2014

Three-dimensional image technology in forensic anthropology: assessing the validity of biological profiles derived from CT-3D images of the skeleton

BOSTON UNIVERSITY

SCHOOL OF MEDICINE

Thesis

THREE-DIMENSIONAL IMAGE TECHNOLOGY IN FORENSIC ANTHROPOLOGY: ASSESSING THE VALIDITY OF BIOLOGICAL PROFILES DERIVED FROM CT-3D IMAGES OF THE SKELETON

by

MARIA JULIA GARCIA DE LEON VALENZUELA

B.Sc., University of Reading, 2010

Submitted in partial fulfillment of the

requirements for the degree of

Master of Science

2014

Approved by

First Reader

Donald F. Siwek, Ph.D.
Associate Professor of Anatomy and Neurobiology
Program in Forensic Anthropology

Second Reader

James T. Pokines, Ph.D.
Assistant Professor
Program in Forensic Anthropology

DEDICATION

To Ángel Valencia and Leticia González. You inspire me.

ACKNOWLEDGEMENTS

I would like to acknowledge the contributions of Dr. Donald F. Siwek and Dr. James T. Pokines to the compilation of this manuscript. Special thanks go to Margaret Lavoye and Justine K. DePonte from Boston Medical Center's Radiology Department for their help with the scanning of my samples, as well as to Darryl Ricketts for his guidance during the course of this project. I would also like to thank Dr. Darlene R. Ketten from Harvard Medical School and the Woods Hole Oceanographic Institution in Woods Hole, MA for her valuable input. Finally, I am grateful for the help of Dr. Jonathan Bethard, Celena Toon, Gregory Moore, Karen Brun and Rebecca King; without them, this project would not have been possible.

THREE-DIMENSIONAL IMAGE TECHNOLOGY IN FORENSIC ANTHROPOLOGY: ASSESSING THE VALIDITY OF BIOLOGICAL PROFILES DERIVED FROM CT-3D IMAGES OF THE SKELETON

MARIA JULIA GARCIA DE LEON VALENZUELA

Abstract

This project explores the reliability of building a biological profile for an unknown individual based on three-dimensional (3D) images of the individual's skeleton. 3D imaging technology has been widely researched for medical and engineering applications, and it is increasingly being used as a tool for anthropological inquiry. While the question of whether a biological profile can be derived from 3D images of a skeleton with the same accuracy as achieved when using dry bones has been explored, bigger sample sizes, a standardized scanning protocol and more interobserver error data are needed before 3D methods can become widely and confidently used in forensic anthropology.

3D images of Computed Tomography (CT) scans were obtained from 130 innominate bones from Boston University's skeletal collection (School of Medicine). For each bone, both 3D images and original bones were assessed using the Phenice and Suchey-Brooks methods. Statistical analysis was used to determine the agreement between 3D image assessment versus traditional assessment. A pool of six individuals with varying experience in the field of forensic anthropology scored a subsample (n = 20) *to explore interobserver error. While a high agreement was found for age and sex estimation for specimens scored by the author, the interobserver study shows that observers found it difficult to apply standard methods to 3D images. Higher levels of experience did not result in higher agreement between observers, as would be expected. Thus, a need for training in 3D visualization before applying anthropological methods to 3D bones is suggested. Future research should explore interobserver error using a larger sample size in order to test the hypothesis that training in 3D visualization will result in a higher agreement between scores. The need for the development of a standard scanning protocol focusing on the optimization of 3D image resolution is highlighted.*

Applications for this research include the possibility of digitizing skeletal collections in order to expand their use and for deriving skeletal collections from living populations and creating population-specific standards. Further research for the development of a standard scanning and processing protocol is needed before 3D methods in forensic anthropology are considered as reliable tools for generating biological profiles.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF TABLES, CONTINUED

LIST OF FIGURES

LIST OF FIGURES, CONTINUED

LIST OF ABBREVIATIONS

μCT	Micro-computed Tomography
2D	Two-dimensional
3D	Three-dimensional
BUDOC	Boston University Donated Osteological Collection
CAESAR	Civilian American and European Surface Anthropometry Resource
CAT	Computed Axial Tomography
CT	Computed Tomography
DICOM	Digital Imaging and Communications in Medicine
GNU GPL	GNU General Public License
MDCT	Multi-detector Computed Tomography
MRI	Magnetic Resonance Imaging
MSCT	Multi-slice Computed Tomography
QTVR	QuickTime Virtual Reality

CHAPTER 1: INTRODUCTION

What in recent years has become known as virtual anthropology consists of a multidisciplinary approach to the study of human anatomical data by combining quantitative analysis with digital technologies (Weber and Bookstein 2011). It has been applied within the discipline of anthropology to answer questions regarding hominoid growth, trauma analysis, individual biological profiles, craniofacial reconstruction, bone development, forensics and in the reconstruction of fossils (Benazzi *et al.* 2009; Benazzi *et al.* 2010; Bilfield *et al.* 2012; Brough *et al.* 2013; De Greef and Willems 2005; Decker *et al.* 2011; DeSilva *et al.* 2013; Ferrant *et al.* 2009; Friess 2012; Grabherr *et al.* 2009; Hoppa 2000; Kettner *et al.* 2011; Lottering *et al.* 2014; Robinson *et al.* 2008; Sansoni *et al.* 2009; Telmon *et al.* 2005; Thali *et al.* 2003a, b, and c; Tocheri *et al.* 2002; Uldin *et al.* 2012; Villa *et al.* 2013a and b; Willis *et al.* 2007; Wink 2011; Wink 2014; Wozniak *et al.* 2012; Zech *et al.* 2012; Zollikofer *et al*. 1998).

The benefits of virtual anthropology have been noted by various researchers in the field. Kullmer (2008) considers the possibility of non-invasive study of remains as one of the method's major contributions to anthropology, allowing researchers to study fragile remains without damaging the original sample, as well as opening the door to the acquisition of structural data that would otherwise be inaccessible, such as the study and reconstruction of a fossil hominid's auditory canals. One of the earlier applications of three-dimensional (3D) imaging techniques in anthropology involved the scanning of hominid fossils using computed tomography (CT) imaging that allowed access to hidden anatomical structures, such as the inner ear (Zollikofer *et al*. 1998). More recently, DeSilva *et al.* (2013) have extracted parts of an *Australopithecus sediba* fossil still embedded in its depositional context using micro-computed tomography (μCT) scans, which permitted the study of a highly valuable and rare fossil without risking its destruction.

Aso *et al.* (2005) also stressed the non-destructive and non-invasive nature of performing *virtuopsies* (which is a term derived from the words "virtual" and "autopsy"), and suggested data digitization as a way to preserve evidence in forensic cases. Other researchers, such as Elton and Cardini (2008) noted the increased potential for data sharing among anthropologists, as well as the preservation of museum collections for future generations, although they warn that the maintenance of digital skeletal collections is likely to be fraught with high costs and maintenance issues. There is also the added benefit of lower costs for researchers who wish to study skeletal collections, since a digital database would translate into fewer travel and accommodation expenses, as well as increased time frame for research and the repeatability of studies, since the source samples would no longer be confined to a museum, but instead theoretically be available to anyone that needs them.

Aside from these benefits, the use of 3D images in anthropology allows researchers to study skeletal material in a non-destructive, non-invasive manner (Aso *et al.* 2005), which is an advantage especially if remains are fragile. A potential application of this technology to forensics is the study and reconstruction of human remains exposed to thermal damage. It would also permit preservation of important collections for future generations, greater replicability of studies, and greater access to skeletal collections. Additionally, Brough *et al.* 2013 have suggested the use of 3D images from CT scans would provide forensic anthropologists with the ability to determine the biological profile of an individual without having to macerate remains (and possibly damage them in the process).

Despite the many applications of virtual anthropology, the validity of conducting research based on 3D images without having access to the actual skeletal remains from which they derive is an issue that is only recently being explored (for examples, see Dedouit *et al.* 2007 and Telmon *et al.* 2005). For that reason, one of the aims of the present study is to assess the reliability of determining an individual's biological profile from 3D images of skeletal remains, with the goal of establishing the validity of conducting anthropological research "from the desk" (Elton and Cardini 2008:209). If scoring 3D images derived from bones proves to be a viable alternative to traditional methods, the applications that can be derived for the fields of biological and forensic anthropology hold the promise to help solve some of the greater issues that have long plagued both disciplines, such as a lack of modern samples with which to derive regression equations that will more appropriately estimate the biological profile of contemporary populations, a current lack of repeatability and data verification both in research and in forensic cases, and the need for wider access to skeletal samples.

Most methods commonly employed in biological anthropology (and its applications in the field of forensics) for estimating the biological profile of individuals were not devised using data from modern populations. Instead, specimens from collections dating to the end of the 19th century and the beginning of the 20th century (e.g., the Hamann-Todd Collection housed in the Cleveland Museum of Natural History and the Terry Collection housed in the National Museum of Natural History at the Smithsonian Institution) were commonly used for data gathering as well as for validating anthropological methods, i.e., replicating a previously published study and comparing the results obtained in order to assess whether a method is reliable. Some of these methods have become a staple of the anthropologist's methodology, such as the Phenice (1969) method for sex estimation.

The issue with applying methods that were developed using these types of collections to forensic cases, however, is that the data derived from those collections can be considered to be 'outdated' in anthropological terms. Research done in anthropology suggests that environmental factors such as nutrition and disease, and not just heredity, have an effect on the overall size and shape of human skeletons. In his famous migration studies, Franz Boas (one of the fathers of modern biological anthropology in the United States) studied the effects of the environment on immigrants and found that first generation migrants (i.e., born in their home countries but living in the United States) were taller than their compatriots and that the children of first generation migrants were taller than their parents (Boas 1912). Meadows Jantz and Jantz (1999) explored secular changes in long bone lengths in U.S. populations using data from the Huntington Collection, the Terry Collection, the skeletons of World War II casualties and the Forensic Anthropology Data Bank (a database composed of data collected by forensic anthropologists using the cases they work on while on the job). They found that environmental forces are the usual causes of secular change in the long bone length and overall size of the individuals studied. In another publication, Jantz and Meadows Jantz (2000) found secular changes in the shape of the cranial vault (and to its size, to a lesser degree) when comparing specimens from the Terry and Hamann-Todd collections to data obtained from the Forensic Anthropology Data Bank that follow a similar trend to what was observed in their previous research on long bone length (Meadows Jantz and Jantz 1999). Malina *et al.* (2004) explored secular change in rural Oaxaca, Mexico, and found significant increases in height, sitting height and estimated leg lengths in children and adolescents between 1978 and 2000. They also found that the secular change observed was greater for adolescent and young adult males compared to females. Research in this field raises the question of whether it is valid to apply common anthropological methods to current forensic cases, since secular change has been proven to be one factor that greatly affect human dimensions.

It is for this reason that there is a need for research to be done using data derived for current populations, especially considering that methods developed using populations from the United States are also being applied to other regions of the world, where the data are even less of a fit for the populations the methods are being applied to.

Unfortunately, there are few collections composed of modern individuals and those that do exist are often composed of donated bodies (e.g., the Bass Collection housed at the University of Tennessee). This presents with the issue of the average age of the individuals in the collections skewing towards the old. Many researchers working in forensic cases gather data from the individuals they work with (hence the existence of the Forensic Data Bank, currently consisting of 3400 cases out of which 2400 have documented sex and ancestry according to the Forensic Anthropology Center at the University of Tennessee, Knoxville), but more research is needed in order to compile enough data so that the current anthropological methods can be adapted to modern populations. This is where virtual anthropology can be of great use.

As it has been already stated, virtual anthropology offers the possibility of deriving modern skeletal collections from living populations. However, as with any method in its infancy, there are practical concerns that have to be addressed before virtual anthropology can be used for this purpose.

The aim of the present study is to contribute to the widening knowledge of the application of 3D technologies to further anthropological goals. There is already some evidence that the use of 3D imaging techniques applied to living individuals can result in the virtual extraction of human skeletons that can then be analyzed using standard anthropological methods (i.e., sex, age, stature, and ancestry estimation from bones). However, the studies published to date have generally worked with low sample sizes and have mostly left the issue of interobserver error (i.e., the way the results obtained will differ between observers when applying the same method to the same bone) at a wider scale that goes beyond the observations recorded by the authors working on a given study largely ignored, so that data concerning a larger sample of observers is currently lacking (Decker *et al.* 2011; Ferrant *et al.* 2009; Grabherr *et al.* 2009; Telmon *et al.* 2005; Uldin *et al.* 2012; Villa *et al.* 2013b; Wink 2014). Furthermore, the studies have used different settings for 3D imaging as well as visualized the 3D images using different software packages, rendering the results incomparable with each other. No research, for example, has yet been conducted on the optimal settings to be used for the purpose of extracting data from living individuals.

The present study adds to the body of knowledge obtained from previous research by analyzing 130 specimens from the Boston University Donated Osteological Collection (BUDOC) using 3D imaging, then visualizing the resulting 3D copies and applying two commonly used anthropological methods: the Phenice (1969) method of sex estimation and Brooks and Suchey (1990) method of age estimation. Furthermore, a group of six observers of various levels of expertise (beginner, intermediate and expert) were recruited for scoring a subsample (n = 20) of the specimens to explore the degree of interobserver error that can be expected from scoring 3D images using methods devised for dry bones. In this way, this study explores the agreement between dry bone scores and 3D bone scores.

The three specific questions to be addressed in the present study are:

1. The degree to which the author's scores derived from dry bones agree with scores recorded from the corresponding 3D images of the bones (n = 130).
2. The degree to which six observers will agree with the scores of their peers when divided into experience levels (n = 20).
3. The degree of agreement between the scores from dry bones and the scores from the corresponding 3D bones as recorded by each of the six observers (n = 20).

Based on the results obtained by Villa *et al.* (2013b), it is expected that assessment following the guidelines in Phenice (1969) will yield a higher degree of agreement between dry and 3D bone scores compared to results obtained when applying the Suchey-Brooks method for age estimation (Brooks and Suchey 1990). Furthermore, it is expected that the degree of agreement between and within the results obtained by the six observers that participated in the interobserver error portion of the present study will depend on the level of expertise of the observers in question. In short, it is expected that agreement will increase with the level of experience of the observers in question.

