values are scaled according to Hounsfield units (HU): air (-1000 HU), water (0 HU), and compact bone (+1000 HU). In CBCT technology, the degree of x-ray attenuation is scaled using grey values, hence voxel values [4]. CBCT grey values are often arbitrary and do not correspond to MDCT HU values [3], [5], [6]. Furthermore, a large variability in the grey values has been reported between different CBCT scanners [7], [8].

At present, medical AM requires the conversion of DICOM images into virtual 3D surface models that are commonly saved as standard tessellation language (STL) files (Figure 1, Step 2). STL models are commonly used to design medical constructs using computer-aided design (CAD) software. The DICOM-to-STL conversion process requires the partitioning and hence the segmentation of voxels into different tissue types. The most

![Figure 1. A schematic diagram of the three steps required to fabricate an AM medical construct.](image-url)
common segmentation method used to date is thresholding. During the thresholding process, all voxels with a grey value that is equal or greater than a selected threshold value \( t \) are included in a segmented volume [9] using a binary mask \( M_{x,y} \) (Equation 1):

\[
M_{x,y} = \begin{cases} 
0 & \text{if } I_{x,y} < t \\
1 & \text{if } I_{x,y} \geq t
\end{cases}
\]

where \( I_{x,y} \) denotes the grey value at coordinates \( x \) and \( y \).

The medical image segmentation software packages available offer only a single, default threshold value for compact bone, soft tissue, and cartilage. However, these default values are often not optimized for all types of MDCT, DECT, and CBCT images and do not take into account the variations in grey values between different scanners [10]. Therefore, in most cases, manual threshold selection is necessary to acquire an optimal STL model. Threshold selection, however, still remains a subjective task [11], especially in the head area due to the plethora of complex bony geometries (Figure 2). Furthermore, minor dislocations in the facial area can have an impact on patient function and aesthetic appearance.

At present, there is a paucity of literature on threshold selection in the head area for medical purposes. Therefore, the aim of this study was to assess the impact of manual and default threshold selection on the reliability and accuracy of skull STL models acquired using different MDCT and CBCT technologies.

**MATERIALS AND METHODS**

One female and one male human cadaver head were anonymously provided by the Department of Anatomy, VU University Medical Center Amsterdam, The Netherlands. The two heads were embedded in a novel embalming liquid “Fix for Life” [12] that produces near life-like cadavers. Ethical approval for this study was provided by the Medical Ethical Committee of the VU University Medical Center (Ref. 2016.401).

![Figure 2. The effect of threshold selection on skull STL models.](image-url)
The two “Fix for Life” cadaver heads were imaged using the following CT technologies: GE Discovery CT750 HD 64-slice MDCT (GE Healthcare, Little Chalfont, Buckinghamshire, UK), NewTom 5G CBCT (NewTom, Verona, Italy), and Vatech PaX Zenith 3D CBCT (Vatech, Fort Lee, USA) (Figure 3, Step 1). The GE Discovery CT750 MDCT scanner was also operated using a dual-energy imaging mode (DECT). All scanners and image acquisition parameters are summarized in Table 1.

After CT image acquisition, all DICOM files were imported into Osirix® MD software (Osirix Foundation, Geneva, Switzerland). This software is FDA-cleared, CE-labeled for primary diagnostics, and is commonly used in medical AM. Osirix® MD software provides options for both manual and default threshold selection.

Four medical engineers were subsequently requested to manually select the optimal threshold value for bone in order to create an accurate STL model of the female and male skull, hence facial bony structures (Figure 3, Step 2). All four engineers were blinded for their own results and those of others. The manual threshold selection procedure was repeated after a six-week interval in order to determine the intra-observer variability and to calculate the mean threshold value. In addition, the inter-observer variability and intra-class correlation coefficients (ICC) were calculated using SPSS® software (SPSS® version 22, Chicago, Il, USA). ICC ranges between 0 and 1, with 0 corresponding to no agreement and 1 corresponding to complete agreement [13].

In order to graphically represent the distribution of grey values in the manually selected and default threshold values, histograms were plotted for each of the four CT scanners.
Figure 3. Outline of the study.

<table>
<thead>
<tr>
<th></th>
<th>GE Discovery CT750 HD 64-slice (MDCT)</th>
<th>GE Discovery CT750 HD 64-slice (DECT)</th>
<th>NewTom 5G (CBCT)</th>
<th>Vatech PaX Zenith 3D (CBCT)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>female</td>
<td>male</td>
<td>female</td>
<td>Male</td>
</tr>
<tr>
<td>Tube voltage (kV)</td>
<td>120</td>
<td>120</td>
<td>80,140</td>
<td>80,140</td>
</tr>
<tr>
<td>Tube current (mA)</td>
<td>300</td>
<td>300</td>
<td>375</td>
<td>375</td>
</tr>
<tr>
<td>Exposure time (s)</td>
<td>0.638</td>
<td>0.912</td>
<td>0.912</td>
<td>0.699</td>
</tr>
<tr>
<td>Spacing between slices (mm)</td>
<td>0.312</td>
<td>0.312</td>
<td>0.312</td>
<td>0.312</td>
</tr>
<tr>
<td>Slices thickness (mm)</td>
<td>0.625</td>
<td>0.625</td>
<td>0.625</td>
<td>0.625</td>
</tr>
<tr>
<td>Number of voxels</td>
<td>512 x 512</td>
<td>512 x 512</td>
<td>512 x 512</td>
<td>512 x 512</td>
</tr>
<tr>
<td></td>
<td>767</td>
<td>919</td>
<td>767</td>
<td>919</td>
</tr>
</tbody>
</table>
using MatLab® software (MatLab v.2012, MathWorks, Natick, Massachusetts, USA) (Figure 4). Only the highest and lowest mean selected threshold values presented on the eight histograms were used to generate STL models (Figure 3, Step 3). The generated STL models were subsequently geometrically compared to each other using GOM Inspect® software (GOM Inspect v8, GOM mbH, Braunschweig, Germany) in order to calculate the variations between the highest and lowest threshold STL models (Figure 3, Step 4).

In a final step, all soft tissues were manually removed from the cadaver heads using standard dissection equipment (i.e., scrapers and scalpels) by a highly experienced technician of the Department of Anatomy. Manual removal was opted for since this procedure ensured minimal dimensional changes in the bony structures of the cadaver skulls [14]. The resulting dry female and male skulls were subsequently scanned using a GOM ATOS™ III optical 3D scanner (GOM GmbH, Braunschweig, Germany) with an accuracy of < 0.05 mm to acquire a “gold standard” STL model of the skulls (Figure 3). These “gold standard” STL models were subsequently superimposed on the STL models generated using the highest and lowest manually selected and default threshold values in order to calculate the accuracy of each thresholded STL model (Figure 3, Step 5).

RESULTS

The intra- and inter-observer reliability results of all manually selected threshold values are presented in Table 2. All selected threshold values ranged from 113 HU to 303 HU for the MDCT and DECT technologies and from 537 gv to 1281 gv for the CBCT technologies (Figure 4 A - H). As shown in the histograms, all the selected threshold values differed from the default threshold value provided by Osirix MD® software (500 HU). Furthermore, the geometric variations between the highest and lowest thresholded STL models were larger in the STL models derived from DECT and CBCT when compared with the MDCT-derived STL models (Figure 5).

When compared to the “gold standard”, all manually and automatically thresholded STL models demonstrated inaccuracies ranging from -0.8 mm to +1.1 mm, -0.7 mm to +2.0 mm, and -2.3 mm to +4.8 mm for all STL models derived from MDCT, DECT, and CBCT, respectively (Figure 6 A - K). The male skull presented comparable accuracies to those observed on the female skull. The MDCT and DECT-derived STL models acquired using the default threshold value demonstrated the highest loss of bone-HU values (Figure 6 C, F). The NewTom CBCT-derived STL model acquired using the default threshold value (500 HU) provided by Osirix MD software resulted in an increase in artifacts and noise (Figure 6 I). The Vatech CBCT DICOM images did not allow the creation of an STL model using the 500-HU default threshold value since the grey values were not scaled to HU values (Figure 4 D,H).
Table 2. Intra- and inter-observer variability of manual threshold selection by four medical engineers on CT images of a female and a male cadaver head.

<table>
<thead>
<tr>
<th>Intra-observer variability</th>
<th>Inter-observer variability between the engineers</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Intra-class correlation coefficient (ICC)</strong></td>
<td><strong>ICC</strong></td>
</tr>
<tr>
<td>Cadaver head</td>
<td>female / male</td>
</tr>
<tr>
<td>Engineer 1</td>
<td>0.999 / 0.997</td>
</tr>
<tr>
<td>Engineer 2</td>
<td>0.995 / 0.995</td>
</tr>
<tr>
<td>Engineer 3</td>
<td>0.992 / 0.999</td>
</tr>
<tr>
<td>Engineer 4</td>
<td>0.969 / 0.989</td>
</tr>
</tbody>
</table>

Figure 4 (A to H). The mean threshold values (HU) selected by four medical engineers and the pre-defined default threshold value (500 HU) are presented in histograms A to H. The y-axis of the histograms (frequencies) are set to a logarithmic scale.
DISCUSSION
To date, thresholding is the most commonly used segmentation method in medical AM. However, accurate bone segmentation often requires manual threshold selection, which still remains a subjective task. Moreover, recent studies suggest that the majority of inaccuracies that occur during the medical AM process are introduced during the image acquisition and image processing phases, rather than during the manufacturing, i.e., the 3D printing process itself [15]–[17]. Such inaccuracies can markedly influence the resulting STL model (see Figure 6) and subsequently lead to ill-fitting AM implants [18]. Therefore, the aim of the present study was to assess the impact of manual and automatic default threshold selection on the reliability and accuracy of skull STL models.

In the present study, all threshold values selected by the four engineers demonstrated a good intra-observer reliability (ICC>0.9). Furthermore, the inter-observer reliability was also good (ICC>0.9), as shown in Table 2. Interestingly, all engineers that were blinded during the experiment selected threshold values for bone that were very close to the grey values of soft tissues (Figure 4). This resulted in small disjointed structures in the STL model (marked red in Figure 7) that represent the transition from bone into soft tissue grey values. Such disjointed “soft-tissue” structures can be manually removed during STL post-processing [19]. All engineers purposely selected the “soft tissue” threshold values during bone segmentation in order to incorporate the maximum number of bone-specific grey values. These grey values are allocated to voxels that represent different tissues during the CT image reconstruction process. However, during this process, voxels on the bone-to-soft tissue boundaries that are partially filled with soft tissue are commonly assigned a lower grey value than bone. This phenomenon is coined the partial volume effect (PVE).
Figure 6 (A - K). Accuracy of all STL models of the female skull acquired using the lowest (left) and highest (middle) mean threshold value selected by the four engineers and the default threshold value of 500 HU (right). The arrows indicate missing data (C, F) or excessive noise (I) in the default-threshold STL models.