CHAPTER 2: PREVIOUS RESEARCH

The use of 3D imaging in human anatomical research has become very popular in fields outside of anthropology, particularly in medicine. The focus of much research has been in developing faster, more cost-effective methods of gathering anatomical data from living people, often for the purpose of diagnosing congenital and acquired craniofacial malformations and for developing new data sets for the industrial design industry, particularly for clothing, furniture, working gear, etc. (see Enciso *et al.* 2003; Enciso *et al.* 2004; Hammond *et al.* 2004; Lu and Wang 2008; Ozsoy *et al.* 2009; Wong *et al.* 2008). Researchers have also explored the potential applications of 3D imaging techniques for the purpose of training future surgeons (Jabir 2012; Shetty 2006), reconstructing highly comminuted bone fractures (Willis *et al.* 2007; Zhou *et al.* 2009), detecting blood diseases (Salih *et al.* 2004), and predicting the biomechanical strength of the proximal femur in order to prevent possible hip fractures in individuals with osteoporosis (Baum *et al.* 2010).

Jones and Rioux (1997) report a wide range of applications of 3D imaging technology in other disciplines, ranging from orthodontics and orthopaedics, to breast cancer detection, human motion analysis, virtual reality, and communications. The utility of 3D facial image analysis for human identification for security purposes has also been explored by Nagamine *et al.* (1992). Many of the studies referenced stress the utility of no-contact scanning methods, both for their high accuracy in data recording and for the lessened discomfort experienced by the human volunteers from whom the data are being derived. Ergonomics is also a field where the use of 3D technologies is being explored in order to develop products designed for particular region-specific target markets (Figure 2.1) (Ball *et al.* 2010:837).

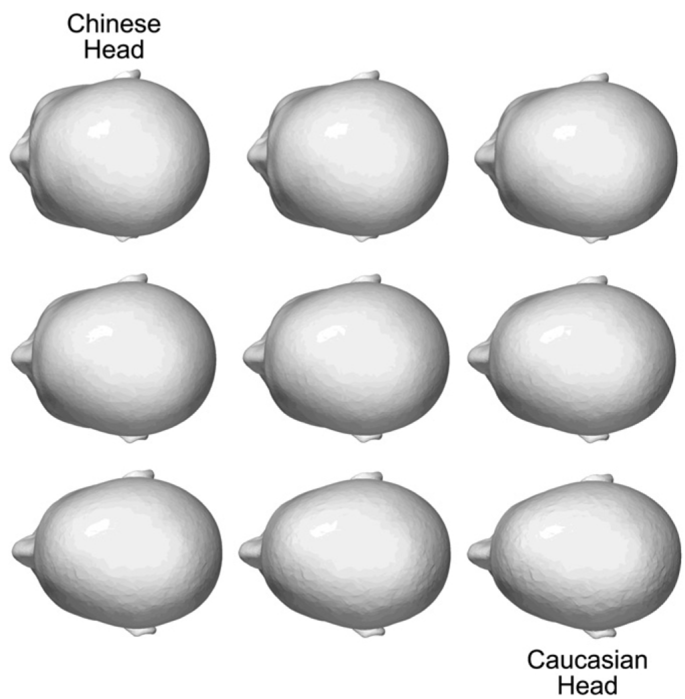

Figure 2.1. Different head shapes obtained using 3D laser scans from SizeChina (a database for adult Chinese individuals) and the Civilian American and European Surface Anthropometry Resource (CAESAR). The study of morphologic variation between populations has been used to design products that are better suited for different target markets. After Ball *et al.* (2010:837).

3D Data Collection

There are various ways to gather data for 3D visualization, depending on the type of structures that need to be observed. One type of data collection method consists of gathering only the information pertaining to the surface of the object to be scanned, a method commonly used for rapid prototyping (i.e., 3D printing). The other type of data collection method scans the whole specimen so that inner structures may be visualized. The method to be used must be selected in function of the research design and the nature of the specimens themselves. Some types of data collection methods are discussed below.

Surface Scanners

These types of scanners record only the surface shape of the specimen in question, although they may also record the texture and color, depending on the particular scanner being used. There are two broad categories of surface scanners. The first, a contact scanner, is commonly used to record specific points in an object, such as in the case of an anthropological digitizer (a piece of equipment used to calculate common anthropological measurements such as the distance between the eye orbits, the width of the nose, the breadth of the cranium, etc.), while the second category, a contactless or contact-free scanner is able to record the whole object by creating a point cloud of its geometry with reflected rays of light (Weber and Bookstein 2011:95). Surface scanners may be suitable in the event where a researcher wishes to digitize dry bones, since there is no soft tissue covering the specimen and it can be scanned as is. Surface scanners would theoretically be the best alternative for digitizing entire osteological collections in order to make them accessible over the internet.

MRI

Magnetic resonance imaging (MRI) produces volume data (i.e., data that can be represented as 3D objects, Webber and Bookstein 2011:49) by virtually 'slicing' the specimen being scanned (Weber and Bookstein 2011:59). The technique is based on the principle of nuclear magnetic resonance and can scan an individual by generating a very strong magnetic field that has no adverse side effects for patients (Weber and Bookstein 2011:60) which makes it very suitable for extracting the skeletal data from a living individual.

CT scanners

Computed tomography (previously referred to as computed axial tomography, or CAT scan) is used to produce an image of a subject that contains volumetric data. Similar to what would take place when taking a radiograph, CT imaging subjects the specimen in question to radiation and then records the resulting attenuation profile, or the degree to which rays were absorbed by different structures within the person or specimen being scanned (Allard 2006:22). The resulting data can be saved in a variety of formats, both two-dimensional (2D) and 3D, including as a Digital Imaging and Communications in Medicine (DICOM) file, which is a format established for medical applications that includes a set of standards for handling the information it contains (Webber and Bookstein 2011:79). Settings that can be customized for a particular specimen include the radiation dose to be used (i.e., the voltage, amperage, and exposure time that will be needed depending on the type of specimen, Webber and Bookstein 2011:55), the thickness of the slices to be scanned (e.g., 0.625 mm, 1.25 mm, etc.), the set increments in which the slices will be taken and the number of images to be scanned (in order to produce bigger or smaller files as needed). The reason why a slice thickness and slice increment must be selected is that scanning the entire length of a specimen would potentially take a long time and result in a large file, so only some parts of the specimen are selected for scanning, e.g., for every 0.625 mm of a specimen's surface, a slice 0.625 mm in thickness is taken. A group of settings that are used for scanning a particular specimen is commonly referred to as a scanning protocol.

Unlike radiography, which uses a single dose of radiation to generate a static, 2D view of a subject, the use of CT imaging technology takes many individual radiographs from different angles of the subject to be studied; these are subsequently combined to form the 3D image (Brenner 2009). However, this results in an increased exposure to x-ray radiation to patients. Although CT technology has revolutionized the field of medical imaging, and while its use is becoming more common in medicine, some researchers are concerned of the effects that this radiation exposure has on patients, particularly children (Brenner 2009:63).

It is for this reason that some investigators, such as Wink (2014), have undertaken retrospective studies, which are studies done that examine data that were not originally collected by the researcher in question. Under this research design, no patients are recruited to be scanned using CT imaging. Instead, stored scans are obtained from a medical facility (in the case of Wink 2014, Boston Medical Center). All identifying information is removed except for information relevant to the research (e.g., age and sex). Once skeletal tissues are digitally isolated from the rest of the body, the researcher effectively has a documented skeletal collection from a modern sample with which to work without the risk of exposing patients to more ionizing radiation than they had already agreed to or required.

The advantage of this type of study is that there are no scanning costs involved (although there may be fees for being allowed access to patient information, and the required paperwork for research done on live human subjects must certainly be submitted and approved before the start of the study). The software required to process CT images (and to visualize 3D images from them) will range in cost but will sometimes be available for free under the GNU General Public License (GNU GPL), which is a copyleft license for software and other works often referred to as 'open-source' (to be contrasted with a copyright license, which does not allow for the free use or modification of a product without incurring royalty fees). Thus, the creation of new, modern virtual skeletal collections for use in research would potentially incur a very low cost, if any, provided that traditional methods can be effectively and confidently applied to 3D images.

μCT Scanners

A subtype of CT scanners, μCT scanning provides the researcher with the possibility of visualizing specimens at a much higher resolution than is possible with a conventional CT scanner (Stock 2008:1). This means that apart from higher quality 3D images, researchers can explore questions relating to virtual histology (i.e., microscopic anatomy in 3D) without the need to section or otherwise damage the specimen (Boyd 2009:1). μCT could potentially provide anthropologists with skeletons derived from modern populations that are accurate; however, the serious drawbacks to this data collection method outweigh the potential benefits. The first drawback is that the machinery required for μCT scanning is very often not easily accessible to researchers (Weber and Bookstein 2011:59). While a researcher affiliated with a hospital or medical school can potentially make use of a CT scanner for their daily research (provided the scanner is available, as priority is given to living patients), not many researchers have similar access to a μCT scanner. If they do, they encounter a second drawback: time and money constraints. It can take hours to scan a single specimen at a resolution of 100 μ (for reference, there are 625 microns in 0.625 mm, which is the finest resolution for a regular CT scanner) and scanning of a single specimen can cost hundreds of dollars (Dr. Darlene R. Ketten, *pers. comm.*), which may be reasonable if a researcher intends to scan few specimens. However, if the researcher wishes to scan an entire skeletal collection, the costs involved skyrocket. Furthermore, the physical window for scanning a specimen is quite small and can only fit a human skull in the very best case scenario (Weber and Bookstein 2011:57), which means that scanning complete long bones and innominates is out of the question. And once a researcher scans a specimen, there is an added limit in that the data volumes involved can be gargantuan, which puts a strain on computers and storage media (Weber and Bookstein 2011:59). Weber and Bookstein (2011:59) note that increased demand in the medical sector might result in better μCT scanning equipment being developed, but for the time being, the use of this data collection method for digitizing full skeletons is not feasible. It does, however, have the potential of being of use if applied to the field of dental anthropology.

Visualization Software

The particular software used for 3D image reconstruction may affect the quality of the rendered image itself. There exists a variety of software packages devised for medical imaging and educational purposes that differ depending on what operating system they were designed for, whether they can be used for surface rendering or volume rendering (or both), what features they offer, and their cost (Martin 2012). An example of this type of software is *OsiriX* (Pixmeo, Geneva, Switzerland), which is protected under the GNU GPL and is thus free for researchers working with Macintosh computers. So far, a review of different software packages for anthropological use has not been published. Guyomarc'h *et al.* (2012) suggest that, because the goal of these medical imaging software packages is, first and foremost, to visualize organs for assessing them as part of patient diagnostic procedures, an appropriate algorithm for the visualization of bone tissue is needed if said software is to be used for anthropological purposes. It is very likely that, before the use of 3D technologies is widespread in forensic anthropology, at least a basic degree of experimentation and standardization of the way 3D images are visualized will be needed.

Attention should also be paid to a feature of medical imaging data that could potentially help or hinder 3D osteological research (Figure 2.2). The quality of the images obtained can be greatly influenced by the chosen reconstruction algorithm or parameter (also known as a filter or kernel): a soft tissue reconstruction parameter (with a kernel of ~30) will instruct a computer to smooth the 3D model's surface; a hard-bone reconstruction parameter (with a kernel of ~90) will result in the emphasis of finer details such as the appearance of cranial sutures (Grabherr *et al.* 2009; Weber and Bookstein 2011:95).

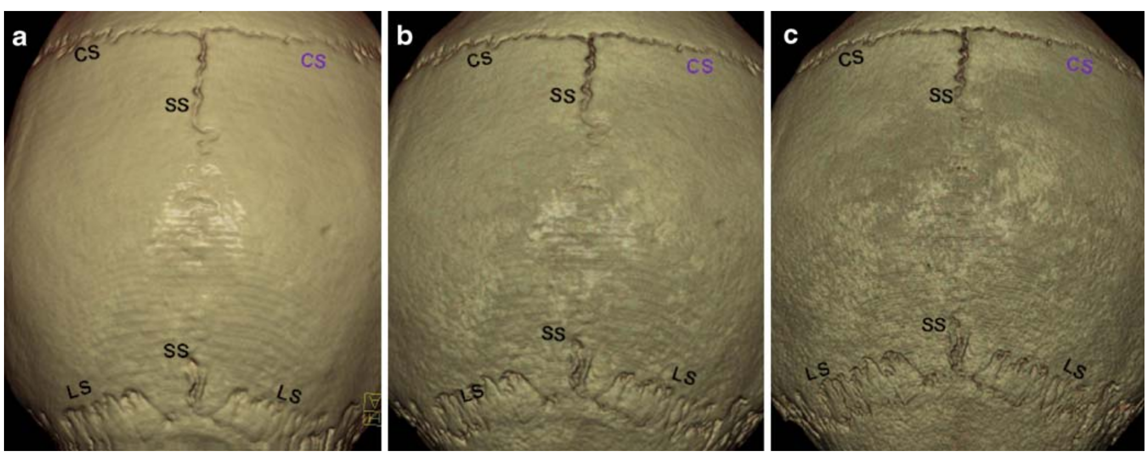

Figure 2.2. A 3D model of the same skull, scanned with the same resolution but calculated from different reconstructed data. (a) soft tissue reconstruction parameter (kernel of 30), (b) parameters for reconstruction between soft tissue and bone (kernel of 50), (c) parameters for reconstruction of hard-bone (kernel of 90). Grabherr *et al.* (2009:427).

Previous Research In Archaeology, Anthropology, and Forensic Science

Within the realm of archaeology, virtual anthropology methods have been used to reconstruct the appearance in life of Ferrante Gonzaga (an important figure in Italy during the Renaissance; Benazzi *et al.* 2010), as well as that of the famous Italian poet Dante Alighieri (Benazzi *et al.* 2009) for the purpose of assessing the accuracy of available portraits for the two historical figures by a superimposition comparison of the portraits and the craniofacial reconstructions derived from the skeletal remains of the individuals (Figure 2.3). The studies, however, achieved mixed results.