[20]. As a consequence of the PVE, voxels may be erroneously allocated to “soft tissue” instead of “bone”, resulting in data loss and hence bone voids in the STL model (Figure 6). Therefore, engineers should be aware of this phenomenon since data loss can lead to large inaccuracies in individualized printed medical constructs [18], [20].

Another major finding in this study was the difference between the MDCT and CBCT DICOM files that were used to construct STL models (Figure 4). One explanation for this phenomenon is the inherent difference between these technologies. CBCT technology is
typically more heavily affected by image noise and distortions due to the “cone-beam” geometry of the X-ray beam [21], [22]. In CBCT, the simultaneously irradiated area is typically larger than in MDCT technology. This causes increased scatter levels and results in cupping, reduced contrast, and other scatter-induced artifacts in the reconstructed image. In addition, CBCT images are more subject to cone beam artifacts due to the large cone beam angle and the imaging geometry comprising a single focal plane. The cone beam artifacts result from violating Tuy’s sufficiency condition [23] that requires that each plane intersecting a region of interest must intersect the focal trajectory, i.e., the path defining the radiation source position during the imaging. The embodiments of cone beam artifacts are dependent on the reconstruction algorithm and the imaging geometry. Typical cone beam artifacts include the elongation of structures in the axial direction and negative undershoots at sharp edges in the transaxial planes [24]. In CBCT,
the focal trajectory consists of a single planar circle or arc that results in a violation of Tuy’s sufficiency condition in all regions outside the focal plane. The resulting cone beam artifacts are more pronounced the further away the region of interest is from the focal plane. In MDCT, the volume that satisfies Tuy’s sufficiency condition is notably larger due to the helical nature of the focal trajectory.

The presence of artifacts makes the segmentation and hence the thresholding of bone-specific grey values in CBCT images more cumbersome [25]. This subsequently leads to a larger variation in manually selected threshold values for CBCT images (Figure 4) and to the larger geometric variations of up to 0.55 mm in CBCT-derived STL models observed in this study (Figure 5). DECT-derived STL models demonstrated geometric variations of up to 0.59 mm (Figure 5). As a consequence of these geometric variations in STL models, the use of DECT and CBCT technology in its present form does not deliver reproducible STL models for medical AM. Therefore, the authors of this study suggest that only MDCT technology should be used for AM applications because of the lower variability (0.13 mm, see Figure 5) and higher accuracy (Figure 6) of the technology.

The present study demonstrates that the “human factor”, i.e., the medical engineer, influences the outcome of the segmentation process. Moreover, no single bone threshold value is applicable for all facial bones. The authors of this study therefore recommend the use of individual threshold values for each anatomical buttress. Recently, attempts have been made to develop novel segmentation algorithms using multi-thresholding [26], adaptive thresholding [11], and semi-automatic region growing [27]. However, these algorithms are still in an early stage of development [28] and do not take the inherent differences between MDCT and CBCT technologies into account. Future research should therefore focus on developing novel medical image segmentation software that is suitable for different CT imaging modalities. Furthermore, new approaches should be developed using pattern recognition and machine learning algorithms.

CONCLUSION
This study shows that manual threshold selection results in better skull STL models than default thresholding since all the medical engineers in our study selected grey values closer to soft tissue to compensate for bone voids. Our study also showed that MDCT-derived STL models offer the lowest variability and highest accuracy, whilst the use of DECT and CBCT technology in its present form does not deliver reliable STL models for medical AM. New approaches based on pattern recognition and machine learning algorithms are required.

STATEMENTS
Conflicts of interest
Author 2 Juha Koivisto and author 3 Kalle Karhu are currently employed by Planmeca Ltd (Finland), a company that specializes in the manufacture of cone-beam computed tomography scanners. The other authors declare that they have no conflict of interest.
Ethical approval
This article does not contain any studies with human participants or animals performed by any of the authors. All human cadaveric materials that were used in the present study (one female and one male head) were anonymously acquired through the body donor program of the Department of Anatomy of the VU University Medical Center Amsterdam, The Netherlands, in full accordance with Article 1 of the Dutch law on funeral services (http://wetten.overheid.nl/BWBR0005009/2015-07-01) and European legislation. Furthermore, ethical approval for this study was provided by the Medical Ethical Committee (METC) of the VU University Medical Center (Ref. 2016.401).

Informed consent
For this type of study, no formal consent was required.
REFERENCES


ACCURACY OF MDCT AND CBCT IN THREE-DIMENSIONAL EVALUATION OF THE UPPER AIRWAY MORPHOLOGY

Maureen van Eijnatten*, Hui Chen*, Ghizlane Aarab, Tymour Forouzanfar, Jan de Lange, Paul van der Stelt, Frank Lobbezoo, Jan Wolff

* Both authors contributed equally to this work

doi: 10.1093/ejo/cjx030
ABSTRACT

Objectives
To assess the accuracy of five different computed tomography (CT) scanners for the evaluation of the oropharynx morphology.

Methods
An existing cone-beam computed-tomography (CBCT) dataset was used to fabricate an anthropomorphic phantom of the upper airway volume that extended from the uvula to the epiglottis (oropharynx) with known dimensions (gold standard). This phantom was scanned using two multi-detector row computed-tomography (MDCT) scanners (GE Discovery CT750 HD, Siemens Somatom Sensation) and three CBCT scanners (NewTom 5G, 3D Accuitomo 170, Vatech PaX Zenith 3D). All CT images were segmented by two observers and converted into standard tessellation language (STL) models. The volume and the cross-sectional area of the oropharynx were measured on the acquired STL models. Finally, all STL models were registered and compared with the gold standard.

Results
The intra- and inter-observer reliability of the oropharynx segmentation was fair to excellent. The most accurate volume measurements were acquired using the Siemens MDCT (98.4%; 14.3 cm$^3$) and Vatech CBCT (98.9%; 14.4 cm$^3$) scanners. The GE MDCT, NewTom 5G CBCT and Accuitomo CBCT scanners resulted in smaller volumes, viz., 92.1% (13.4 cm$^3$), 91.5% (13.3 cm$^3$), and 94.6% (13.8 cm$^3$), respectively. The most accurate cross-sectional area measurements were acquired using the Siemens MDCT (94.6%; 282.4 mm$^2$), Accuitomo CBCT 95.1% (283.8 mm$^2$), and Vatech CBCT 95.3% (284.5 mm$^2$) scanners. The GE MDCT and NewTom 5G CBCT scanners resulted in smaller areas, viz., 89.3% (266.5 mm$^2$) and 89.8% (268.0 mm$^2$), respectively.

Limitations
Images of the phantom were acquired using the vendor-supplied default airway scanning protocol for each scanner.

Conclusion
Significant differences were observed in the volume and cross-sectional area measurements of the oropharynx acquired using different MDCT and CBCT scanners. The Siemens MDCT and the Vatech CBCT scanners were more accurate than the GE MDCT, NewTom 5G, and Accuitomo CBCT scanners. In clinical settings, CBCT scanners offer an alternative to MDCT scanners in the assessment of the oropharynx morphology.
INTRODUCTION

Obstructive sleep apnea (OSA) is a sleep-related breathing disorder, often associated with a compromised upper-airway space and an increase in upper-airway collapsibility [1]. The most common complaints of OSA patients are excessive daytime sleepiness, unrefreshing sleep, poor concentration, and fatigue [2]. OSA also has a range of deleterious consequences that include increased cardiovascular morbidity, neurocognitive impairment, and overall mortality [3]–[6]. An important role in the pathogenesis of OSA is played by anatomical and functional abnormalities of the upper airway [7].

The three-dimensional (3D) morphology of the upper airway is currently assessed using computed tomography (CT) technologies [8]. The two most common CT technologies used to date for the assessment of the upper airway are multi-detector row computed tomography (MDCT) [9] and cone-beam computed tomography (CBCT) [10]. The major advantages of CBCT scanners are their lower radiation dose and costs [11], [12]. As a result, CBCT scanners are being increasingly used for upper-airway imaging in OSA patients [13]–[15]. CBCT scanners use a single, partial gantry rotation [16], which not only accounts for lower radiation dose, but also produces acceptable diagnostic image quality [17]. However, it remains unclear whether MDCT and CBCT scanners can provide accurate 3D images of the upper airway.

According to a systematic review by Alsufyani et al. [18], only one out of sixteen studies focusing on the use of CBCT to automatically or semi-automatically model the upper airway had a sufficiently sound methodology to test the accuracy of the upper airway dimensions. The main challenge faced in the assessment of the upper airway accuracy using MDCT or CBCT is the lack of a “gold standard” model with known dimensions. Recent studies have used artificial models, hence phantoms of the upper airway, as a gold standard [19]–[21]. However, commercially available phantoms used in most of the aforementioned studies are commonly manufactured in simple, generic forms and sizes and therefore do not resemble the clinical situation. In the present study, a novel 3D printed anthropomorphic phantom of the upper airway volume (oropharynx) with known dimensions was manufactured that closely resembled a real patient in terms of size, shape, structure, and attenuation profiles.

The aim of this study was to assess the accuracy of two different MDCT scanners and three different CBCT scanners using a novel 3D printed anthropomorphic phantom for the evaluation of the oropharynx morphology.

MATERIALS AND METHODS

A CBCT dataset of a 27-year-old female that had been previously acquired using a NewTom 5G CBCT scanner (QR systems, Verona, Italy), was used to design and 3D print an anthropomorphic phantom of the airway space (Figure 1 A and B). The aforementioned CBCT dataset was converted into a virtual 3D surface, hence standard tessellation language (STL) model of the upper airway volume that extended from the uvula to the epiglottis: the oropharynx (Figure 3). This STL model served as the gold standard
in this study. The gold standard STL model of the oropharynx was subsequently used to manufacture the phantom. All bony structures surrounding the oropharynx were 3D printed using a High Performance Composite powder ZP151 (3D Systems, Rock Hill, USA). This composite material was chosen due to its bone-like density that resembles the attenuation profile of bone [22]. The soft tissue surrounding the oropharynx was fabricated using soft-tissue-equivalent silicon (Dragon Skin 30, Smooth-On, Inc., Macungie, Pennsylvania, USA). During the assembling of the phantom, three metal markers were positioned in a defined plane to acquire a reproducible reference-point system (RPS) for the cross-sectional area measurement of the oropharynx (Figure 1).