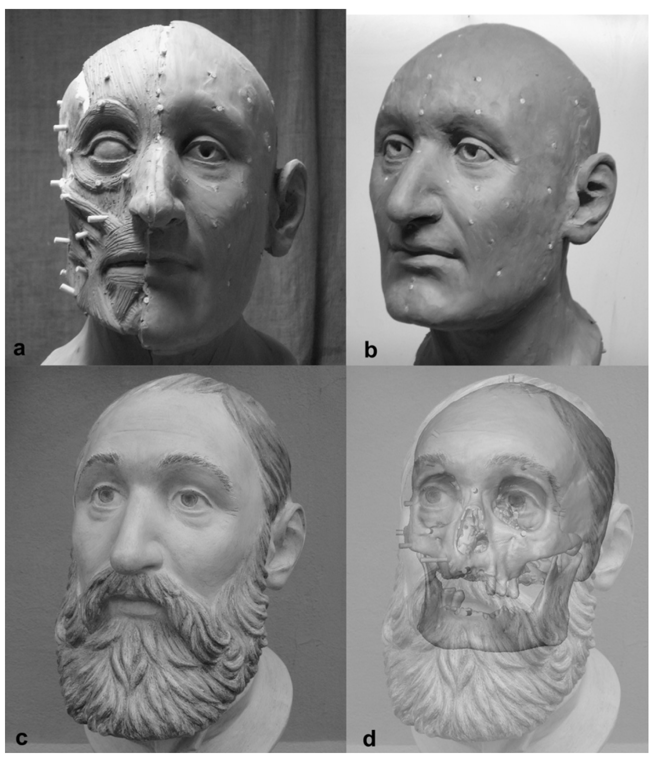

Figure 2.3. Craniofacial reconstruction of Ferrante Gonzaga. The authors obtained a CT scan of the individual's cranium, then a physical model was made using rapid prototyping. Traditional methods in forensic facial reconstruction (A-D) were then applied to the skull. Benazzi *et al.* (2010:1576).

Within forensic science, 3D imaging has proven quite successful in the study of trauma and weapon identification (Sansoni *et al.* 2009; Thali *et al.* 2003a, b, and c; Wozniak *et al.* 2012), where researchers have reported many advantages over traditional casting methods. One study in particular by Kettner *et al.* (2011) suggests the use of 3D printed models (a type of rapid prototyping) derived from 3D images used for trauma studies, which can also be presented as demonstrative evidence in court (Figure 2.4). Their suggestion, however, will not likely be taken as an alternative to current methods used in court due to the high costs which rapid prototyping currently entail.

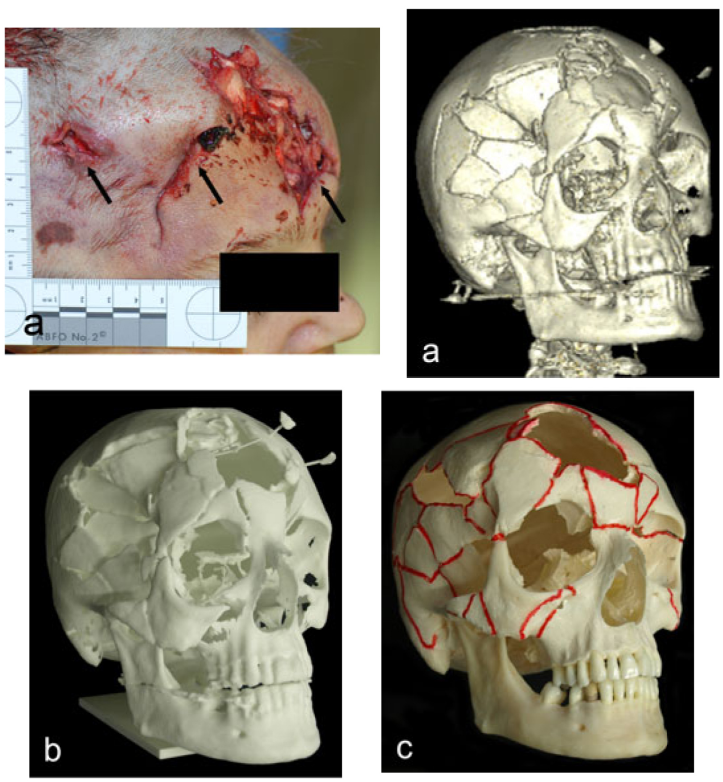

Figure 2.4. Trauma reconstruction using images derived from the victim's skull. Upper left: right frontotemporal photograph of the head with three distinct lacerations (arrows). (A) right frontotemporal view of the 3D volume reconstruction (OsiriX V3.3.2), (B) rapid prototyping/selective laser sintering model, (C) reconstructed macerated skull. Adapted from Kettner *et al.* (2011:1016).

3D technology has also been suggested for the documentation of crime scenes using contact-less scanners for the purpose of analyzing both the scenes themselves and related tissue wounds for medicolegal death investigations (Sansoni *et al.* 2007; Sansoni *et al.* 2011). Attempts have been made to apply the technology to identify individuals caught on 2D media by comparing a 3D model of their faces to the incriminating evidence (De Angelis *et al.* 2009). Following the authors' rationale, one could apprehend a suspect, derive a 3D image of his or her face, and then compare it to a video of a perpetrator caught in a surveillance camera while a crime was taking place, which could then be potentially used as evidence in a trial if the technique proves to be robust enough.

In the field of forensic anthropology in particular, 3D cranio-facial reconstruction (i.e., software-based reconstruction of a face from a skull) for aiding forensic identification appears to be the most popular application of 3D technology. The reason for this popularity seems to stem from the assertion that 3D technology constitutes an alternative approach to traditional methods (i.e., those made using clay, thus requiring artistic talent but incorporating a degree of subjectiveness to the reconstruction) that may potentially remove the artistic element from what many authors state should be a scientific endeavour (e.g., De Greef and Willems 2005). However, the studies agree that further research involving the depths of facial musculature is needed in order to sharpen the software's ability to provide an accurate depiction of individuals, and as such, the digital methods are far from complete.

Research in the particular issue of determining a biological profile from 3D images derived from bones is a relatively recent development in forensic anthropology, and the studies that are currently available have obtained mixed results. Telmon *et al.* (2005) applied the Suchey-Brooks method to 3D images derived from the pubic symphysis and found the method not to differ significantly from traditional approaches (Figure 2.5). However, their study only included 21 specimens, and their methods are questionable because they scored the bony elements using still images (four different views: front, anterioposterior, ventral, and dorsal) of the 3D images that they had obtained, an approach that defeats the purpose of using 3D images in the first place. Uldin *et al.* (2012), however, found that their own method using multi-detector computed tomography (MDCT, a type of CT scanner that can capture multiple cross-sections at the same time) scanning did not allow for sufficient resolutions to adequately derive a biological profile from their 3D images, and suggested that a specialized scanning protocol be devised for this purpose.

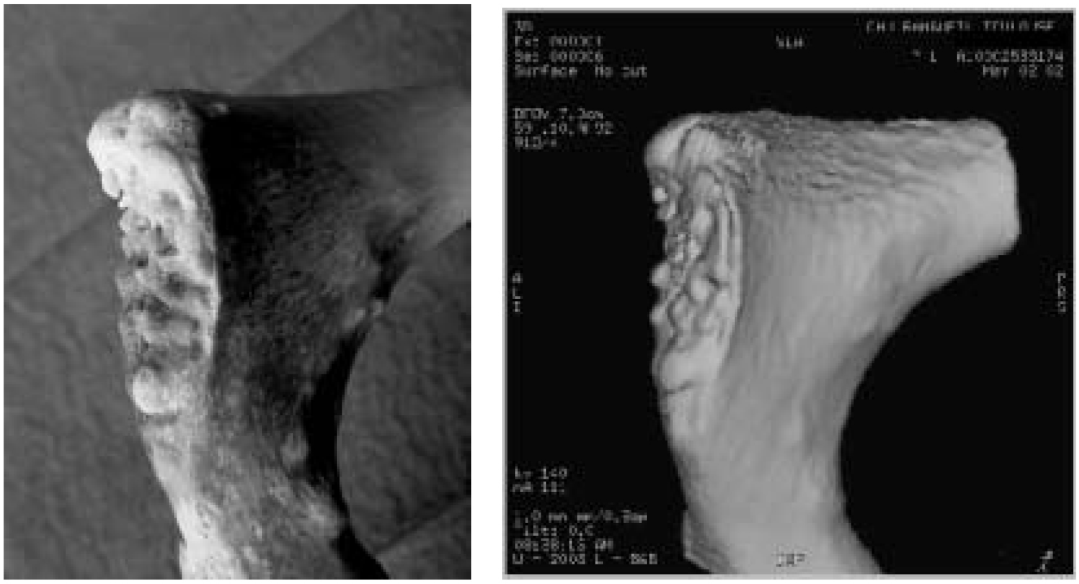

Figure 2.5. A comparison of a dry pelvic bone and the 3D image derived from it. Suchey-Brooks Phase II: 21 year old male, 3/4 dorsal view (Telmon *et al.* 2005:3).

Other researchers obtained much more encouraging results. Zech *et al.* (2012) attempted sex determination from the sacrum using postmortem CT scans of 95 corpses (49 men and 46 women). Two observers performed measurements on axial images with a slice thickness of 1.25 mm. Sex was determined using discriminant function analysis on the first sacral vertebra. The authors found no significant differences between the measurements taken from the CT images and caliper measurements performed on the same bones. However, they do not recommend the use of sacral measurements using postmortem CT scans for sex determination due to the method itself being variable depending on the target populations it is applied to, as well as evidence that sex estimation using the sacrum involves a wide overlap between the sexes. Villa *et al.* (2013a) devised a method to estimate an individual’s age from postmortem CT scans by scoring trabecular bone changes in the pelvis. Their sample consisted of 319 CT scans (145 females and 174 males) taken within three days of the individual's death. The authors found low interobserver error and low inaccuracy using three different image settings that varied in slice thickness and slice increments (i.e., there was no statistically significant difference in the settings used). They observed that individuals age 60 years and older were evaluated with better accuracy, compared to the macroscopic methods commonly used in anthropology and argued that their method is a valid alternative to the often time-consuming and potentially destructive process of bone maceration for the study of defleshed bones. Lottering *et al.* (2014) developed standardized protocols for the measurement of subadult crania using 3D images derived from multi-slice computed tomography (MSCT, a type of CT scanner that allows for a higher number of slices to be taken in order to increase the image resolution), obtaining high levels of repeatability between five observers of varying expertise and experience with imaging software. The authors state that, while their present study focused on subadult crania, standardized protocols are currently being derived from this study for application in postcranial skeletons, with the aim of constructing population-specific standards for human identification, due to the lack of documented skeletal collections available for research in the region of the world in which they currently work (Australia). This highlights the possibility of using virtual anthropology as a tool for constructing skeletal samples that are derived from modern, region-specific populations.

Other researchers have focused on the validation of the use of 3D imaging technologies for the estimation of biological profiles. Brough *et al.* (2013), for example, developed a standardized scanning protocol that was applied to 33 juvenile clavicles from the Scheuer Juvenile Skeletal Collection (housed in the Centre for Anatomy and Human Identification at the University of Dundee, in Scotland) using MDCT. The authors argue that because there are no currently established scanning protocols and guidelines on which virtual measurements are most valuable for the successful identification of human remains, it is important to take a step towards standardization of the use of 3D technologies for forensic practice. They also suggest future research that may aid in the recognition of these new approaches as legally admissible anthropological techniques. In a retrospective study, Bilfeld *et al.* (2012) applied geometric morphometrics methodologies to 3D images of 65 Caucasian adults (30 males, 35 females) obtained using MSCT imaging at their institution in Toulouse, France. The authors found intra- and interobserver variabilities lower than 3% and dimorphic results that are in accordance with more traditional (i.e., dry bone) studies, as well as a higher sexual dimorphism for some bones and bone complexes (i.e., groups of bone features) as observed in the 3D images compared to the original bones. They show that their method is reliable, and their plans for future research include a higher sample of adults, inclusion of juveniles in the sample, and sexual dimorphism analysis using type III landmarks (points defined relative to other, distant structures, Hallgrímsson *et al.* 2011). Robinson *et al.* (2008) scanned 15 human lower limbs in various states of decomposition using MDCT in order to obtain traditional bone measurements from 3D images. The measurements were taken by four individuals with various levels of expertise (a radiographer, a radiologist, a trainee forensic pathologist with little experience in anthropological or radiological measurements, and a forensic anthropology graduate). Their findings revealed that there was no significant difference between measurements taken using the images from the MDCT scans and measurements taken once the bone was defleshed. The limbs were scanned through sealed, labeled bags, which in the opinion of the researchers is a good way of avoiding any visual unpleasantness, as well as health and safety issues that would be involved with handling the remains using plain x-ray and fluoroscopy. They recommended that bones be scanned at a slice thickness of no more than 1.25 mm for best results. The authors tested the amount of time it would take to scan the bones, upload them to the internet, have a researcher download them at a different site from the scanning site, measure them, and report their findings (what they have termed "tele-anthro-radiology"; Robinson *et al.* 2008:1289). They found that the average time needed for this process was between 20 and 30 minutes following the scanning (which took around 5 minutes). The authors indicate that this method would allow the international community to "build up and maintain population data with regards to radiological osteological assessments as well as providing the facility for standardization of protocols, international peer review, and quality assurance schemes" (Robinson *et al.* 2008:1295).

More recently, researchers working at the University of Granada, Spain, have announced that they have developed a method using CT images that is able to determine the age and sex of a corpse with a reliability of 95% using histogram analysis; while their research has yet to be published, their results were announced by the university at a popular science news outlet (University of Granada 2013). Interestingly, the researchers announced that they obtained excellent results in estimating the age of individuals older than 50 years of age. Research done by Tocheri *et al.* (2002) in alliance with the Partnership for Research in Spatial Modeling (PRISM) at Arizona State University have explored the possibility of applying quantitative analyses to 3D images of the Suchey-Brooks casts, in order to explore the age-related changes to the pubic symphysis as described by Brooks and Suchey (1990). The authors hope that, by making osteological reference samples widely available with the help of 3D technologies (for which the scanning of the casts was a test-run), and by continuing to refine statistical techniques for osteological data, there will be an increase in the understanding of the way age-related changes to the pubic symphyseal face occur, enabling the refinement of the Suchey-Brooks method. By understanding age-related changes to the pubic symphysis, researchers would then be able to adjust the method to specific populations. The need for this type of research stems from findings indicating differences in the way age-related changes manifest depending on the population being studied; thus, some populations will exhibit certain changes (as defined by the Suchey-Brooks method) earlier, later, or the changes might take longer or shorter to progress or complete (Hoppa 2000). It is for this reason that population-specific methods are needed in order to refine the ability of forensic anthropologists to arrive at a biological profile.

The Phenice (1969) and Suchey-Brooks Methods

In recent years, researchers have explored the degree to which the Phenice (1969) method of sex estimation and Suchey-Brooks (Brooks and Suchey 1990) method of age estimation can be applied to 3D images of skeletal remains, since these are two of the most widely used methods for deriving biological profiles from individuals in forensic anthropology. Such studies have often used 3D images obtained from the innominate bones of cadavers, have involved tests of inter- and intra-observer reliability, and have obtained results that greatly depended on the scanning protocols they used for data collection, as well as their sample size.