As Figure 2 shows, MDCT images of the oropharynx phantom were acquired using two MDCT scanners: GE Discovery CT750 HD 64-slice MDCT (GE Healthcare, Little Chalfont, Buckinghamshire, UK) and Siemens Somatom Sensation 64-slice MDCT (Siemens Medical

---

**Figure 1.** A. Design of the oropharynx of the phantom (red = maxilla and mandible; green = cervical vertebrae; yellow = supports of the markers; black = markers at the level of the minimum cross-sectional area of the oropharynx; blue = upper airway; purple = base plane; grey and pale-yellow = mould of the skin). B. The 3D printed phantom. C. Sagittal image of the phantom using GE scanner. Arrow = a marker. D. Segmentation based on GE images (purple = oropharynx; green = base plane of the phantom).
Solutions, Malvern PA, USA). Furthermore, the following three CBCT scanners were used to acquire CBCT images of the phantom: NewTom 5G CBCT (QR systems, Verona, Italy), 3D Accuitomo 170 CBCT (J Morita, Kyoto, Japan), and Vatech PaX Zenith 3D CBCT (Vatech, Fort Lee, USA). All MDCT and CBCT images were acquired using the vendor-supplied default airway scanning protocol. All imaging parameters of the five CT scanners are presented in Table 1.

The acquired CT datasets were saved as Digital Imaging and Communications in Medicine (DICOM) files and were imported into Amira® software (v4.1, Visage Imaging Inc., Carlsbad, CA, USA) (Figure 2C). Using thresholding, one maxillofacial radiologist and one orthodontist then segmented all the acquired DICOM datasets of the oropharynx

---

**Figure 2.** Flowchart of this study.

**Table 1.** Image acquisition parameters for MDCT and CBCT scans.

<table>
<thead>
<tr>
<th></th>
<th>GE MDCT</th>
<th>Siemens MDCT</th>
<th>NewTom 5G CBCT</th>
<th>Accuitomo CBCT</th>
<th>Vatech CBCT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tube voltage (kV)</td>
<td>120</td>
<td>120</td>
<td>110</td>
<td>90</td>
<td>115</td>
</tr>
<tr>
<td>Tube current (mA)</td>
<td>103</td>
<td>57</td>
<td>5.8</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Scan time (s)</td>
<td>0.5</td>
<td>0.5</td>
<td>3.6</td>
<td>17.5</td>
<td>24</td>
</tr>
<tr>
<td>Slices thickness (mm)</td>
<td>0.625</td>
<td>0.600</td>
<td>0.300</td>
<td>1</td>
<td>0.200</td>
</tr>
<tr>
<td>Number of voxels</td>
<td>512 x 512 x 180</td>
<td>512 x 512</td>
<td>610 x 610 x 539</td>
<td>684 x 684 x 480</td>
<td>800 x 800 x 632</td>
</tr>
<tr>
<td>Reconstruction</td>
<td>Soft</td>
<td>B40f</td>
<td>Standard</td>
<td>N.A.</td>
<td>N.A.</td>
</tr>
<tr>
<td>DLP (mGy-cm)</td>
<td>46.49</td>
<td>49</td>
<td>N.A.</td>
<td>N.A.</td>
<td>N.A.</td>
</tr>
<tr>
<td>DAP (mGy-cm²)</td>
<td>N.A.</td>
<td>N.A.</td>
<td>12.232</td>
<td>N.A.</td>
<td>17.67</td>
</tr>
<tr>
<td>CTDIvol (mGy)</td>
<td>N.A.</td>
<td>N.A.</td>
<td>N.A.</td>
<td>8.70</td>
<td>N.A.</td>
</tr>
</tbody>
</table>

CBCT: cone beam computed tomography; CTDI: computed tomography dose index; DAP: dose-area product; DLP: dose-length product; Gy: Gray; MDCT: multi-detector row computed tomography; N.A.: not applicable.
(Table 2). Both observers were blinded for their own results and those of each other. The segmentation procedure was performed five times for each CT scanner and was subsequently repeated after a ten-day interval. This resulted in a total of 20 threshold values per CT scanner (Table 2). These 20 threshold values per scanner were subsequently used to segment the oropharynx. All segmented oropharynx volumes acquired from the five MDCT and CBCT scanners (Figure 2D) were converted into STL models. These STL models were subsequently imported into GOM Inspect v8® metrology software (GOM, Braunschweig, Germany) for further airway analysis.

The region of interest (ROI) in this study was the volume of the oropharynx between two parallel planes located 1.5 mm and 42 mm above the base plane of the phantom (Figure 3). In order to obtain comparable oropharynx volume measurements, all STL models derived from the five CT datasets were cropped accordingly. The volume of the oropharynx was subsequently calculated using GOM Inspect software (Figure 4A). Furthermore, the cross-sectional area of the oropharynx at the level of the metal markers in the phantom was calculated (Figure 4B).

To determine the accuracy of the CT-derived STL models, all acquired STL models were superimposed onto the gold standard STL model of the oropharynx using a verified surface registration (local best-fit) algorithm in GOM Inspect software with an accuracy of 0.05 mm [23]. All geometric deviations between the oropharynx STL models and the printed gold standard phantom STL model are depicted in Figure 5.

![Figure 3. The segmented volume of the oropharynx. (Red line = upper boundary of region of interest (ROI); yellow line = lower boundary of ROI; green line = the location of the markers)](image-url)
Finally, statistical analysis was performed, using IBM Statistical Package for Social Sciences for Windows (SPSS® version 21, Chicago, Il, USA). Statistical significance was set at $\alpha=0.05$. To determine the intra- and inter-observer reliability of the oropharynx measurements, intraclass correlation coefficients (ICCs) were calculated. Reliability was divided into three categories: poor (ICC<0.40); fair to good (0.40≤ICC≤0.75); excellent (ICC>0.75) [24]. The accuracy of the scanners was calculated as the ratio of the phantom measurements to the gold standard (%). One-way analysis of variance (ANOVA) was used to test the differences in the volume and cross-sectional area measurements between the five different CT scanners. Post-hoc analysis (Tukey’s honest significant difference test) was run to establish which CT scanners produced significantly different results.

RESULTS

All threshold values used for the segmentation of the oropharynx are shown in table 2. Intra- and inter-observer reliability hence ICCs of the threshold values ranged from 0.436 (fair to good) to 0.966 (excellent).

There were significant differences between the volume measurements of the oropharynx STL models acquired using the five different CT scanners (F=84.21; P=0.00) (Figure 4A). Tukey’s test showed that there were no significant differences in volume measurements between the Siemens MDCT and Vatech CBCT scanners, and between the GE MDCT and NewTom 5G CBCT scanners. The Siemens MDCT and Vatech CBCT scanners provided the most accurate volume measurements of the oropharynx (Figure 4A). The NewTom 5G CBCT, Accuitomo CBCT, and the GE MDCT scanners resulted in smaller volume measurements of the oropharynx (Figure 4A, Table 3).

There were also significant differences between the cross-sectional area measurements of the oropharynx STL models acquired using the five different CT scanners (F=43.11; P=0.00) (Figure 4B). The Siemens MDCT, Vatech CBCT and Accuitomo CBCT scanners

Table 2. Intraobserver and interobserver reliability of the threshold values (Hounsfield Units) chosen for five CT scanners estimated by intraclass correlation coefficients (ICCs), based on 20 measurements in total per scanner.

<table>
<thead>
<tr>
<th>Scanner</th>
<th>Threshold values (HU) (Intra-observer reliability)</th>
<th>Threshold values (HU) (Inter-observer reliability)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Observer 1 Mean±SD (ICC)</td>
<td>Observer 2 Mean±SD (ICC)</td>
</tr>
<tr>
<td>GE (MDCT)</td>
<td>-339±47 (.966)</td>
<td>-389±87 (.864)</td>
</tr>
<tr>
<td>Siemens (MDCT)</td>
<td>-204±54 (.573)</td>
<td>-250±77 (.748)</td>
</tr>
<tr>
<td>NewTom5G (CBCT)</td>
<td>-114±35 (.481)</td>
<td>-102±40 (.720)</td>
</tr>
<tr>
<td>Accuitomo (CBCT)</td>
<td>-289±40 (.661)</td>
<td>-244±62 (.438)</td>
</tr>
<tr>
<td>Vatech (CBCT)</td>
<td>-361±75 (.787)</td>
<td>-330±63 (.436)</td>
</tr>
</tbody>
</table>
Figure 4. A. Mean and standard deviation of the volume of the oropharynx derived from five CT scanners. GS = gold standard; NS = no significant difference. B. Mean and standard deviation of the cross-sectional area of the oropharynx derived from five CT scanners. GS = gold standard; NS = no significant difference. The accuracy (%) was calculated as the ratio between the phantom measurements and the gold standard values.

Table 3. Mean and standard deviation of the volume and the cross-sectional area of the upper airway derived from five CT scanners. GS = gold standard.

<table>
<thead>
<tr>
<th>Variable</th>
<th>GS</th>
<th>GE MDCT</th>
<th>Siemens MDCT</th>
<th>NewTom 5G CBCT</th>
<th>Accuitomo CBCT</th>
<th>Vatech CBCT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volume of the upper airway (cm$^3$)</td>
<td>14.5</td>
<td>13.4 ± 0.32</td>
<td>14.3 ± 0.31</td>
<td>13.3 ± 0.15</td>
<td>13.8 ± 0.21</td>
<td>14.4 ± 0.19</td>
</tr>
<tr>
<td>Area of the upper airway (mm$^2$)</td>
<td>295.5</td>
<td>266.5 ± 6.1</td>
<td>282.4 ± 6.8</td>
<td>268.0 ± 3.8</td>
<td>283.8 ± 7.9</td>
<td>284.5 ± 5.2</td>
</tr>
</tbody>
</table>

provided the most accurate cross-sectional area measurements of the oropharynx (Figure 4B). The GE MDCT and NewTom 5G CBCT scanners resulted in smaller area measurements of the oropharynx (Figure 4B, Table 3).

Figure 5 shows the oropharynx STL models acquired using five different CT scanners. The largest geometric deviations were observed in the vicinity of the uvula and the epiglottis region (Figure 5).
DISCUSSION

3D evaluation of the oropharynx offers new possibilities of assessing anatomical abnormalities in OSA patients. In this study, significant differences ($P < 0.001$) were found between the volume and cross-sectional area measurements of the oropharynx acquired using different MDCT and CBCT scanners (Figure 4, Figure 5).

The most accurate volume measurements of the oropharynx were acquired using the Siemens MDCT (98.4%; 14.3 cm$^3$) and Vatech CBCT (98.9%; 14.4 cm$^3$) scanners.

Figure 5. The results of registration between all CT-derived oropharynx STL models (acquired using the mean threshold value) and the gold standard STL model. A = GE; B = Siemens; C = NewTom 5G; D = Accuitomo; E = Vatech; F = Gold standard STL model. Scale: Red = CT-derived STL model is larger than the gold standard; blue = CT-derived STL model is smaller than the gold standard; green = CT-derived STL model is close to the gold standard; black and white = outliers (> 0.8 mm).
The GE MDCT, NewTom 5G CBCT and Accuitomo CBCT scanners resulted in smaller volume measurements, viz., 92.1% (13.4 cm³), 91.5% (13.3 cm³), and 94.6% (13.8 cm³), respectively. The most accurate cross-sectional area measurements of the oropharynx were acquired using the Siemens MDCT (94.6%; 282.4 mm²), Accuitomo CBCT (95.1%; 283.8 mm²) and Vatech CBCT (95.3%; 284.5 mm²) scanners (Figure 4 B). The GE MDCT and NewTom 5G CBCT scanners resulted in smaller area measurements, viz., 89.3% (266.5 mm²) and 89.8% (268.0 mm²), respectively. These results are in good agreement with previous studies that reported an underestimation of the airway area in both MDCT and CBCT images [25], [26]. However, it should be noted that the absolute values of the aforementioned inaccuracies ranged between 13.3 cm³ and 14.4 cm³ (volume), and between 266.5 mm² and 284.5 mm² (cross-sectional area) (Table 3). Consequently, the authors of this study hypothesize that the reported inaccuracies should not affect the radiological evaluation of OSA patients in clinical settings [27].

Even though the authors of this study used the recommended soft tissue imaging protocols on all CT scanners, the STL models acquired using the GE and NewTom 5G scanners were generally smaller in size than the gold standard STL model (Figure 4). The smaller STL models caused a shift in the histograms towards the negative direction (Figure 5 A, C). This phenomenon is probably due to the partial volume effect [28], in which voxels in the vicinity of the air-to-soft tissue boundary are commonly allocated to “soft tissue” instead of “air” during the segmentation process.

The higher accuracy of the Vatech STL models could be a result of the smaller spatial resolution used in the default airway scanning protocol (Table 1) [29]. However, a very recent study by Sang et al. [30] that investigated the influence of voxel size on the accuracy of NewTom 5G and Vatech CBCT reported that increasing the voxel resolution from 0.30 to 0.15 mm does not always result in increased accuracy of 3D tooth reconstructions, while different CBCT modalities (i.e. NewTom 5G vs. Vatech) can significantly affect the accuracy.

The largest geometric deviations were found in the uvula and epiglottis area (Figure 5). Interestingly, the acquired oropharynx STL models were generally too large in the epiglottis region and too small in the vicinity of the uvula. One explanation for this phenomenon could be that the epiglottis has a concave-like geometry and the uvula is convex. These findings are in good agreement with a previous study by Barone et al. [31] who observed discrepancies between the segmentation of concave and convex shapes in teeth.

The results of the present study show that CBCT scanners offer an alternative to MDCT scanners in the assessment of the oropharynx. This is in good agreement with a previous study by Suomalainen et al. [32], who reported that CBCT scanners offer images similar to those acquired using low-dose MDCT protocols. Therefore, taking the lower CBCT radiation dose into consideration [11], [12], clinicians should preferably use CBCT modalities for the analysis of the oropharynx. Moreover, all appropriate measures should be undertaken to minimize the dose according to the International Commission on Radiological Protection (ICRP) and the As Low As Reasonably Achievable (ALARA) principles [33].
Limitations of the current study

One limitation of this study was that for each of the five CT scanners, only one image dataset was acquired using a single image acquisition protocol. Since there are multiple image acquisition protocols available for each MDCT and CBCT scanner, different protocols should be considered in a future study. Another limitation was that the gold standard values in the present study were obtained from the original STL model of the oropharynx that was used to 3D print the phantom. 3D printing was performed using a Zprinter 250 inkjet powder printer (3D Systems, Rock Hill, USA), which has a layer thickness of 0.1 mm and an in-plane resolution of approximately 0.05 mm [23]. Therefore, this process may have introduced a manufacturing error, hence measurement uncertainty, of up to 0.2 mm [34]. Nevertheless, this uncertainty can be considered clinically insignificant [35].

CONCLUSION

Significant differences were observed in the volume and cross-sectional area measurements of the oropharynx acquired using different MDCT and CBCT scanners. The Siemens MDCT and the Vatech CBCT scanners were more accurate than the GE MDCT, NewTom 5G, and Accuitomo CBCT scanners. In clinical settings, CBCT scanners offer an alternative to MDCT scanners in the assessment of the oropharynx morphology.
REFERENCES


diagnostic task for cone beam computed tomography with different fields of view. Eur. J. Radiol. 80:483–488


THE ACCURACY OF ULTRASHORT ECHO TIME MRI SEQUENCES FOR MEDICAL ADDITIVE MANUFACTURING

Maureen van Eijnatten, Erik-Jan Rijkhorst, Mark Hofman, Tymour Forouzanfar, Jan Wolff

doi: 10.1259/dmfr.20150424
ABSTRACT

Objectives
Additively manufactured bone models, implants and drill guides are becoming increasingly popular amongst maxillofacial surgeons and dentists. To date, such constructs are commonly manufactured using computed tomography (CT) technology that induces ionizing radiation. Recently, ultrashort echo time (UTE) magnetic resonance imaging (MRI) sequences have been developed that allow radiation-free imaging of facial bones. The aim of the present study was to assess the feasibility of UTE MRI sequences for medical additive manufacturing (AM).

Methods
Three morphologically different dry human mandibles were scanned using a CT and MRI scanner. Additionally, optical scans of all three mandibles were made to acquire a “gold standard”. All CT and MRI scans were converted into Standard Tessellation Language (STL) models and geometrically compared to the gold standard. To quantify the accuracy of the AM process, the CT, MRI and gold standard STL models of one of the mandibles were additively manufactured, optically scanned and compared to the original gold standard STL model.

Results
Geometric differences between all three CT-derived STL models and the gold standard were less than 1.0 mm. All three MRI-derived STL models generally presented deviations below 1.5 mm in the symphyseal and mandibular area. The AM process introduced minor deviations of less than 0.5 mm.

Conclusions
This study demonstrates that MR imaging using UTE sequences is a feasible alternative to CT in generating STL models of the mandible, and would therefore be suitable for surgical planning and AM. Further in-vivo studies are necessary to assess the usability of UTE MRI sequences in clinical settings.
INTRODUCTION

Virtual three-dimensional (3D) surgical planning and additive manufacturing (AM) technologies are increasingly used in oral and maxillofacial surgery. To date, these advanced technologies have proven to be invaluable in dental implant surgery [1], in the restoration of mandibular fractures [2], in tumour resections [3] and in maxillofacial reconstructions [4], [5]. Moreover, additively manufactured medical constructs such as drill guides, [6] saw guides [7] or individualized implants [8] offer a predictive, functional and aesthetic outcome, especially in patients with anatomical limitations, insufficient mandibular or maxillary bone or poor bone density [9].

Medical AM comprises three basic steps. The first step is the acquisition of high-resolution medical 3D images that are archived as a digital imaging and communications in medicine (DICOM) file. The second step is to load the DICOM file into virtual planning software that commonly converts the DICOM file automatically into a virtual 3D surface model in the Standard Tessellation Language (STL) file format [10]. The resulting STL model can then be used to design a medical construct using computer aided design (CAD) software. The last step is the conversion of the STL model into a g-code that is required to control the AM process.

Currently, 3D medical images intended for oral and maxillofacial AM are acquired using multi-detector row computed tomography (MDCT) or cone beam computed tomography (CBCT) technologies [11]. Both MDCT and CBCT provide high resolution images of the maxillary and mandibular bone and the surrounding tissues [12]. However, the aforementioned technologies have one major disadvantage: they expose the patient to ionizing radiation. The dose levels range between 20 and 400 µSv for CBCT and around 1000 µSv for MDCT modalities, which is about 10 to 50 times more than a conventional panoramic radiograph (about 20 µSv) [13], [14]. Even though device manufacturers have been developing “ultra” low-dose CT scanning protocols [15]–[17], any X-ray-based imaging technology inevitably results in some radiation exposure. Therefore, an alternative, radiation-free imaging modality is still being sought, especially in applications that require a large number of follow-up scans [18].

Over the past decade, MRI technology has been discussed as an alternative to CT technology for implant planning [9], [19], [20]. To date, most conventional MRI sequences only offer optimal soft tissue images and lack the ability to image bone/air interfaces in the head area. This is because MR image acquisition is based on water microenvironments and other sources of protons in the human body. Since cortical bone has a low proton density and a very short T2 relaxation time of about 1.5 ms [21], conventional spin echo and gradient echo MRI sequences are too slow to acquire a sufficient signal from bone. As a result, new ultrashort echo time (UTE) sequences have been recently developed and successfully used for bone imaging [22], [23]. The unique feature of these novel UTE sequences is that they sample the MR signal with a minimum echo time due to radial sampling, and rapidly switch to data recording after the radiofrequency (RF) pulse has been delivered [23].
The MRI of bone using UTE sequences is being applied in several emerging clinical technologies such as PET-MRI attenuation correction [22], high-intensity focused ultrasound (HIFU) applications [24], MR-guided surgery [25] and MR-only treatment planning and guidance in radiotherapy [26]. However, to the best of our knowledge no studies have been performed on the feasibility of UTE MRI sequences in generating STL models for AM purposes. In this study, we compare the geometric accuracy of UTE MRI-derived STL models of three morphologically different human mandibles with CT-derived STL models. Furthermore, additional geometric inaccuracies introduced during the AM process are assessed.

MATERIALS AND METHODS

Three morphologically different human cadaver mandibles of succumbed Dutch patients with unknown clinical histories and intact bony structures were obtained from the Department of Anatomy, VU University Medical Center, Amsterdam, The Netherlands. The mandibles were boiled for 20 hours and all soft tissues were meticulously removed manually.