Sex Estimation Using Phenice (1969)

The Phenice (1969) method consists of the scoring of three nonmetric morphological traits that generally distinguish between male and female individuals: the ventral arc, subpubic concavity, and ischiopubic ramus (Figure 2.6). The possible scores for each trait are (1) female, (2) indeterminate, and (3) male. The method was derived from a sample of 275 adults belonging to the Terry Skeletal Collection, housed in the National Museum of Natural History of the Smithsonian Institution. Researchers applying the method to various target samples have reported accuracy rates of 83-95% (Buikstra and Ubelaker 1994; Lovell 1989; Ubelaker and Volk 2002).

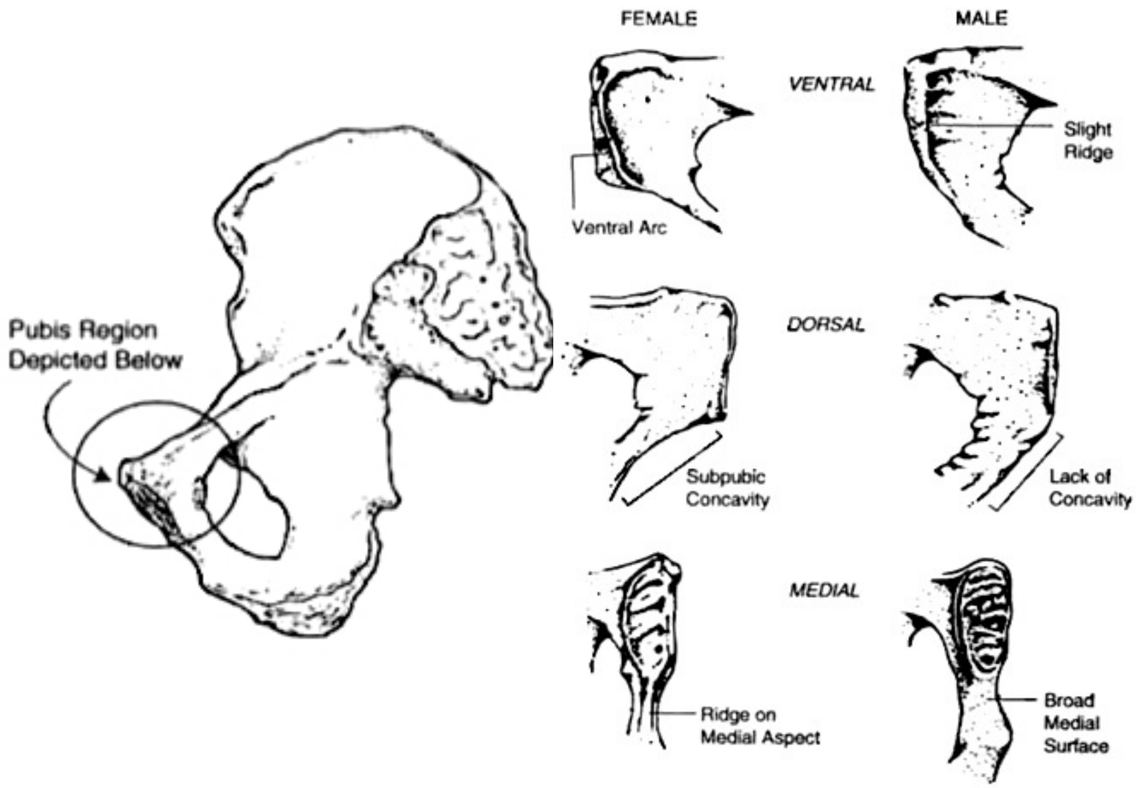

Figure 2.6. The Phenice (1969) method of age estimation relies on the scoring of the ventral arc, subpubic concavity, and medial aspect of the ischiopubic ramus. Adapted from Buikstra and Ubelaker (1994:17).

A retrospective study by Decker *et al.* (2011) analyzed abdominal CT scans of 100 individuals of known sex using both obstetric measurements and metricized versions of the Phenice (1969) nonmetric traits (i.e., they expressed the Phenice nonmetric traits in terms of groups of measurements to decrease observation subjectivity when scoring them). The scans were set at a 1.25 mm slice thickness, and *Mimics* (a visualization software similar to *OsiriX*, developed by *Materialise* in Leuven, Belgium) was used to analyze the 3D images. The authors developed a four-variable equation for sex estimation derived from the obtained measurements and found that their equation correctly classified their cases 100% of the time, with a reported inter-observer error of 2.2%. They input their measurements into *FORDISC 3.0* (a computer program commonly used in anthropology to estimate sex, stature and ancestry, developed by Jantz and Ousley 2005) and compared those results to the results obtained using their equation; they found that, overall, *FORDISC 3.0* was able to classify their specimens 86% of the time, although the classification rate for males only was 67.5%. The authors stated that, by including their metricized nonmetric Phenice traits in the variables used by *FORDISC 3.0* for sex estimation, classification rates should increase.

Age Estimation Using Brooks and Suchey (1990)

The Suchey-Brooks method for phase-based age estimation was developed by Brooks and Suchey (1990) using a documented sample of 1,225 pubic bones obtained from the autopsies of modern individuals ranging from 14 to 99 years of age at death at the Office of the Chief Medical Examiner, in the County of Los Angeles. The method provides guidelines for scoring age-related changes to the pubic symphyseal surface of both males and females, dividing these changes into six stages (I-VI) and providing an estimated age range for each phase.

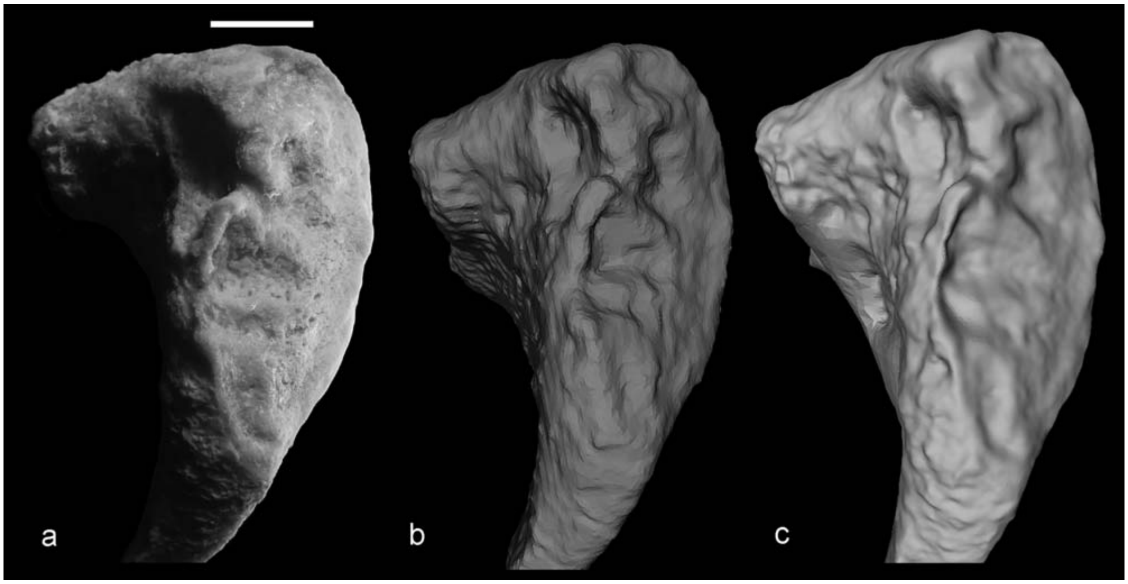

Figure 2.7. Comparison between (a) dry bone, (b) 3D reconstruction from a CT scan, and (c) 3D visualization from a contact-free laser scan. Scale is 1 cm. After Villa *et al.* (2013b:161).

Studies that applied the Suchey-Brooks (1990) method to 3D images have obtained mixed results. Villa *et al.* (2013b) applied the Suchey-Brooks method to 3D images derived from both CT and laser (surface) scans of twelve pubic bones from cadavers obtained from autopsies at the University of Milan, Italy (Figure 2.7). The CT images were scanned using 0.75 mm slice thickness at a 0.3 mm slice increment interval and visualized in 3D using *Mimics* imaging software. The authors found that the Suchey-Brooks age estimation method can be applied to 3D images but at the cost of a lowered accuracy compared to its use on dry bone. Furthermore, they found that images that were obtained using laser scanning had a higher level of quality and were easier to score compared to images obtained using CT scanning. They also note that micro- and macroporosity, as well as surface texture, were very difficult to score.

Figure 2.8. 3D reconstruction of the pubic symphysis belonging to a 55-year-old male (Phase IV). After Wink (2014:4).

Wink (2014) scored 3D images derived from CT scans of 44 living individuals of known age and sex using the Suchey-Brooks method and found that the true age of a scored individual fell into the estimated age ranges in 79.5% of cases (Figure 2.8). The bones were scanned using a 1.25 mm slice thickness, and the data were visualized using the 64-bit version of *OsiriX* Imaging Software. Wink (2014) also found a good intra-observer reliability, but inter-observer reliability was not scored in her study. Ferrant *et al.* (2009) analyzed 33 innominates of adult males of known age obtained at autopsy, with the purpose of converting macroscopic criteria into variable measurements that can be quantified and thus scored (for example, the symphyseal dorsal ridge was measured using the angle in degrees created by the symphyseal surface and the symphyseal posterior edge; porosity was measured using osseous density; Ferrant *et al.* 2009:15). The images were reconstructed using a slice thickness of 0.9 mm and a slice increment interval of 0.45 mm, and viewed in 3D using the CT scanner's console CT viewer software. They found that their CT scan criteria data were correlated to age-at-death, although they admit that their study was limited in great part due to their small sample size.

A study by Grabherr *et al.* (2009) incorporated both age and sex estimation in the scoring of 3D images derived from cadavers (Figure 2.9). MDCT was used to scan various regions of the body of 22 deceased individuals of known age and sex, and the resulting images were analyzed by three anthropologists with varying levels of experience. A slice thickness of 0.63 mm with a slice increment of 0.5 mm was selected, and reconstructions were done using a kernel of B30 (for soft-tissue reconstruction) and B90 (for bone reconstruction).

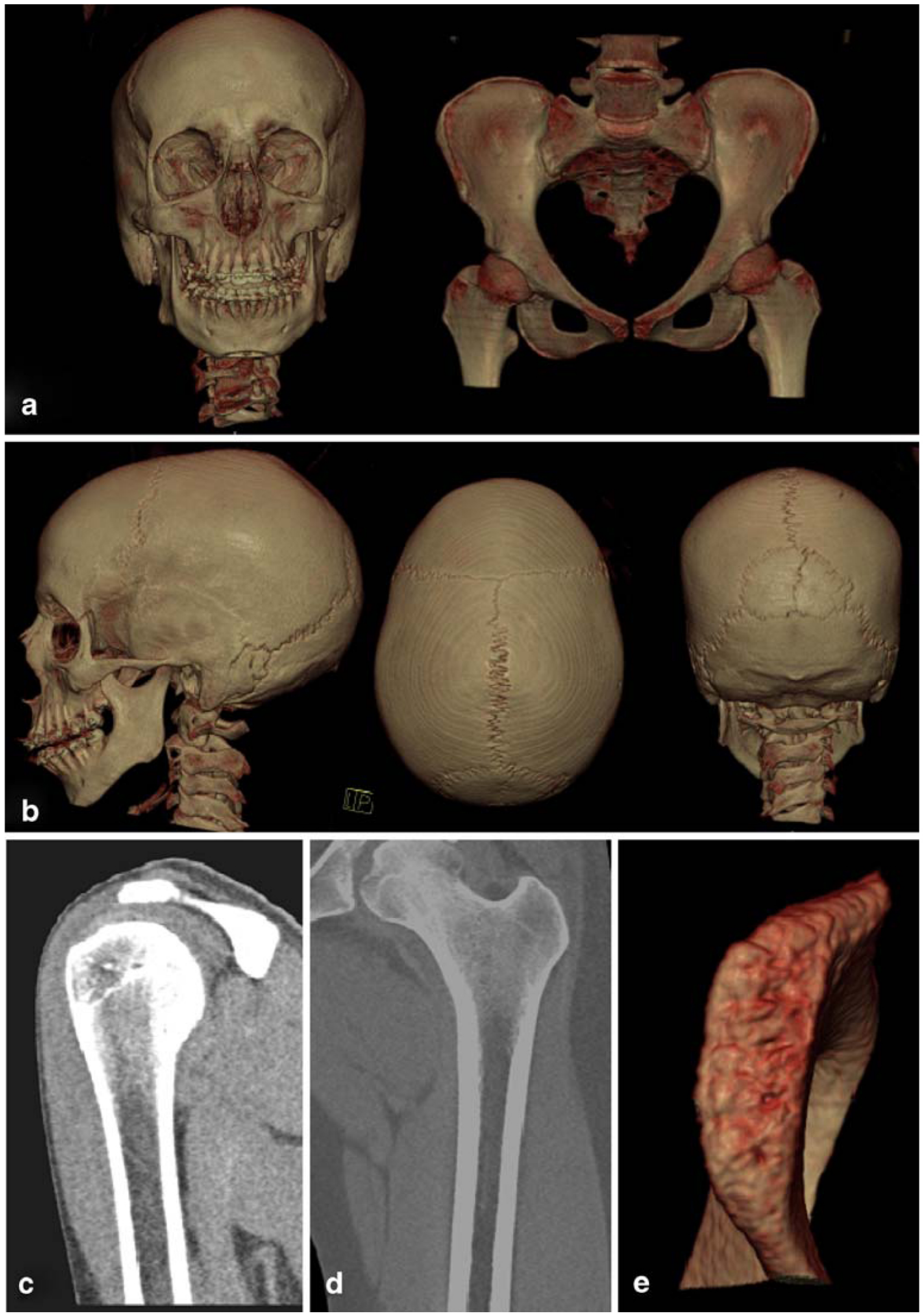

Figure 2.9. 3D reconstruction of the skeleton of a 17-year-old. Reconstruction was done using a B90 kernel. (a) frontal view of skull and pelvis, (b) lateral, superior, and posterior views of the skull, (c) and (d) show 2D visualizations of the CT images, (e) 3D reconstruction of the pubic symphysis. After Grabherr *et al.* (2009:422).