The outline of this study is summarized in Figure 1. The three dry mandibles were scanned using a CT and an MRI scanner (Step 1). In addition, an optical scan of all three mandibles was acquired using an optical 3D scanner (Artec Spider™, Artec group, Moscow, Russia) with a point accuracy of 0.05 mm. The optical scans of the three mandibles were used as gold standard STL models. The CT and MRI scans were then converted to STL files (Step 2) and aligned with the gold standard STL models. Subsequently, the geometric deviations between the CT and MRI-derived STL models and the gold standard models were calculated. Furthermore, the CT, MRI and gold standard STL model of one mandible were additively manufactured (Step 3) and scanned again with the optical scanner (Step 4) in order to map additional geometric deviations introduced during the AM process.

MRI

The mandibular cortical bone has a very short $T_2^*$ of approximately 1.5 ms, and therefore conventional spin or gradient echo sequences are too slow to acquire any bone signal. In UTE imaging, the free induction decay signal is sampled directly with a sub-millisecond echo time, resulting in a MR signal from the surrounding cortical bone [22], [23].

All three mandibles were scanned using an UTE sequence on a Philips Achieva 3T MRI scanner (Philips Healthcare, Best, the Netherlands) with an eight-channel head radiofrequency coil using a UTE single echo sequence. The following scan parameters were used: repetition time $= 4.8$ ms, echo time $= 0.14$ ms and flip angle $= 15^\circ$. A non-selective hard radiofrequency pulse was used for excitation, followed by tuning the receive coil. Immediately thereafter, free induction decay sampling was started during ramp-up of the gradients. This procedure resulted in a time between excitation and readout of 0.14 ms. The 3D k-space was radially sampled and regridded to a Cartesian coordinate system of isotropic 0.5 x 0.5 x 0.5 mm voxels.
CT imaging

CT imaging was performed with a Siemens Somatom Sensation 64-slice CT scanner (Siemens Medical Solutions, Erlangen, Germany) using a low-dose scanning protocol. The scan parameters were as follows: tube voltage = 80 kV, tube current exposure time product = 150 mAs, pixel size = 0.3 mm x 0.3 mm, slice thickness = 0.6 mm. The images were then reconstructed using a sharp bone convolution kernel (H60h).

Image processing and STL deviation analysis

All MRI and CT datasets were saved in DICOM file formats and subjected to the image processing steps described in Figure 2. All CT datasets were imported into Insight Segmentation and Registration Toolkit® software (Kitware Inc., Clifton Park, NY) and manually thresholded using 0 Hounsfield Units (HU). The MRI datasets were thresholded using 15 arbitrary units (AU) (mandible 1 and 2), and 12 AU (mandible 3) (Figure 2, Step 1). These threshold values were chosen just above the MRI background noise signal level to acquire an optimal model of the mandibular bone. Subsequently, all segmented datasets were converted to STL models (Figure 2, Step 2) and imported into MeshLab software (Visual Computing Lab, Pisa, Italy). The MRI and CT-derived STL models were aligned with their corresponding gold standard STL models using an iterative closest point algorithm (Figure 2, Step 3). The signed STL-to-STL distances, that is geometric deviations, between all MRI and CT-derived STL models and the gold standard STL model were computed.
The accuracy of ultrashort echo time MRI sequences

using Visualization Toolkit® software (Kitware Inc.) (Figure 2, Step 4) and visualised using colour maps and histograms (Figure 3).

The 95th percentiles of all STL-to-STL distances were calculated using MatLab software v.2012, (MathWorks®, Natick, MA). To determine the geometric accuracy in the different areas of interest for surgical planning, the 95th percentiles were calculated for the whole mandible as well as for five different anatomical regions: 1) right ramus with condyle; 2) right body; 3) symphyseal area with alveolar ridge; 4) left body; 5) left ramus with condyle (Table 1).

Additive manufacturing
The CT and MRI-derived STL models of mandible 1 were imported into GOM Inspect® software (GOM GmbH, Braunschweig, Germany) and smoothed with a surface tolerance of 1.0 mm to correct stair-step artefacts. The smoothed CT-derived STL model, the smoothed MRI-derived STL model and the gold standard STL model of mandible 1 obtained using the optical scanner were all printed using a Zprinter 250 inkjet powder printer (3D Systems Inc., Rock Hill, SC). The resulting additively manufactured models were once again scanned with an optical scanner, aligned with the gold standard STL model, and subjected to STL deviation analysis according to Steps 3 and 4 in Figure 2.

RESULTS
All geometric deviations between the optical scan “gold standard” STL models and the CT-derived STL models were below 1.0 mm (Figure 3 A-C), with 95th percentiles < 0.5 mm (Table 1), with the exception of Mandible 1 (0.71 mm). The MRI-derived STL models generally presented geometric deviations < 2.0 mm (Figure 3 D-F), with 95th percentiles < 1.5 mm (Table 1), again except for Mandible 1 (1.69 mm).

Figure 4 presents the mean and standard deviation (SD) of the signed geometric deviations in the five different anatomical regions of all CT and MRI STL models. The mean value indicates to what extent the STL model was smaller or larger than the gold standard, and the SD indicates the spread in geometric deviations between the STL model and...
the gold standard. The deviations in all CT-derived STL models were uniformly distributed over the different anatomical regions, whereas all MRI-derived STL models showed a small spread of deviations in the symphyseal area and a larger spread of deviations in the mandibular condyles (Figure 4).

Minor additional geometric deviations were reported in the additively manufactured 3D models of Mandible 1 (Figure 5). The additively manufactured gold standard 3D model demonstrated geometric deviations of up to 0.5 mm (Table 2), especially in thin structures such as the coronoid process and alveolar ridge. Geometric deviations in the additively manufactured 3D model obtained using CT data (Figure 5 B) were mostly in the vicinity of the coronoid process. In the 3D model obtained using MRI data (Figure 5 C), deviations were observed in the coronoid process, mandibular condyles and alveolar ridge.

Figure 3. (A-F) Deviations between the CT-derived Standard Tessellation Language (STL) models and the gold-standard STL model (a–c), and the MRI-derived STL models and the gold-standard STL model (d–f), shown in colour maps and histograms.
Figure 4. Mean ± standard deviation of the geometric deviations of five anatomical regions of the mandible in CT- and MRI-derived Standard Tessellation Language models. Std., standard.

Table 1. The 95th percentiles of the geometric deviations of five anatomical regions of the mandible in CT and MRI-derived Standard Tessellation Language models.

<table>
<thead>
<tr>
<th>Region</th>
<th>CT (mm)</th>
<th>MRI (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole mandible</td>
<td>0.71</td>
<td>0.47</td>
</tr>
<tr>
<td>Right ramus with condyle</td>
<td>0.49</td>
<td>0.33</td>
</tr>
<tr>
<td>Right body</td>
<td>0.30</td>
<td>0.45</td>
</tr>
<tr>
<td>Symphyseal area</td>
<td>0.45</td>
<td>0.56</td>
</tr>
<tr>
<td>Left body</td>
<td>0.62</td>
<td>0.56</td>
</tr>
<tr>
<td>Left ramus with condyle</td>
<td>0.93</td>
<td>0.45</td>
</tr>
</tbody>
</table>

Table 2. The 50th, 95th and 99th percentiles of the geometric deviations in additively manufactured three-dimensional models of Mandible 1.

<table>
<thead>
<tr>
<th>Percentile</th>
<th>Gold standard (mm)</th>
<th>CT (mm)</th>
<th>MRI (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>50th</td>
<td>0.11</td>
<td>0.21</td>
<td>0.48</td>
</tr>
<tr>
<td>95th</td>
<td>0.46</td>
<td>0.88</td>
<td>1.73</td>
</tr>
<tr>
<td>99th</td>
<td>1.12</td>
<td>1.32</td>
<td>2.63</td>
</tr>
</tbody>
</table>
Figure 5. (A-C) Deviations between the additively manufactured (AM) gold-standard three-dimensional (3D) model and the gold-standard Standard Tessellation Language (STL) model; the AM CT-derived 3D model and the gold-standard STL model; and the AM MRI-derived 3D model and the gold-standard STL model shown in colour maps and histograms.
DISCUSSION

Today, there is an ever increasing need for high-resolution 3D images for use in surgical planning [27], [28]. In this context, CT and CBCT imaging technologies are commonly used and subsequently induce harmful radiation to the patient [29]. The cumulative lifetime radiation dose of the population needs to be minimized by reducing the number of overall examinations and the dose resulting from each individual exposure [30]. Hence, radiation-free imaging modalities are still being sought for medical additive manufacturing.

MR and CT imaging

The novel UTE MRI sequence used in this study produced STL models that were morphologically comparable to those acquired using CT technology (Figure 3). In the MRI-derived STL models, the 95th percentile was generally found to be < 1.5 mm (Table 1). The only exception was Mandible 1, which had a 95th percentile of 1.69 mm. These larger geometric deviations were partially due to the submillimetre-thin alveolar ridge of the mandible that could not be imaged due to the limited spatial resolution of the MRI scanner (0.5 mm). Indeed, the 95th percentile of the geometric deviations in the symphyseal area of mandible 1 was 1.64 mm, compared with 0.96 mm and 0.70 mm for Mandible 2 and 3, respectively (Table 1). Another anatomical region that showed slightly larger geometrical deviations in the MRI-derived STL models of all three mandibles were the condyles (Table 1). This phenomenon was most likely caused by a drop in the UTE MRI signal below the chosen threshold value at the edge of the condyles. This lower signal could be due to the higher cortical bone density in the condyle regions. In clinical settings, the presence of soft tissues would offer the possibility to use more advanced segmentation methods [22], instead of the thresholding approach used in this study.

In MRI technology, the spatial resolution is dependent on the number of frequency encoding steps and the size of the field of view. Since the size of the mandible determines the minimal field of view, the resolution is dependent on the frequency encoding, that is how often the free induction decay signal is sampled. Therefore, the resolution of the current UTE MRI sequence could be further improved by increasing the duration of the scan; however, this would be a disadvantage in a clinical setting.

In the CT-derived STL models, the 95th percentile of the geometric deviations in the whole mandible was generally found to be < 0.5 mm (Table 1). All geometrical deviations were < 1.0 mm (Figure 3). The capability of CT to capture small bone structures is also dependent on the spatial resolution of the scanner (0.3 mm), which is directly linked to the hardware properties such as the detector configuration of the scanner.

The aforementioned results are in agreement with a previous study by White et al. [31], who obtained MRI and CT-derived additively manufactured 3D models of 10 ovine knees, and compared caliper measurements of these models with caliper measurements taken from the real bony anatomy. They reported a mean deviation of 2.15 ± 2.44 mm in their MRI models and a mean deviation of 0.61 ± 0.41 mm in their CT models. Lee et al. [32]
reported a mean deviation of 0.7 ± 0.1 mm between their MRI and CT-derived STL models of the femur. A more recent study by Van Den Broeck et al. [33] reported RMS errors of 0.55 mm in their CT-derived STL models of the tibia, and 0.56 mm in their MRI-derived STL models. The results from our study differ from the ones reported by Lee et al. [32] and Van Den Broeck et al. [33]. This may have been due to differences in the evaluated anatomical structures and the measuring protocols. Moreover, the differences may also have been caused by the different MRI sequences used. White et al. [31], Van Den Broeck et al. [33] and Lee et al. [32] all used conventional spin or gradient echo sequences that commonly require more manual segmentation of the bone, which in return can be subjective and very time-consuming.

The UTE MRI protocol used in this study has certain clinical advantages when compared to CT technologies. The major advantage of MR imaging is the lack of radiation exposure, which in return allows for multiple imaging sessions to be undertaken. Furthermore, the UTE sequence used in this study combined with other sequences available on the same MRI device allows for hard and soft tissue visualisation, respectively. This can be especially valuable when treating cancer, ameloblastoma patients and orbital fractures [34], [35].

**Additive manufacturing**

In order to quantify the accuracy of the AM process, the gold standard STL model of Mandible 1 was first additively manufactured and then optically scanned and compared to the original STL model (Figure 5 A). The 95th percentile of geometric deviations introduced during the AM process using the gold standard STL model was 0.46 mm (Table 2). This result suggests that only minor deviations were introduced during the AM process. However, because an optical scanner was used to image the bone, the aforementioned results are not obtainable in clinical settings.

The 95th percentiles of geometric deviations in the MRI and CT-derived additively manufactured models were higher than those obtained with the optical scanner: 1.73 mm and 0.88 mm, respectively (Table 2). These results suggest that both CT and MR imaging have a greater influence on the accuracy of medical additive manufactured constructs than the additive manufacturing process itself.

**Limitations of this study**

The major limitation of the present study was that mandibles without soft tissues were used. Therefore, the results of this study are not simply generalizable to clinical conditions. In clinical settings, a second UTE MRI echo image is required to discriminate bone from air and soft tissues [22]. Furthermore, motion artefacts due to the relatively long duration of the UTE sequence (several minutes) and metal artefacts caused by orthodontic appliances or fillings can affect the image quality in vivo. Further cadaver and patient studies are recommended in order to assess the feasibility of UTE MRI sequences in clinical settings.
CONCLUSION
This study demonstrates that MR imaging of bone using an UTE sequence is a feasible alternative to CT imaging in the generation of STL models of mandibles with different morphologies, and would therefore be suitable for surgical planning and additive manufacturing. The CT and MRI modalities generally introduced geometrical deviations in the STL models of < 1.0 mm and < 1.5 mm, respectively. The additive manufacturing process introduced minor deviations of < 0.5 mm. Further in vivo studies are necessary to assess the feasibility of the UTE MRI sequence in clinical settings.
REFERENCES


30. Implications of commission recommendations that doses be kept as low as readily achievable (1973) ICRP publication 22. Oxford, UK: Pergamon Press


GENERAL DISCUSSION
Even though surgical techniques have improved over time, surgery still remains a manual task. Current surgical techniques in the field of oral and maxillofacial surgery are still heavily dependent on the dexterity and experience of the surgeon. In this context, medical AM opens new avenues to improve the dexterity of the surgeon during surgical procedures. The resulting decrease in operation time and improved accuracy can markedly reduce treatment costs. For example, AM can be combined with virtual planning to create accurate and low-cost individualized orbital floor implants. This novel treatment method has shown to reduce operation time by approximately 40% [1]. However, creating accurate medical AM constructs still requires technical and computational knowledge that can be a challenge for clinicians. Therefore, AM constructs are currently most often designed by engineers with in-depth knowledge of the field of anatomical modelling and reverse engineering.

While the aforementioned AM constructs can be fabricated with great accuracy (< 0.1 mm) [2] using current printing technologies, their design is currently limited by the accuracy of the medical images used during the fabrication process. It must be noted that image quality can differ markedly between MDCT, DECT and CBCT scanners. [3] Moreover, scanner-specific acquisition parameters can also influence the image quality [4]. Of all the scanners on the market, MDCT scanners offer the best image quality and accuracy for medical AM [3]. However, radiation safety still needs to be taken into consideration when acquiring images for AM, given the current concerns of radiation-induced secondary malignancies [5]. Therefore, if possible, low-dose MDCT scanning protocols should be used to minimise the effective dose in patients who require AM constructs [3]. Even though current CBCT devices can attain lower dose values than MDCT devices in the maxillofacial area [6], their image quality remains problematic for medical AM purposes. The CBCT-derived STL models evaluated in this PhD thesis demonstrated large non-uniform inaccuracies of up to 0.9 mm in the maxillofacial area (Chapter 3), which was probably due to the relatively high amount of noise and number of artefacts [7] and inhomogeneities in the CBCT X-ray beam [8]. Bearing the above-mentioned concerns in mind, more emphasis should be put into developing UTE MRI sequences that are specially designed for AM purposes [9].

One possible method of improving the overall image quality of CBCT scanners in the maxillofacial region is to change the head position in the gantry. According to our studies, the best CBCT image quality can be achieved on a mobile gantry device using prone or supine imaging positions [10]. More specifically, the prone imaging position resulted in the best image quality for the sinus, mandible and maxilla when compared with the supine and oblique imaging positions [11]. The oblique imaging position offered inadequate image quality except in the sinus region. These novel insights regarding the image quality of CBCT scanners may help to improve the accuracy of CBCT images for medical AM purposes. A more advanced method of improving the accuracy of CBCT images would be to develop adaptive, flexible scan protocols that would allow physicians to zoom into anatomical areas of interest during image acquisition. For example, an initial “coarse”
CBCT scan of a patient’s head could be acquired using a large field of view, followed by a more precise local scan of the region of interest, e.g., a tumour or fracture site. This process would, however, require real-time reconstruction of CBCT images using powerful cluster computers. Another promising new technology that could revolutionise the overall image quality required in the AM process is spectral imaging [12]. Spectral imaging refers to a technique that uses photon-counting detectors to separate photons coming from a polychromatic X-ray source into a series of energy bins, for which the energy threshold levels can be configured using dedicated software. Each tissue type in the human body has a different response curve to photon energies. Therefore, it is possible to discriminate tissues based on their response to different photon energies. This could greatly improve tissue discrimination and delineation in CT and CBCT images.

In addition to the aforementioned CT imaging process, the conversion of the resulting images into STL models remains a challenge in medical AM. The most challenging step in the conversion process is image segmentation. Our literature review on bone segmentation methods used in medical AM (Chapter 6) revealed a large spread in accuracies amongst the current methods used ranging from 0.04 mm to 1.9 mm. Image segmentation is often compromised by factors such as the partial volume effect, scattering artefacts and noise [13].

To date, global thresholding remains the most commonly used segmentation method in medical AM. However, the selection of a suitable threshold value for bone still remains a manual task that can introduce geometric variations amongst engineers [14]. Generally, global thresholding achieves accuracies under 0.6 mm [15]. Advanced thresholding methods could improve the accuracy of CT-derived STL models to under 0.38 mm [16]. However, such methods are, to the best of our knowledge, not included in the FDA-approved software packages for medical AM. Finally, it should be noted that the triangle-based STL file format was not initially developed to describe complex biological structures that encompass many curved surfaces and different tissue densities. Consequently, many small triangles are needed to accurately describe a curved surface. A solution could be to employ novel file formats such as STL2 [17] or the additive manufacturing file format (AMF) [18]. However, these novel file formats are not yet supported by the software packages and printers currently used in medical AM.

Current CT and CBCT scanners deliver voxel-based datasets in which anatomical structures are represented by a continuous spectrum of grey values that are related to attenuation coefficients. Medical AM, on the other hand, requires binary datasets in the form of STL models. Therefore, the current medical AM process always requires a conversion from continuous data into binary data, i.e., image segmentation. A major drawback with the image segmentation process is that a large amount of the valuable biological information that was initially present in the CT image in the form of grey values is discarded. This can be considered as a waste of photons. One solution to this problem would be to develop an algorithm that could directly compute mesh volumes (STL models) from raw CT projection data. This could be achieved by restricting the number
of grey values in a reconstructed CT image to a small, discrete number of values known in advance: one for each tissue type, i.e., “air”, “soft tissue”, “cancellous bone” and “cortical bone”. Consequently, since each voxel is already assigned to a specific tissue type during the reconstruction phase, image segmentation as a whole becomes redundant. This technique is known as discrete tomography [19]. The major advantage of discrete tomography over conventional CT is that it maximises the amount of information obtained from a CT scan for a given dose level: “making the most out of each photon”. Therefore, I hypothesize that discrete tomography could markedly reduce the number of projections and radiation dose needed to acquire an accurate STL model.

As mentioned above, the creation of accurate medical AM constructs still requires a large amount of human intervention by experienced clinicians and engineers, particularly in the imaging and image processing steps. This problem could be tackled by automating certain processes in the medical AM workflow using novel artificial intelligence (AI) approaches [20]. The basic mathematical concept of AI dates back to the 1940s when researchers proposed the development of an electrical network that mimics the function of neurons in the human brain [21]. This was the onset for further developments in the field of AI and led to a variety of machine learning (ML) algorithms that are today used in the fields of computer vision and medical imaging, among others. In recent years, the development of graphic processing units (GPUs) has enabled the use of multi-layered ML algorithms such as convolutional neural networks (CNNs) [22]. Currently, CNNs are becoming the methodology of choice for analysing medical images [23] and show great potential in performing accurate and automated image segmentation tasks [24]–[26]. Furthermore, a CNN can be taught to reconstruct complex 3D shapes from CT data [27] and could therefore theoretically be used to auto-complete STL models of disfigured skulls. A CNN can be taught in a supervised or unsupervised manner. To date, the majority of CNNs in the field of medical imaging use supervised learning, in which a CNN “learns from experience” based on a large amount of labelled training data (e.g., medical images). It should be pointed out that the performance of this class of CNN depends heavily on the availability and the quality of the images used for training. As opposed to supervised learning, in unsupervised learning a CNN is taught to independently discover patterns in data. Taking the above into account, one can postulate that recent developments in AI combined with medical imaging (“big”) data will revolutionise healthcare and will subsequently offer new possibilities in the field of medical AM, virtual/augmented reality and robotics. These developments will most likely lead to more automation, which, in turn, will have a fundamental socio-economic impact on the global labour market in the 21st century.
REFERENCES


SUMMARY

AM offers possibilities to fabricate different types of patient-specific constructs such as anatomical models, surgical guides and implants. Such constructs can improve the surgeon’s dexterity, decrease operation time and improve accuracy of surgical procedures. As demonstrated in Chapter 2, AM can be combined with virtual planning to create accurate and low-cost individualised orbital floor implants. A virtual 3D model of a patient with an orbital floor fracture was generated from CT images of the patient. The fracture was virtually closed using spline interpolation. The resulting 3D model was used to design and manufacture a mould of the defect site using an inkjet printer. This tangible mould was subsequently used during surgery to sculpture an individualised autologous orbital floor implant. This novel treatment method enhanced the overall accuracy and efficiency of the surgical procedure. The sculptured autologous orbital floor implant showed an excellent fit in vivo and reduced the operation time by approximately 40%.