Grabherr *et al.* (2009) used MDCT to examine the skull and pelvis for sex estimation, and the endocranial sutures of the skull, the trabecular structure of both humeri and femora, and the pubic symphysis for age estimation. They encountered no issues with the accuracy of sex estimation; however, age estimation was found to be more difficult, an issue that has also been noted by Villa *et al.* (2013b). A difficulty stemming from a lack of experience in evaluating 3D images was reported by Grabherr *et al.* (2009). They also reported an issue with the appearance of artefacts due to the presence of gas bubbles within bone marrow cavities on some of the decomposing bodies, which was not observed in any of the other studies reviewed here that used cadavers as their source of data. Much like other researchers, Grabherr *et al.* (2009) praised the ability of 3D imaging techniques for allowing researchers to study skeletal remains with no damage to the original specimens, the possibility for multiple researchers to analyze the same specimen regardless of geographic constraints, the avoidance of time-consuming and potentially damaging maceration procedures (an issue also noted by Villa *et al.* 2013a), the chance for interactive data exchange, and thanks to the increasing application of MDCT in the aid of victim identification, the potential use of 3D technologies in cases of mass disasters.

Methodological Considerations

As with any research method, it is important to keep in mind the strengths and weaknesses of 3D imaging technologies. Time constraints, data storage and preservation, the type of 3D software used, algorithms for deriving 3D images from scans, as well as the expertise of the researcher, all have an effect on the end product achieved using these methods (Kuzminksy and Gardiner 2012). Investigators should keep in mind that different scanning methods will yield different results. A no-contact surface 3D scanner will create a surface model of a bone, whereas the use of CT scanning will allow both outer and inner structures to be viewed. Abel *et al.* (2012), for example, used μCT scanning technology to derive a model of a prehistoric shark tooth which showed both the surface details as well as the tooth's inner blood supply (Figure 2.10). A surface scanner, on the other hand, only registers the surface form of an object, with a smaller data volume and the option of recording the object's texture and color as well as its shape (Weber and Bookstein 2011). Friess (2012) has suggested the use of surface scanning for measuring external morphology, citing the possibility of a higher measurement precision associated with the use of this technology. The type of scanning technology to be used for research will thus depend on the type of data to be collected.

Thus, in order to use 3D imaging technologies effectively, it is vital to consider the type of scanning method appropriate for the source material in question so that it will yield appropriate results.

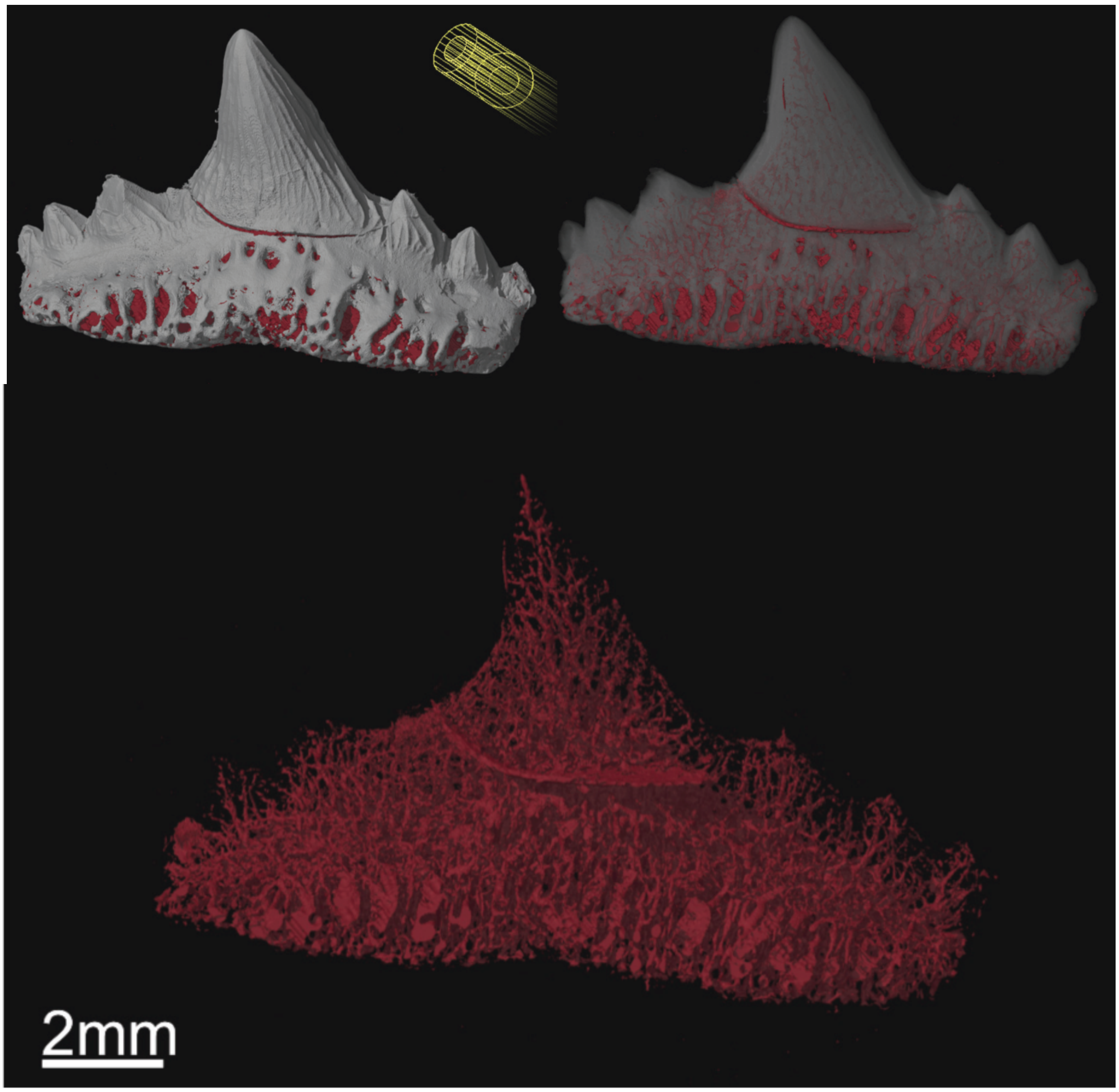

Figure 2.10. 3D reconstruction of a µCT image of a shark's fossilized tooth (*Sphenacanthus hybodoides*). CT-3D technologies retain volumetric data, allowing researchers to observe inner structures of objects that would otherwise be unavailable for study. Adapted from Abel *et al.* (2012:7).

CHAPTER 3: METHODS

The aim of the present study was to assess the degree of agreement between scores obtained using the Phenice (1969) and the Brooks and Suchey (1990) methods when applied to dry bones and their corresponding 3D versions. It was also of interest to assess whether the results are comparable in terms of accuracy and inter- and intraobserver variability. The specific questions addressed were the degree of agreement between scores obtained from dry bones and their corresponding 3D bones as recorded by the author of the present study and the degree of agreement between scores obtained from a subsample of 20 specimens by a group of observers divided into experience levels. It was hypothesized that the agreement between dry and 3D bone scores would be high but lower than expected due to the fact that the methods being used are meant to be applied to dry bones only.

A sample of 147 specimens consisting of left and right innominates was obtained from the Boston University Donated Osteological Collection (BUDOC) and from the teaching collection available at Boston University Medical School for gross anatomy teaching. From that group, a total of 130 specimens were selected based on their consideration as good candidates for scoring of the Phenice (1969) traits and the pubic symphysis using the Suchey-Brooks method (Brooks and Suchey 1990). The criteria for inclusion consisted mostly of the completeness of the pubic bone and pubic symphysis, and thus any specimens in the collection that were considered too damaged for assessment using standard osteological methods were not considered. Since most specimens had previously been used as anatomical teaching aids, they sometimes presented with glue or felt on the pubic symphysis. In these cases, residue was removed using acetone prior to observation.

The age and sex of the selected specimens for this study were unknown to all participants. The knowledge of this information was not required because the aim of this study was not to estimate a biological profile of a specimen but to rate the degree to which sex and age estimation from dry bones agrees with estimations done using 3D images.

Scanning Protocol

The selected specimens were scanned at the Boston Medical Center using a 64-slice scanner (GE Healthcare Lightspeed VCT, Waukesha, Wisconsin) in Helical mode, with 0.625 mm per slice thickness at an interval of 0.625 mm, with a radiation dose of 120 kV and 10 mA. The reconstruction parameters used were configured for scanning bone specimens.

Image Rendering

The CT scans were visualized using the 64-bit version of *OsiriX* Imaging Software (Pixmeo, Geneva, Switzerland) on a Macintosh computer. The images were reconstructed from CT slacks using the 3D Volume Rendering (i.e., the resulting images were in 3D as opposed to 2D). A series of rendering settings were tested to determine which were the most beneficial for the analysis of the specimens, with particular emphasis on the detail of the features of the pubic symphysis. Previous researchers (e.g., Wink 2011) suggested the use of the 'Glossy Bone' preset that is integrated within *OsiriX* in aiding the highlighting of features such as the dorsal rampart, but these settings hindered the scoring of the pubic symphysis and they were thus not used. Table 3.1 shows the settings that were ultimately selected for viewing and scoring the specimens.

Table 3.1. 3D Volume rendering settings.	
Window Level (WL)	-520
Window Width (WW)	410
Opacity	Logarithmic Inverse Table
Color Look-Up Table (CLUT)	VR- Bones
3D Presets	Mid Contrast
Detail	Fine
Shading Settings: Default	
Ambient Coefficient	0.15
Diffuse Coefficient	0.90
Specular Coefficient	0.30
Specular Power	15.00

Scoring

The specimens, both dry bones and their corresponding 3D images, were scored within two weeks of each other. A special code was given to the 3D images corresponding to each labeled bone, and a cross-reference log of this information was kept for subsequent statistical analysis and comparison of results.

Scoring for the Phenice (1969) method was done following the guidelines in Buikstra and Ubelaker (1994) (Figure 3.1). The scoring guidelines for the Suchey-Brooks method were followed as defined in Brooks and Suchey (1990) (Table 3.2). The Suchey-Brooks male and female cast series (France Casting, http://www.francecasts.com/) were used as aids in scoring the specimens.

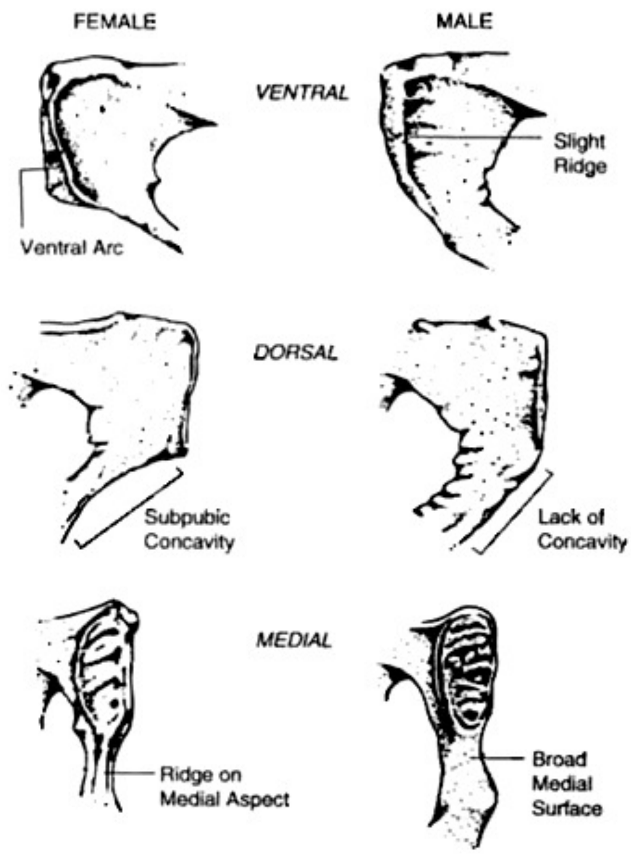

Figure 3.1. The Phenice (1969) method of age estimation relies on the scoring of the ventral arc, subpubic concavity, and medial aspect of the ischiopubic ramus. A score of 1 denotes a female manifestation of the trait, while a score of 3 indicates a male manifestation. A score of 2 was given in cases where neither manifestation was obvious. Adapted from Buikstra and Ubelaker (1994:17).

Table 3.2. Morphological description of Suchey-Brooks phases; as listed by Brooks and Suchey (1990:232). Emphasis retained.	
Phase	**Description**
I	Symphyseal face has a billowing surface (ridges and furrows) which usually extends to include the pubic tubercle. The horizontal ridges are well-marked and ventral beveling may be commencing. Although ossific nodules may occur on the upper extremity, *a key to the recognition of this phase is the lack of delimitation of either extremity*.
II	The symphyseal face may still show ridge development. *The face has commencing delimitation of lower and/or upper extremities occurring with or without ossific nodules*. The ventral rampart may be in beginning phases as an extension of the bony activity at either or both extremities.
III	Symphyseal face shows lower extremity and *ventral rampart in process of completion.* There can be a continuation of fusing ossific nodules forming the upper extremity and along the ventral border. Symphyseal face is smooth or can continue to show distinct ridges. Dorsal plateau is complete. Absence of lipping or symphyseal dorsal margin; no bony ligamentous outgrowths.
IV	Symphyseal face is generally fine grained although remnants of the old ridge and furrow system still remain. Usually the oval outline is complete at this stage, but a hiatus can occur in upper ventral rim. Pubic tubercle is fully separated from the symphyseal face by definition of upper extremity. The symphyseal face may have a distinct rim. Ventrally, bony ligamentous outgrowths may occur on inferior portion of pubic bone adjacent to symphyseal face. If any lipping occurs it will be slight and located on the dorsal border.
V	Symphyseal face is completely rimmed with some slight depression of the face itself, relative to the rim. Moderate lipping is usually found on the dorsal border with more prominent ligamentous outgrowths on the ventral border. There is little or no rim erosion. Breakdown may occur on superior ventral border.
VI	Symphyseal face may show ongoing depression as rim erodes. Ventral ligamentous attachments are marked. In many individuals the pubic tubercle appears as a separate bony knob. The face may be pitted and porous, giving an appearance of disfigurement with the ongoing process of erratic ossification. Crenulations may occur. The shape of the face is often irregular at this stage.

Statistical Analysis

Intraobserver Error

Twenty percent of the sample (n = 26) was scored twice. Intraobserver error was assessed for the scores obtained from each method separately (Phenice in dry bone, Phenice in 3D image, Suchey-Brooks in dry bone, and Suchey-Brooks in 3D image) by calculating Krippendorff's alpha reliability coefficient, using the web-based service ReCal3 (Reliability Calculator, Freelon 2010, found at http://dfreelon.org/utils/recalfront/) for two observations of nominal data for each case.

A reliability coefficient evaluates whether or not a set of instructions given to different observers for the scoring of a set of phenomena (in this case, the methods outlined in Phenice 1969 and Brooks and Suchey 1990) yield the same data within a reasonable margin of error (Hayes and Krippendorff 2007). In other words, it measures the degree to which the observers agree in their observations, as well as providing a basis for assessing whether or not the data can be trusted to represent something real (versus what one would expect to result from variations in human observation). Krippendorff's alpha reliability coefficient has the advantage of being applicable to a set of data regardless of the number of observers involved, the amount of measurement levels, sample sizes, and the presence or absence of missing data (Hayes and Krippendorff 2007). The closer the alpha coefficient is to a value of 1, the more reliable the data.

Degree of Agreement Between Dry Bones and 3D Images

Krippendorff's alpha reliability coefficient was calculated to compare the scores obtained from the dry bones and 3D images. In the case of the Suchey-Brooks method, a percentage was also calculated that reflected the proportion of specimens that were scored exactly the same, as well as those that were scored within one phase of the score for the dry bones (either up or down).

Interobserver Error Study with Observers of Various Levels of Expertise

A pool of 6 individuals with various levels of experience (Table 3.3) affiliated with Boston University's Forensic Anthropology program volunteered for an interobserver error study which consisted of scoring a subset of the specimens (n = 20). As in the main study, both dry bones and 3D images were scored by each individual. A small survey was also provided to test the difficulty in assessing different features of the pubic bone based on the opinion of the observer scoring the specimens.

Table 3.3. Summary of interobserver error participants.

Observer	Level of Experience
A	1st year student, Forensic Anthropology, beginner
B	1st year student, Forensic Anthropology, beginner
C	2nd year student, Forensic Anthropology, intermediate
D	2nd year student, Forensic Anthropology, intermediate
E	Forensic Anthropologist, Expert
F	Forensic Anthropologist, Expert

Due to logistical issues with access to 3D imaging software, and because the observers had no previous experience with the rendering of 3D-CT images, each 3D specimen involved in the interobserver error study was exported as a standalone file. The images were exported into a .MOV file that can be viewed using *QuickTime* Viewer (Apple, Inc.). Exporting from *OsiriX* to a *QuickTime* Virtual Reality (QTVR) would have allowed the observers to tilt the specimen in whichever direction they preferred (very similar to scoring the specimens in *OsiriX* itself, but without the need to alter the image whatsoever). However, the feature is no longer supported by Apple and *OsiriX*. For this reason, a 360º vertical and horizontal view was exported for each bone. This is not a method that has been previously reported in the literature and it was devised to address the specific issue of outside observers having to score 3D images of bones without access to visualization software such as *OsiriX*. Due to concerns that this untested method would limit the ability of the observers to score the specimens, and thus negate the utility of the interobserver error portion of the study, the author scored the 20 specimens using the vertical and horizontal views. No significant obstacles were found in the scoring of these 3D images compared to the virtual environment of *OsiriX* by the author of the present study. A Krippendorff reliability coefficient was calculated for the author's interobserver error scores, and the results show little difference between the images scored using *OsiriX* and the images scored using the 360º vertical and horizontal views.

For the interobserver error study, the Krippendorff coefficient was calculated for each group of observers (globally), for each group of observers divided into level of experience, and for each observer individually. Each participant worked independently. The study was divided into two phases, referred from here on as **Week 1** and **Week 2**. Participants were given a time limit of seven days per phase to report their results, with a waiting period of two weeks between phases. Week 1 consisted on scoring dry bones, while Week 2 consisted of the scoring of the corresponding 3D images.

A short survey was also provided for the participants to complete. The participants were asked to rate the ease with which they scored specific traits (1 = easy, 2 = intermediate, 3 = hard). For the Phenice (1969) method, the traits consisted of the ventral arc, the subpubic concavity, and the ischiopubic ramus. For the Suchey-Brooks method, the traits were the presence or absence of billowing, ridges, and furrows; rampart formation and breakdown; and the presence of micro- and macroporosity.

It is important to note that none of the studies reviewed in Chapter 2 mentioned the exporting of 3D-CT images that would allow observers to score the images without the use of *OsiriX.* The researchers involved in those studies scored the images using the visualization software of their choice, and any intra- and interobserver testing was done the same way. No third parties were involved in the testing of the interobserver error for the vast majority of the studies reviewed in Chapter 2, and thus the logistics of testing interobserver error using third parties (researchers not directly involved in the studies) have not been addressed as of yet. This proved an issue when deciding how to best allow the participants in this interobserver error study to score the 3D images. Considering the lack of experience on the part of the participants in using *OsiriX*, it was decided that the 3D images had to be exported somehow, but no guidelines for this process were found.

Since some investigators have suggested the collaboration of multiple parties working on the same cases or research projects (e.g., Elton and Cardini 2008), establishing a bare minimum of requirements for visualizing 3D images is necessary. A familiarity with *OsiriX* (or any of the visualization software alternatives) and the process of transforming CT scans of bones into 3D images that can be confidently scored is currently a prerequisite in order to apply anthropological methods to 3D images.

There is evidence that standard anthropological methods can be applied to 3D images derived from bones, as seen in recent literature (Decker *et al.* 2011; Ferrant *et al.* 2009; Grabherr *et al.* 2009; Telmon *et al.* 2005; Uldin *et al.* 2012; Villa *et al.* 2013b; Wink 2014). However, research is currently affected by issues with small sample sizes, differences in scanning resolution, and a lack of interobserver error assessment involving multiple observers that takes individual skill and experience into account.

The present study not only assessed the validity of obtaining a biological profile from 3D images derived from bones using a larger sample size compared to that of the current literature, but it also explores interobserver error using participants from the beginner, intermediate and expert levels. The present study also highlights methodological issues that must be resolved before 3D technologies can be widely applied to the field of forensic anthropology. The results of this study are presented in the next chapter.

CHAPTER 4: RESULTS

The results obtained for this study are presented in the following order: first, the degree of agreement found between the scoring of dry bones versus 3D images (n = 130) using each anthropological method is presented, followed by the results from the interobserver error study (divided into global results and into results achieved by each level of observer expertise) and the interobserver survey, followed by the degree of agreement achieved by each individual observer that participated in the interobserver error study. The last two portions of this chapter report the intraobserver agreement achieved by the author of the study (n = 130) as well as the degree of agreement that the author found when scoring the interobserver error materials (for the purpose of comparing results obtained using the vertical and horizontal views from the interobserver error study to the results obtained when scoring 3D images using *OsiriX*).

The results for the main portion of this study are found in Table 4.1.

Table 4.1. Summary of statistics for main study (n = 130).

Method	Krippendorff's Alpha	N Decisions
Phenice (1969)	0.788	260
Brooks and Suchey (1990)	0.382	260

The number of cases that were in agreement when applying the Phenice (1969) method to both types of specimens (physical bones and 3D images) was 122 out of 130

(94%). The number of specimens in agreement for the Suchey-Brooks (1990) method was 66 out of 130 (51%). However, when one considers an agreement as any score that comes within one phase from the score obtained from the dry bones (either upper or lower), the number of scores in agreement rises to 108 out of 130 (83%).

Interobserver Error

Before assessing the degree of agreement between Observers A-F, statistics for each separate phase of the study were calculated (Week 1 and Week 2, Tables 4.2-4.8). This was done to explore the level of agreement observers had for scoring physical bones only and to compare it to the agreement for scoring 3D-CT images only. In this way, the degree of proficiency with each method (Phenice and Suchey-Brooks) of each observer at the time of their taking part in this study is being tested. This was necessary, because a baseline had to be established regarding the degree to which observers agree with one another and with themselves in their scoring abilities. If a marked trend toward global and individual agreement was seen when scoring 3D images, it was important to discern whether this was the result of inherent limitations with 3D imaging methods, or if this diminished agreement was comparable to the results of the same methods being applied to dry bones.

Furthermore, if a marked difference in agreement between the main researcher and the observers that participated in the study was seen, it would be important to discern whether previous or more extensive experience working with 3D-CT images might be a factor to account for. If so, there is a possibility that 3D imaging methods may be applicable to forensic anthropological methods, with the caveat that the observer undergo a degree of training in 3D technologies and that he or she become acquainted with the ways different scoring traits will present themselves in 3D images.

Results: Week 1 (Scoring of Physical Bones)

Tables 4.2 to 4.5 summarize the results obtained for the interobserver error portion of this study for **Week 1**, both globally and subdivided into level of experience of the observers involved.

Table 4.2. Summary of statistics for interobserver error (n = 20).		
Results: Global (Observers A-F)		
Week 1 (Physical Bones)		
Method	**Krippendorff's Alpha**	**N Decisions**
Phenice (1969)	0.243	100
Brooks and Suchey (1990)	0.286	100

Global results for **Week 1** show that the observers were in agreement in 5 out of 20 cases for the Phenice (1969) method (25%) and in only 2 out of 20 cases for the Suchey-Brooks (1990) method (10%). If one considers being within a phase (upper or lower) of a score as agreement, the number rises to 10 out of 20 (50%).

Table 4.3. Summary of statistics for interobserver error (n = 20).		
Results: Beginner (Observers A and B)		
Week 1 (Physical Bones)		
Method	**Krippendorff's Alpha**	**N Decisions**
Phenice (1969)	0.235	40
Brooks and Suchey (1990)	0.336	40

Results for Observers A and B (beginner level) show agreement in 11 out of 20 cases for the Phenice (1969) method (55%) and in 10 out of 20 cases for the Suchey-Brooks (1990) method (50%). If one considers being within a phase (upper or lower) of a score as agreement, the number rises to 17 out of 20 (85%).

Table 4.4. Summary of statistics for interobserver error (n = 20).		
Results: Intermediate (Observers C and D)		
Week 1 (Physical Bones)		
Method	**Krippendorff's Alpha**	**N Decisions**
Phenice (1969)	0.120	40
Brooks and Suchey (1990)	0.514	40

Results for Observers C and D (intermediate level) show agreement in 8 out of 20 cases for the Phenice (1969) method (40%) and in 12 out of 20 cases for the Suchey-Brooks (1990) method (60%). If one considers being within a phase (upper or lower) of a score as agreement, the number rises to 17 out of 20 (85%).

Table 4.5. Summary of statistics for interobserver error (n = 20).		
Results: Expert (Observers E and F)		
Week 1 (Physical Bones)		
Method	**Krippendorff's Alpha**	**N Decisions**
Phenice (1969)	0.391	40
Brooks and Suchey (1990)	0.420	40

Results for Observers E and F (expert level) show agreement in 16 out of 20 cases for the Phenice (1969) method (80%) and in 11 out of 20 cases for the Suchey-Brooks (1990) method (55%). If one considers being within a phase (upper or lower) of a score as agreement, the number rises to 15 out of 20 (75%).

Results: Week 2 (Scoring of 3D-CT Images)

Tables 4.6 to 4.9 summarize the results obtained for the interobserver error portion of this study for **Week 2**, both globally and subdivided into level of experience of the observers involved.

Table 4.6. Summary of statistics for interobserver error (n = 20).		
Results: Global (Observers A-F)		
Week 2 (3D-CT)		
Method	**Krippendorff's Alpha**	**N Decisions**
Phenice (1969)	0.259	80
Brooks and Suchey (1990)	0.123	80

Global results for **Week 2** show that the observers were in agreement in 6 out of 20 cases for the Phenice (1969) method (30%) and in 0 out of 20 cases for the Suchey-Brooks (1990) method (0%). If one considers being within a phase (upper or lower) of a score as agreement, the agreement goes from being null to 7 out of 20 (35%).

Table 4.7. Summary of statistics for interobserver error (n = 20).		
Results: Beginner (Observers A and B)		
Week 2 (3D-CT)		
Method	**Krippendorff's Alpha**	**N Decisions**
Phenice (1969)	0.432	40
Brooks and Suchey (1990)	0.302	40

Results for Observers A and B (beginner level) show agreement in 13 out of 20 cases for the Phenice (1969) method (65%) and in 9 out of 20 cases for the Suchey-Brooks (1990) method (45%). If one considers being within a phase (upper or lower) of a score as agreement, the number rises to 18 out of 20 (90%).

Table 4.8. Summary of statistics for interobserver error (n = 20).		
Results: Intermediate (Observers C and D)		
Week 2 (3D-CT)		
Method	**Krippendorff's Alpha**	**N Decisions**
Phenice (1969)	-0.040	40
Brooks and Suchey (1990)	0.030	40

Results for Observers C and D (intermediate level) show agreement in 5 out of 20 cases for the Phenice (1969) method (25%) and in 6 out of 20 cases for the Suchey-Brooks (1990) method (30%). If one considers being within a phase (upper or lower) of a score as agreement, the number rises to 15 out of 20 (75%).

Table 4.9. Summary of statistics for interobserver error (n = 20).		
Results: Expert (Observers E and F)		
Week 2 (3D-CT)		
Method	**Krippendorff's Alpha**	**N Decisions**
Phenice (1969)	0.568	40
Brooks and Suchey (1990)	0.345	40

Results for Observers E and F (expert level) show agreement in 18 out of 20 cases for the Phenice (1969) method (90%) and in 11 out of 20 cases for the Suchey-Brooks (1990) method (55%). If one considers being within a phase (upper or lower) of a score as agreement, the number rises to 17 out of 20 (85%).

Table 4.10 presents a summary of the statistical results obtained for **Week 1** and **Week 2**, both globally and subdivided into experience level.

Table 4.10. Summary of statistics for interobserver error (n = 20).		
Results: Global (Observers A-F)		
Method	**Krippendorff's Alpha (Physical Bones)**	**Krippendorff's Alpha (3D-CT)**
Phenice (1969)	0.243	0.259
Brooks and Suchey (1990)	0.286	0.123
Results: Beginner (Observers A and B)		
Phenice (1969)	0.235	0.432
Brooks and Suchey (1990)	0.336	0.302
Results: Intermediate (Observers C and D)		
Phenice (1969)	0.120	-0.04
Brooks and Suchey (1990)	0.514	0.03
Results: Expert (Observers E and F)		
Phenice (1969)	0.391	0.568
Brooks and Suchey (1990)	0.420	0.345

Global results for Week 1 and Week 2 using the Phenice (1969) method were very similar, with a Krippendorff's alpha coefficient of 0.243 for dry bones versus 0.286 for 3D images (Tables 4.2 and 4.6). Even though the alpha coefficient is not very high (closer to 0 than to 1), the fact that agreement between Week 1 and Week 2 results was not very different suggests that sex estimation from 3D images using the Phenice traits is viable. Global results for each week using the Brooks and Suchey (1990) method, however, showed a lower level of agreement when applied to 3D images, with a Krippendorff's alpha coefficient of 0.269 when scoring dry bones but a coefficient of 0.127 when scoring 3D images. These results are consistent with those reported by Grabherr *et al.* (2009) and Villa *et al.* (2013b) where there was a decrease in reliability of the Suchey-Brooks method when applied to 3D images. It should be noted, however, that Krippendorff's alpha reliability coefficient measures only perfect agreement, so that it does not consider two observers that scored a specimen within one phase of the Suchey-Brooks age estimation method as an agreement.