Despite recent advances and promising case studies involving medical AM, notable scientific, technological and regulatory challenges remain. Each step in the medical AM process, i.e., imaging (step 1), image processing (step 2) and manufacturing (step 3) can introduce inaccuracies in the resulting construct. However, since the aforementioned AM constructs can be fabricated with great accuracy (< 0.1 mm) using current printing technologies, their design is currently limited by the accuracy of the imaging and image processing steps. Therefore, the general aim of this thesis was to provide an insight into the current scientific and technological challenges faced in medical additive manufacturing with respect to imaging and image processing. More specifically, the different parameters that influence the accuracy of patient-specific AM constructs were identified.

In Chapter 3, the image quality and the accuracy of STL models acquired using different CT scanners and acquisition parameters is assessed. Images of three dry human skulls were acquired using different scanning protocols on two multi-detector row computed tomography (MDCT) scanners, a dual energy computed tomography (DECT) scanner and one cone beam computed tomography (CBCT) scanner. All images were ranked according to their image quality and converted into STL models. These STL models were subsequently compared with corresponding gold standard models acquired using an optical 3D surface scanner. It was found that the image quality differed between the MDCT, DECT and CBCT scanners. Images acquired using low-dose MDCT protocols were preferred over images acquired using routine protocols. All CT-based STL models demonstrated non-uniform geometrical deviations of up to +0.9 mm. The largest deviations were observed in CBCT-derived STL models. Therefore, it can be concluded that CT imaging technologies and their acquisition parameters can markedly affect the accuracy of medical AM constructs.

In Chapter 4, the impact of head positioning on CBCT image quality is evaluated. The impact of supine, prone and oblique patient imaging positions on the image quality, contrast-to-noise ratio and figure of merit value in the maxillofacial region of a fresh frozen cadaver head was assessed using a CBCT scanner. In addition, the CBCT supine images
were compared with supine MSCT images. The best CBCT image quality was achieved in the prone imaging position for sinus, mandible and maxilla, followed by the supine and oblique imaging positions. The MSCT scanner offered similar image qualities to the 7.5-mA supine images acquired using the mobile CBCT scanner. The prone imaging position offered the best CNR and FOM values on the mobile CBCT scanner. These findings can help to improve CBCT image quality in the maxillofacial region, which, in turn, may improve the overall accuracy of CBCT-derived AM constructs.

In addition to the CT imaging step, the accuracy of medical AM constructs is affected by errors introduced during image processing, particularly during the image segmentation process. Therefore, in Chapter 5, thirty-two publications that reported on the accuracy of different CT image segmentation methods for bone used in medical AM are reviewed. The advantages and disadvantages of the different image segmentation methods used in these studies were evaluated and the reported accuracies were compared. The spread between the reported accuracies was large (0.04 mm to 1.9 mm). Global thresholding was the most commonly used segmentation method with accuracies under 0.6 mm. The disadvantage of this method is the extensive manual post-processing required. Advanced thresholding methods could improve the accuracy to under 0.38 mm. However, such methods are currently not included in commercial software packages. Statistical shape model methods resulted in accuracies from 0.25 mm to 1.9 mm but they are only suitable for anatomical structures with moderate anatomical variations. To improve the accuracy and reduce the cost of patient-specific AM constructs, more advanced image segmentation methods are required.

Chapter 6 provides more insight into the most commonly used image segmentation method in medical AM, namely global thresholding. The impact of manual and default threshold selection on the reliability and accuracy of skull STL models acquired using different CT technologies was evaluated. One female and one male human cadaver head were imaged using MDCT, DECT and two CBCT scanners. Four medical engineers manually thresholded the bony structures on all CT images. The lowest and highest selected mean threshold values and the default threshold value were used to generate skull STL models. Geometric variations between all manually thresholded STL models were calculated. Additionally, in order to calculate the accuracy of the manually and default thresholded STL models, all STL models were superimposed on an optical scan of the dry female and male skulls (“gold standard”). The intra- and inter-observer variability of the manual threshold selection was good (intra-class correlation coefficients >0.9). All engineers selected grey values closer to soft tissue to compensate for bone voids. Geometric variations between the manually thresholded STL models were 0.13 mm (MDCT), 0.59 mm (DECT) and 0.55 mm (CBCT). All STL models demonstrated inaccuracies ranging from −0.8 mm to +1.1 mm (MDCT), −0.7 mm to +2.0 mm (DECT) and −2.3 mm to +4.8 mm (CBCT). The findings of this study demonstrate that manual threshold selection results in better STL models than default thresholding.
In Chapter 7, a novel method was developed to assess the accuracy and reliability of CT scanners and software packages. To this end, an anthropomorphic phantom that provided ground truth measurements for the evaluation of the oropharynx morphology was 3D printed. This phantom was used to assess the accuracy of two multi-detector row computed tomography (MDCT) scanners (GE Discovery CT750 HD, Siemens Somatom Sensation) and three CBCT scanners (NewTom 5G, 3D Accuitomo 170, Vatech PaX Zenith 3D). All CT images were segmented by two observers and converted into standard tessellation language (STL) models. The volume and the cross-sectional area of the oropharynx were measured on the acquired STL models. Finally, all STL models were registered and compared with the gold standard. Significant differences were observed in the volume and cross-sectional area measurements of the oropharynx acquired using the different MDCT and CBCT scanners. The Siemens MDCT and the Vatech CBCT scanners were more accurate than the GE MDCT, NewTom 5G, and Accuitomo CBCT scanners.

AM medical models, implants and drill guides are becoming increasingly popular amongst maxillofacial surgeons and dentists. This rise in popularity has subsequently increased the number of CT and CBCT scans performed worldwide. However, all X-ray based imaging modalities induce harmful ionizing radiation to the patient. Therefore, Chapter 8 assesses the feasibility of using radiation-free UTE MRI sequences for medical AM. Three morphologically different dry human mandibles were scanned using a CT and MRI scanner. All CT and MRI scans were converted into STL models and geometrically compared with a corresponding gold standard STL model acquired using an optical 3D scanner. To quantify the accuracy of the AM process, the CT, MRI and gold-standard STL models of one of the mandibles were additively manufactured, optically scanned and compared with the original gold standard STL model. Geometric differences between all three CT-derived STL models and the gold standard were < 1.0 mm. All three MRI-derived STL models generally presented deviations < 1.5 mm in the symphyseal and mandibular area. It should be noted that the AM process only introduced minor deviations of < 0.5 mm. Hence, UTE MRI sequences offer an alternative to CT in generating STL models of the mandible and could be suitable for surgical planning and AM.
SAMENVATTING

Additive manufacturing (AM), oftewel driedimensionaal (3D-) printen, maakt het mogelijk om patiëntspecifieke producten te fabriceren zoals anatomische modellen, chirurgische boormallen en implantaten. Dergelijke toepassingen verkorten de operatieduur en kunnen de nauwkeurigheid van de operatie verbeteren. In Hoofdstuk 2 wordt een nieuwe behandelmethode gepresenteerd waarbij 3D-printen gecombineerd wordt met virtuele chirurgische planning om nauwkeurige en goedkope patiëntspecifieke implantaten te fabriceren voor de reconstructie van de bodem van de oogkas (orbitabodem). De eerste stap van deze nieuwe behandelmethode is om computertomografie (CT-) beelden van een patiënt met een orbitabodemfractuur te converteren naar een virtueel 3D-model. Vervolgens wordt het defect in de orbitabodem virtueel gesloten. Op basis van het resulterende 3D-model wordt vervolgens een mal ontworpen die geprint wordt met een inkjet 3D-printer. De chirurg kan een dergelijke fysieke mal tijdens de operatie als hulpmiddel gebruiken om van autoloog bot een patiëntspecifiek orbitabodemimplantaat optimaal vorm te geven. Deze nieuwe behandelmethode leidt tot een hogere chirurgische precisie en efficiëntie en kan de operatietijd significant verkorten.

Ondanks de recente ontwikkelingen en veelbelovende medische casuïstiek op het gebied van 3D-printen zijn er nog vele wetenschappelijke, technologische en reglementaire uitdagingen te overwinnen. Het 3D-printen van een medisch product vereist drie stappen: 1) medische beeldvorming; 2) beeldverwerking; en 3) 3D-printen. Elk van deze stappen kan de nauwkeurigheid van medische 3D-geprinte producten beïnvloeden. Echter, omdat moderne 3D-printtechnieken zeer nauwkeurig zijn (< 0,1 mm), is de nauwkeurigheid met name afhankelijk van de medische beeldvorming en de daaropvolgende beeldverwerking. Het doel van deze dissertatie is inzicht te verkrijgen in de huidige wetenschappelijke en technologische uitdagingen op het gebied van medisch 3D-printen, met name omtrent medische beeldvorming en beeldverwerking. Ons onderzoek heeft geleid tot de identificatie van verschillende parameters die de nauwkeurigheid van medische 3D-geprinte producten beïnvloeden.

In Hoofdstuk 3 wordt de invloed van verschillende CT-scanners en scanparameters onderzocht op de beeldkwaliteit en de nauwkeurigheid van de resulterende virtuele 3D-modellen. Drie droge menselijke schedels werden gescand met twee multidetectorcomputertomografie (MDCT) scanners, een dual-energy computertomografie (DECT) scanner, en een cone-beam computertomografie (CBCT) scanner.