Furthermore, there were marked differences when accounting for observer experience level. Observers at the beginner level had a higher degree of agreement compared to the global results. Observers at the intermediate level, however, had a markedly lower degree of agreement compared to the global figures. Observers at the expert level had a higher degree of agreement than the rest of the observers. Results for Observers A and B had a higher degree of agreement using 3D-CT images than they did using standard methods. On the other hand, Observers C and D had a lower agreement for the Phenice (1969) method but had a higher rate of agreement with the Suchey-Brooks (1990) method when scoring dry bones. At the expert level, the level of agreement was lower for sex estimation using dry bones compared to 3D images, but the agreement for age estimation using dry bones was higher than for 3D images.

When one takes the calculated percentages into account (such that being within one phase of the Suchey-Brooks scoring system is considered as an agreement), observers at the beginner level were more in agreement with each other when scoring 3D images, observers at the intermediate level had a lower level of agreement when scoring 3D images, and observers at the expert level saw an increase in their agreement when scoring 3D images. If experience level is what determines the degree to which two observers will agree on their assessments (due to having a similar level of knowledge and practice with applying age and sex estimation methods), then it would be expected to see a clear pattern of increasing agreement from beginners to intermediates and all the way to experts. However, this was not the case for the present study. It is possible that a repetition of this study using a larger sample size would yield very different results.

Week 1 and Week 2 Agreement

The degree of agreement between the results obtained for dry bones and for 3D images by each individual observer (A-F) are shown in Table 4.11. The first percentage for the Suchey-Brooks method consists of perfect matches. The second percentage consists of scores within one phase of agreement (upper or lower).

Table 4.11. Summary of statistics for each observer.

Method	Krippendorff's Alpha	Percentage of Agreement
Observer A		
Phenice (1969)	0.112	45% (9 out of 20)
Brooks and Suchey (1990)	0.316	50% (10 out of 20)
		70% (14 out of 20)
Observer B		
Phenice (1969)	0.460	65% (13 out of 20)
Brooks and Suchey (1990)	0.217	40% (8 out of 20)
		85% (17 out of 20)
Observer C		
Phenice (1969)	0.826	90% (18 out of 20)
Brooks and Suchey (1990)	0.032	30% (6 out of 20)
		90% (18 out of 20)
Observer D		
Phenice (1969)	0.546	70% (14 out of 20)
Brooks and Suchey (1990)	0.023	25% (5 out of 20)
		60% (12 out of 20)
Observer E		
Phenice (1969)	0.568	85% (17 out of 20)
Brooks and Suchey (1990)	0.494	55% (11 out of 20)
		80% (16 out of 20)
Observer F		
Phenice (1969)	0.391	80% (16 out of 20)
Brooks and Suchey (1990)	0.165	45% (9 out of 20)
		65% (13 out of 20)

By far, the observer that agreed the most with themselves was Observer E (expert level). However, when considering percentages alone, Observer C (intermediate) agreed with 90% of their sex estimation assessment, and 90% for age estimation in dry versus 3D bones, although their Krippendorff's alpha coefficient was remarkably low (0.032) because it only takes perfect agreement into account.

All observer except Observer A (beginner) had a higher rate of agreement with their own assessments when estimating sex than when estimating age.

Interobserver Scoring Survey

Figures 4.1 to 4.6 reflect the results for the interobserver scoring survey. Overall, the observers reported that the features they had most difficulty scoring were the ischiopubic ramus and rampart formation and breakdown in dry bones (on average, rated as 2 = intermediate). For 3D-CT images, micro- and macroporosity, the ischiopubic ramus and rampart formation and breakdown were considered the most difficult to observe (on average, rated as 3 = hard). One observer (Observer A, beginner level of experience) found it "basically impossible" to score 3D images using the Suchey-Brooks (1990) method, despite results showing a relatively high degree of agreement between Observer A's scores using this method for dry bones and 3D images, and even higher than the results the observer obtained when applying the Phenice (1969) method. In other words, even though Observer A's results were reasonable in their agreement, the observer's confidence in scoring 3D images was very low.

Overall, there was a tendency towards participants considering 3D-CT images harder to score, except for Observer C (intermediate level of experience), who found it easier to score the ventral arc and rampart formation and breakdown in 3D images.

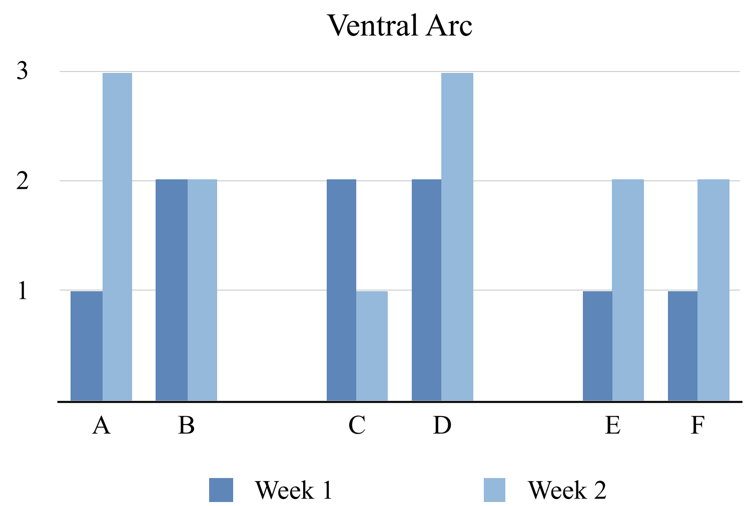

Figure 4.1. Combined results of Week 1 and Week 2, scoring of the ventral arc. Phenice (1969) method.

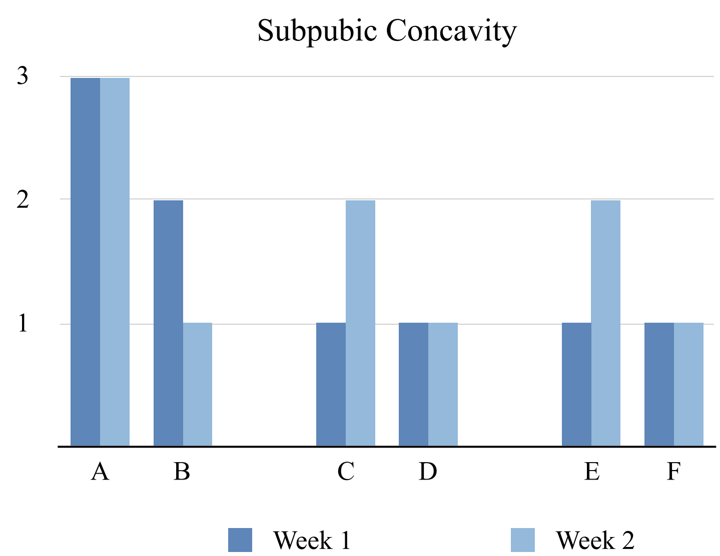

Figure 4.2. Combined results of Week 1 and Week 2, scoring of the subpubic concavity. Phenice (1969) method.

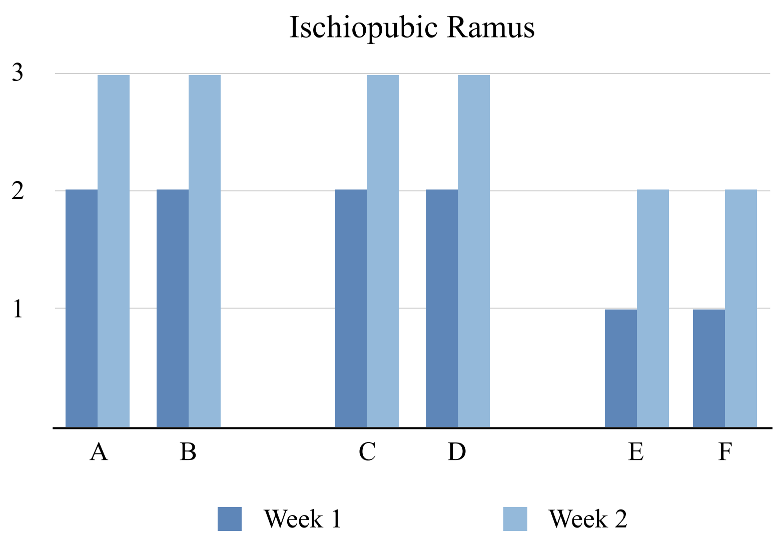

Figure 4.3. Combined results of Week 1 and Week 2, scoring of the ischiopubic ramus. Phenice (1969) method.

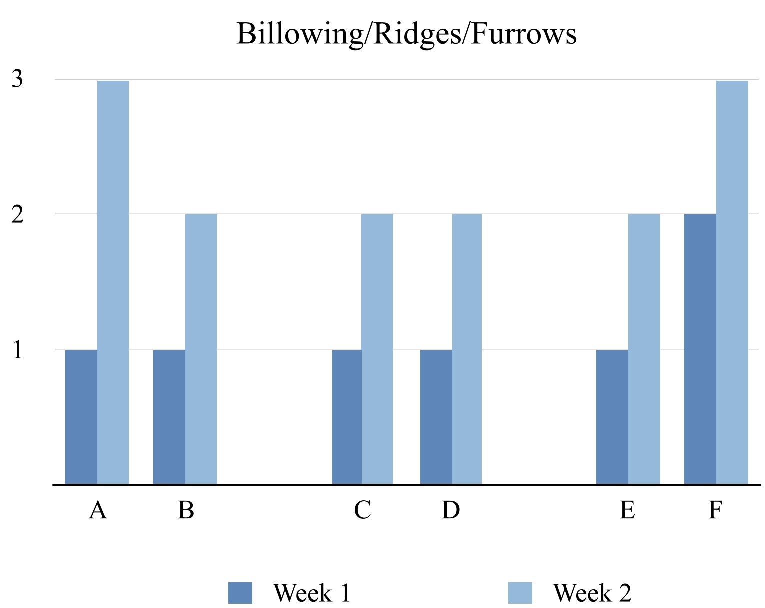

Figure 4.4. Combined results of Week 1 and Week 2, scoring of the presence or absence of billowing, ridges and furrows. Suchey-Brooks method.

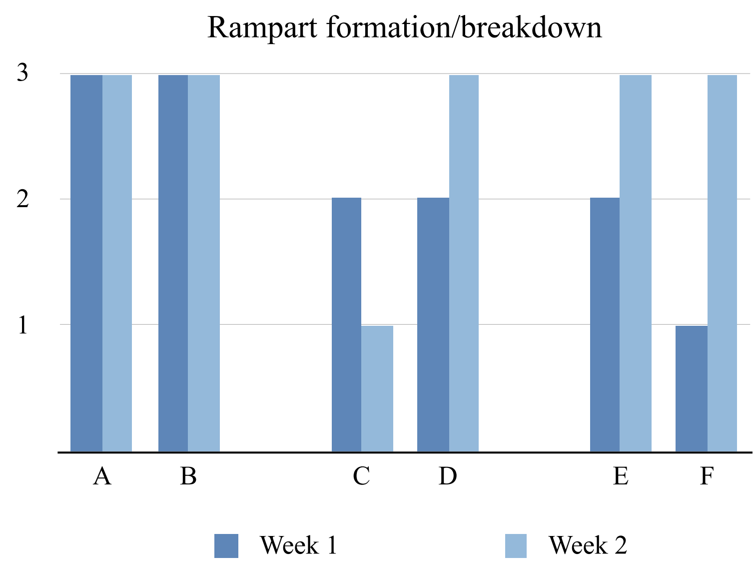

Figure 4.5. Combined results of Week 1 and Week 2, scoring of rampart formation and breakdown. Suchey-Brooks method.

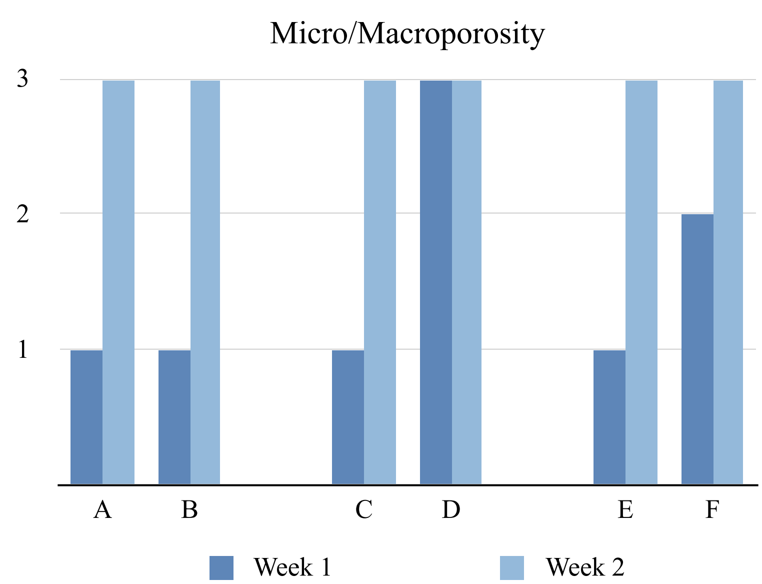

Figure 4.6. Combined results of Week 1 and Week 2, scoring of the observability of micro- and macroporosity. Suchey-Brooks method.

Intraobserver Agreement Results

Intraobserver agreement for the author was scored twice for each method: once for the scores obtained on dry bone and once for scores obtained from 3D-CT images (Table 4.12). The agreement between dry bone and 3D-CT scores was also calculated for the second time the specimens were scored.