Op elk van deze CT-scanners werden verschillende scanprotocollen getest. Twee ervaren radiologen beoordeelden de beeldkwaliteit van de resulterende CT-scans. Hierna werden de scans geconverteerd naar virtuele 3D-modellen en werd de nauwkeurigheid van deze 3D-modellen bepaald door ze te vergelijken met een “gouden standaard” (optische 3D-scan). De radiologen namen significante verschillen in beeldkwaliteit waar tussen de MDCT-, DECT- en CBCT-scanners. Bovendien prefereerden de radiologen het lage-dosis-scanprotocol boven het standaard scanprotocol op de MDCT-scanner. Alle
virtuele 3D-modellen vertoonden niet-uniforme onnauwkeurigheden tot +0,9 mm, waarvan de modellen verkregen met CBCT de grootste afwijkingen vertoonden. De belangrijkste conclusie van Hoofdstuk 3 is dat het type CT-scanner en het gebruikte scanprotocol een grote invloed kunnen hebben op de nauwkeurigheid van medische 3D-geprinte producten.

In Hoofdstuk 4 wordt de invloed van de hoofdpositie in de CBCT-scanner op de beeldkwaliteit, contrast-to-noise ratio (CNR) en figure-of-merit (FOM) onderzocht. Mobiele CBCT-scanners kunnen het hoofd van een patiënt in rugligging, in buikligging, of zittend met het hoofd in een schuine positie scannen. In elk van deze drie posities werd een CBCT-scan gemaakt van een vers ingevroren kadaver hoofd. Ook werd er ter vergelijking een MDCT-scan gemaakt van het kadaverhoofd. De beste beeldkwaliteit voor de sinussen, mandibula en maxilla op de CBCT-scan werd verkregen in buikligging, gevolgd door de rugligging. De schuine positie gaf de laagste beeldkwaliteit. De beeldkwaliteit van de CBCT-scan van het kadaver hoofd in rugligging met een buisstroom van 7,5 mA was vergelijkbaar met de MDCT-scan. De rugligging resulteerde in de beste CNR en FOM voor de CBCT-scanner. Deze nieuwe inzichten over de kwaliteit van CBCT-scan kunnen de algehele nauwkeurigheid van 3D-geprinte producten verkregen uit CBCT-scan ten goede komen.

Het is niet alleen de medische beeldvorming die de nauwkeurigheid van medische 3D-printen beïnvloedt maar ook de beeldverwerking. De belangrijkste stap van beeldverwerking is beeldsegmentatie. Hoofdstuk 5 geeft daarom een overzicht van tweeëndertig wetenschappelijke artikelen die rapporteren over de nauwkeurigheid van verschillende CT-beeldsegmentatiemethodes voor botstructuren. De voor- en nadelen van deze methodes worden samengevat en de nauwkeurigheden worden met elkaar vergeleken in een tabel. Deze tabel laat zien dat de nauwkeurigheden in de literatuur sterk uiteenlopen: 0,04 mm tot 1,9 mm. De meest gebruikte segmentatiemethode in medisch 3D-printen is de zogenaamde drempelmethode (“thresholding”). Deze methode resulteert in nauwkeurigheden onder 0,6 mm. Een belangrijk nadeel van de drempelmethode is dat men vaak handmatig ruis of artefacten van het virtuele 3D-model moet verwijderen. Geavanceerde drempelmethodes kunnen de nauwkeurigheid van virtuele 3D-modellen verbeteren tot onder 0,38 mm. Echter, dergelijke methodes zijn tot op heden niet beschikbaar in de commerciële softwarepakketten voor medisch 3D-printen. Ten slotte resulteren de segmentatiemethodes gebaseerd op statistische vormmodellen in nauwkeurigheden tussen 0,25 mm en 1,9 mm, al zijn dergelijke methodes alleen geschikt voor het segmenteren van botstructuren met beperkte anatomische variaties. Kortom, betere beeldsegmentatiemethodes zouden de nauwkeurigheid en de kosten van medisch 3D-printen ten goede komen.

Hoofdstuk 6 gaat dieper in op de drempelmethode. Een medisch ingenieur selecteert meestal handmatig de juiste grijswaarde als drempelwaarde om botstructuren te segmenteren in CT-beelden, al bevatten de meeste softwarepakketten ook een automatische drempelwaarde. Daarom wordt in Hoofdstuk 6 de invloed onderzocht van handmatige en automatische drempelwaardeselectie op de betrouwbaarheid en
de nauwkeurigheid van virtuele 3D modellen verkregen met verschillende CT-scanners. Een mannelijk en een vrouwelijk gebalsemde kadaverhoofd werden gescand met een MDCT-, een DECT- en twee CBCT-scanners. Vier medisch ingenieurs selecteerden vervolgens in iedere CT-scan een drempelwaarde om bot te segmenteren. De laagste en de hoogste geselecteerde drempelwaarde werden vervolgens gebruikt om virtuele 3D-modellen te genereren. Dit werd ook gedaan met de automatische drempelwaarde. De geometrische variaties tussen alle handmatig gegenereerde 3D-modellen werden berekend. Ook werd de nauwkeurigheid van alle 3D-modellen bepaald door deze te vergelijken met een gouden standaard (een optische 3D-scan van de schedels waarvan het zachte weefsel verwijderd was). De intra- en inter-observer variabiliteit van de handmatige drempelwaardeselectie was goed (intra-class correlatiecoëfficiënt > 0,9). Alle medisch ingenieurs selecteerden drempelwaardes die leken op de grijswaarden van de zachte, omliggende weefsels om zo veel mogelijk bot te segmenteren. Echter, er werden geometrische variaties tussen de resulterende 3D-modellen waargenomen van 0,13 mm (MDCT), 0,59 mm (DECT), en 0,55 mm (CBCT). De nauwkeurigheden liepen uiteen van -0,8 mm tot +1,1 mm (MDCT), −0,7 mm tot +2,0 mm (DECT), en van −2,3 mm tot +4,8 mm (CBCT). Een andere interessante bevinding was dat de handmatig geselecteerde drempelwaardes resulteerden in betere 3D-modellen dan die verkregen met de automatische drempelwaarde.


Medisch 3D-printen wint snel in populariteit onder mond-, kaak- en aangezichtschirurgen en tandartsen. Dit heeft echter ook tot gevolg dat het aantal gemaakte CT- en CBCT-scans wereldwijd toeneemt. Alle beeldvormende modaliteiten waarbij röntgenstraling gebruikt wordt, stellen de patiënt bloot aan schadelijke straling. Daarom wordt in Hoofdstuk 8 een nieuwe, stralingsvrije methode onderzocht om bot in beeld te brengen: een ultrashort echo time (UTE) sequentie op een magnetic resonance imaging (MRI-) scanner. Drie droge mandibulae werden gescand met een UTE MRI-sequentie en een CT-scanner. De resulterende scans werden geconverteerd naar virtuele 3D-modellen en vergeleken
met een gouden standaard (een optische 3D-scan). Ook werd de nauwkeurigheid van het 3D-printproces onderzocht door alle virtuele 3D-modellen te printen op de inkjet printer van het VU medisch centrum en deze opnieuw te vergelijken met de gouden standaard. Alle virtuele 3D-modellen verkregen met de CT-scanner resulteerden in nauwkeurigheden < 1,0 mm. Ter vergelijking, alle virtuele 3D-modellen verkregen met de UTE MRI-sequentie waren < 1,5 mm nauwkeurig in de symphysis mandibulae en de corpus mandibulae. Een andere interessante bevinding was dat het 3D printproces slechts minimale afwijkingen introduceerde (< 0,5 mm). UTE MRI-sequenties kunnen dus een alternatief bieden voor CT voor het genereren van virtuele 3D-modellen van de mandibula. Dergelijke modellen kunnen gebruikt worden voor virtuele chirurgische planning en 3D-printen.
ACKNOWLEDGEMENTS

The 3D Innovation Lab of the VU University Medical Center (VUmc) constantly explores and invests in 3D imaging and printing technologies and a wide array of medical virtual planning software solutions required for the fabrication of individualized medical constructs hence human spare parts. This unique collaboration and education space formed the basis of my PhD research. Therefore, I would like to express my gratitude to all people involved in the 3D Innovation Lab: Niels Liberton, Sjoerd te Slaa, Frank Verver, Micha Paalman, Dafydd Visscher, dr. Jack van Loon, Angela Ridwan-Pramana, dr. Sophie Kommers, Omar Hertgers, Ivar Kommers, Rianne van Loenen, Professor Ruud Verdaasdonk and our colleagues from the Academic Medical Center (AMC), Professor Theo Smit, dr. Geert Streekstra en dr. Iwan Dobbe. Also thanks to our Bachelor and Master students that contributed to our research over the past years: Roelof van Dijk, David Spies, Fedde Kuilman, Niki Tokkari, Daan Brons, Karen Kiers, Guus van Baar, Janne Ploeger, Jordi Minnema. Meghan Simons, Sevde Yen, Claire Olthoff, and Ezra Esmeijer. In this context, I would like to thank Theo Faes for coordinating our Master students.

I would like to thank the department of Oral & Maxillofacial Surgery/Pathology and my department head and promotor Professor Tymour Forouzanfar for facilitating my PhD. Also thanks to all of my colleagues from the department. I am particularly grateful to Annelies van der Geest and dr. Martijn van Steenbergen for their consistent support and excellent organization skills.

Dr. Eric-Jan Rijkhorst and dr. Dennis Heijtel, I would like to thank you for the interesting physics-related discussions and your contributions to our research. Eliane Kaaij and Jasmine Rubirayoxall, thank you for your outstanding anatomical skills and passion. And Professor Paul van der Stelt and dr. Hui Chen, thank you for the productive collaboration.

A special thanks to Planmeca Oy for providing financial support for my PhD project. Timo Müller, Vesa Varjonen, Jaakko Lähelmä, dr. Kalle Karhu, dr. Ari Hietanen, and dr. Juha Koivisto, I hope we can continue our excellent collaboration in the future. I would also like to offer my thanks to DSM Biomedical B.V., and particularly Jac Koenen, for their financial support. Finally, I would like to thank Anke Kleinveldd and her colleagues from the Institute for Education and Training for allowing me to become a tutor at the VUmc School of Medical Sciences and thereby supporting my research at the VUmc. It was an amazing experience for me to guide Bachelor students through their first years of medical education.

Thanks so much to my family and friends for their unlimited encouragement and support. In this context, I would like to thank my paronymphs Elise van Eijnatten and Mike van Rijssel. Floor Maarse, thank you for being such a great swim teacher. Furthermore, I am especially
grateful to Daniël Doorman, for always taking the time to listen to me and standing by my side. A special thanks to Peter Heath for revising the manuscript and to my dad for revising the Dutch summary. I am also grateful to all of the co-authors of the related publications.

Finally, I am extremely grateful to my mentor and co-promotor Jan Wolff for believing in me from the beginning of my research career and for his consistent support.