Table 4.12. Intraobserver error for the author.		
Method	**Krippendorff's Alpha**	**Percentage of Agreement**
Dry Bones		
Phenice (1969)	1	100% (26 out of 26)
Brooks and Suchey (1990)	0.472	58% (15 out of 26)
		88% (23 out of 26)
3D-CT		
Phenice (1969)	0.710	92% (24 out of 26)
Brooks and Suchey (1990)	0.610	69% (18 out of 26)
		92% (24 out of 26)
Dry Bones versus 3D-CT		
Phenice (1969)	0.514	88% (23 out of 26)
Brooks and Suchey (1990)	0.508	62% (16 out of 26)
		88% (23 out of 26)

Scoring Method Validity Assessment: DICOM versus .MOV

Table 4.13 shows the results obtained for the interobserver error portion of the study as scored by the author for the purpose of comparing the use of *OsiriX* for 3D images (DICOM format) versus an exported, standalone *QuickTime* Viewer file (.MOV format).

Table 4.13. Summary of statistics for interobserver error scored by the author (n = 20) using .MOV files.

Agreement Between Physical Specimens and 3D-CT Images		
Method	**Krippendorff's Alpha**	**N Decisions**
Phenice (1969)	0.777	40
Brooks and Suchey (1990)	0.488	40

Results for the author show agreement in 19 out of 20 cases for the Phenice (1969) method (95%), and in 11 out of 20 cases for the Suchey-Brooks method (55%). If one considers being within a phase (upper or lower) of a score as agreement, the number rises to 18 out of 20 (90%). This shows good agreement between the scores obtained from the physical bones and the scores obtained using the vertical and horizontal 3D views of the bones.

Table 4.14 compares the the degree of agreement between the results obtained when scoring 3D images using *OsiriX* versus the vertical and horizontal images viewed using *QuickTime* Viewer.

Table 4.14. Comparison of 3D images scored using *OsiriX* (DICOM) versus *QuickTime* Viewer (.MOV).		
Method	**Krippendorff's Alpha**	**N Decisions**
Phenice (1969)	1	40
Brooks and Suchey (1990)	0.612	40

Scoring the 3D images using both methods outlined above showed an agreement in 20 out of 20 cases for the Phenice (1969) method (100%), and in 14 out of 20 cases for the Suchey-Brooks method (70%). If one considers being within a phase (upper or lower) of a score as agreement, the number rises to 18 out of 20 (90%).

These results showed that there was a very good degree of agreement between the scores obtained when visualizing images using *OsiriX* and when scoring the vertical and horizontal views. This suggests that exporting 3D images into vertical and horizontal 360º views is an acceptable alternative to scoring 3D images viewed with *OsiriX*.

CHAPTER 5: DISCUSSION

The results obtained by the author were encouraging, although a higher Krippendorff's alpha reliability coefficient was found when comparing dry bone and 3D bone scores for the Phenice (1969) method than for the Brooks and Suchey (1990) method (0.788 versus 0.382, respectively). As it has been mentioned before, the reliability coefficient does not consider a score within +/-1 of the score it is being compared to as an agreement. For this reason, the raw percentages of agreement were also provided. The percentage of agreement achieved for the Phenice (1969) method was 94% (122 out of 130), while the percentage for the Brooks and Suchey (1990) method was 83% (108 out of 130). The decrease in agreement for the Suchey-Brooks method is consistent with what has been previously noted in the literature, particularly in the research published by Grabherr *et al.* (2009) and Villa *et al.* (2013b). The present results were also comparable to those of Wink (2014), who found that the true age of her subjects was captured 79.5% of the time when scoring 3D images of their pubic symphyses. Wink (2014) obtained a Krippendorff's alpha reliability coefficient of 0.59 for her intraobserver error, while the author of the present study had a coefficient of 0.61 and an agreement between scores 92% of the time when applying Brooks and Suchey (1990) to 3D images.

It should be noted that the agreement between specimens being scored twice for the intraobserver test using 3D-CT images in the present study was higher than when scoring dry bones twice using the same method (see Table 4.12). When scoring dry bones using

Phenice (1969), an agreement of 100% and a reliability coefficient of 1 was found. The agreement went down to 92% with a reliability coefficient of 0.71 when testing intraobserver error for 3D-CT images. In other words, the author of the present study had a higher agreement for the Suchey-Brooks method when scoring 3D images, but a lower agreement for the Phenice (1969) method compared to the agreement between the scores for dry bones.

In several instances, the author of the present study scored specimens as a Phase IV (Brooks and Suchey 1990) using 3D-CT images but scored the corresponding dry bones as a Phase VI. It was found that being able to tell the point at which the pubic symphyseal face broke down was difficult using 3D images. Furthermore, although dorsal rampart lipping was identified as a sharp-looking margin in 3D images, the bony spurs that would show on the ventral aspect in Phases V or VI were obscured, potentially due to the tendency for 3D visualization software to smooth out the surface of the specimen being visualized, despite attempts to minimize this effect by tweaking rendering settings in *OsiriX* (particularly, by choosing the finest level of detail under the 3D Volume Rendering option of visualization). While the scoring of specimens following Phenice (1969) using 3D images was generally deemed more simple by all observers involved in this study (the author included), it was found that on occasion, it was difficult to score the presence or absence of a ventral arc, as it was not immediately evident in the 3D visualizations whether it was a true ventral arc or the result of very prominent bony spicules along the ventral aspect of the pubic bone.

Interobserver Error Results

The results for the interobserver error portion of this study were not as expected. It was hypothesized that score agreement would vary given the level of expertise for each observer. Such a pattern would have signaled that the level of observer expertise was the main factor that affected the degree to which they agreed with each other. It was expected that this pattern would be even more evident when scores from dry versus 3D specimens for each individual observer were compared (see Table 4.11). However, no such pattern was found. It was noted that observers at the beginner level had a higher agreement with each other compared to the global results (i.e., the results comparing the agreement of all observers that participated, regardless of level of expertise). Observers at the intermediate level tended to greatly disagree with each other, to the point where a negative Krippendorff's alpha reliability coefficient was calculated for their scoring of the Phenice (1969) traits from 3D images (see Table 4.10). Their results when scoring dry bones using the same method were not very encouraging, either. Observers at the expert level had a higher degree of agreement with each other compared to global results.

The most interesting results obtained from the interobserver error study can be seen when one explores the degree to which each observer agreed with their own assessment of the bones. By far, the observer that achieved the highest level of agreement was Observer C (intermediate level), with an agreement of 90% when following Phenice (1969) and 90% when following Brooks and Suchey (1990). No other participant in the interobserver error study approached this level of agreement. Surprisingly, Observer B (beginner level) had similar degrees of agreement compared to Observer F (expert level) (see Table 4.11). When the interobserver error results are compared to the results obtained by the author of this study for the purpose of testing the vertical and horizontal 360º method of visualizing the 3D specimens (see Table 4.13), it can be seen that the author had a higher degree of agreement in her assessments compared to every observer that participated in the study (94% for the Phenice traits, 90% for the Suchey-Brooks method). The Krippendorff's alpha reliability coefficients were also much higher for the author of the study than for the other observers, despite the previously noted downside of the coefficient when it is calculated for the Suchey-Brooks method (i.e., it only counts perfect agreement between scores). This results were obtained despite the fact that the author belongs to the intermediate level of experience.

There is evidence to suggest that experience level is not the main factor that affected the degree of agreement found when scoring dry bones and 3D bones in this study. It is suggested, then, that a level of familiarity with the way dry bones look when they are digitized and visualized in a computer, as well as the way in which the traits being scored present themselves in 3D images (e.g., dorsal rampart lipping will look like a sharper ridge compared to a pubic symphysis with no dorsal lipping) is a prerequisite before an observer can apply methods to 3D bones that were initially meant for dry bones. It is suggested, for the Brooks and Suchey (1990) method in particular, that digitizing of pubic symphyses that are comparable to those illustrated in the Suchey-Brooks casts would be helpful to observers attempting to estimate age from 3D images in that it would help illustrate the way in which particular Suchey-Brooks phases will look in 3D form.

Methodological Issues

Exporting 3D images to be scored by observers who have no knowledge of how to use 3D visualization software posed a dilemma. *OsiriX* presents observers with the ability to rotate specimens in 360º views using all orientations as well as being able to zoom in to better visualize areas of interest. However, and despite much searching, no suitable ways of visualizing standalone 3D specimens in a similar way was found. As it was mentioned before, *OsiriX* used to support the exporting of 3D models to a QuickTime Virtual Reality file, which allowed observers to handle the images the same as if they were visualized in *OsiriX* itself. However, Apple, Inc. has ceased its support of QTVR technology and such files can no longer be exported as standalone 3D images using *OsiriX.*

Tests conducted on the vertical and horizontal views scored by the interobserver error participants suggest that these views are an acceptable means for scoring 3D bone specimens. That being said, it would be ideal to find a way that helps observers visualize the 3D bones the same way a QTVR file would.

Practical Considerations

Some practical considerations were identified during the course of this study, which are outlined below. Although not directly related to the current study, they are nevertheless issues that will arise as the applications of virtual anthropology become more widespread.

Scanning Protocol

After searching the literature for previous research that focused on scoring 3D images of bones in anthropology (Decker *et al.* 2011; Ferrant *et al.* 2009; Grabherr *et al.* 2009; Telmon *et al.* 2005; Uldin *et al.* 2012; Villa *et al.* 2013b; Wink 2014), a need for a more standardized scanning protocol was noted, which is consistent with the findings of Uldin *et al.* (2012). This includes not only the type of scanning used for different types of specimens (e.g., MRI versus CT) but also particular settings such as slice thickness, radiation dosage and reconstruction logarithm (kernel) to be used.

As it was mentioned before, a kernel for hard tissues is preferable for best possible image quality since it will yield the least amount of surface smoothing when using 3D visualization software. Robinson *et al.* (2008) recommended a slice thickness of no more than 1.25 mm, but a slice thickness of 0.625 mm is optimal for a CT scanner. The type of specimen to be scanned also should be considered when choosing a type of scanner.

Villa *et al.* (2013b) achieved a high level of detail for the digitization of dry bones using a contact-free laser scanner. The specimens scored in the present study were scanned using CT technology because the main research objective of the author is to be able to extract complete skeletons from CT scans of living individuals. However, for the purpose of digitizing osteological collections and the like, a contact-free scanner is most probably the best alternative to achieve the highest level of detail possible.

CT scanners and MRI machines can be used for both living populations and cadavers. Although MRI machines are theoretically preferable due to their lack of adverse side effects as compared to CT scanners, a comparison of the quality of images obtained between the two types of technology for the specific purpose of visualizing bone tissue is needed.

μCT scanners are, unfortunately, prohibitive in terms of costs (both time and money) and in terms of the size of the specimen that can be fit into the physical scanner window. There is a chance that increased demand in the medical sector will result in scanners that are easier and cheaper to use, but the issue of whether or not the maximum specimen size allowed will increase is yet to be seen. Because advances in scanning technology are mostly dependant on the needs of medical professionals and their patients, machines that incur lower costs in terms of both time and money are likely to be developed; however, the quality of the images these machines produce is not likely to improve (Weber and Bookstein 2011:59).

3D Visualization Software

A review of 3D visualization software for anthropological purposes could not be found in the literature currently available for virtual anthropology. Unfortunately, since this type of software is usually developed with a specific operating system in mind, it will be difficult to find a single program that works on all or most platforms (Windows, OS X, Linux, etc.). This obstacle could theoretically be overcome if a way of visualizing standalone 3D images is found.

Databases

Researchers in the field of virtual anthropology have stressed the possibility of sharing 3D files across the web so that multiple researchers in different parts of the globe are able to work on a single case by remote access. There is also the possibility of digitizing entire osteological collections to increase the ease with which researchers can collect data for their projects as well as decreasing travel costs and the deterioration of valuable osseous material due to constant handling.

However, the size of the files generated by scanners can vary depending on the technology being used and it could potentially be very costly to maintain large files in digital databases. There is also the question of who would own the digital osteological collections. A good way of raising funds for the maintenance of digital servers would be to charge researchers an access fee. Before large amounts of data can be stored digitally, it is necessary to determine the answers to these questions.

CHAPTER 6: CONCLUSIONS

The benefits of applying 3D technologies to standard forensic anthropological approaches hold the promise of helping the discipline overcome some methodological issues that have plagued it for decades. Namely, 3D technologies, if applied properly, would enable the forensic anthropologist to gather population-specific data that could be translated into methods devised for particular regions of the globe, wherever they are needed, with the knowledge that the results obtained will fit the population the methods are being applied to and will help identify missing individuals with a greater accuracy.

The present study contributes to the growing knowledge of virtual anthropology by performing a study using a larger sample size (n = 130) and by including a test of interobserver error using a subsample of 20 specimens. This study adds to the evidence already extant in the literature which suggests that applying standard anthropological methods commonly used to construct a biological profile to 3D images of a skeleton can yield similar results comparable to those obtained from dry bones (Decker *et al.* 2011; Ferrant *et al.* 2009; Grabherr *et al.* 2009; Telmon *et al.* 2005; Uldin *et al.* 2012; Villa *et al.* 2013b; Wink 2014). However, more practical issues such as the kind of software needed to visualize 3D images, the specific settings needed for the scanning protocol to yield suitable images and the type of scanning technology needed for specific types of specimens to be digitized have been largely left unaddressed in the literature. A move towards standardization of 3D scanning and visualizing methodologies is suggested. The interobserver error results also suggested that a familiarity with applying methods commonly used in anthropology to 3D-CT images of bones could be a factor that affects the scores obtained from them. Since the results obtained in the interobserver error portion of the present study could be due to a low sample size (n = 20), it is suggested that this new hypothesis be tested using a larger sample size as well as a larger pool of observers with varying years of experience. In any case, researchers with no background in radiology or knowledge of the way 3D images of bones are visualized are encouraged to familiarize themselves with the way bones are rendered in 3D and what features commonly scored in traditional anthropology such as the pubic symphysis and the Phenice traits will look like when digitized this way, before attempting to score them using methods devised for dry bones.

Future research on the subject includes the digitizing of pubic symphyses that are comparable to those illustrated in the Suchey-Brooks casts to aid researchers in applying the method to their specimens. An alternate version of the present study that explores interobserver error using a documented sample is also suggested for the purpose of exploring the degree to which scores are consistent with the age and sex of the individuals in question.

REFERENCES

Abel, R.L., Laurini, C.R., Richter, M. (2012). A palaeobiologist's guide to 'virtual' micro-CT preparation. *Paleontologia Electronica* 15(2):6T,17p.

Allard, T.T. (2006). The Role of 3D Printing in Biological Anthropology. M.A. thesis, University of Manitoba, Winnipeg. Accessed: http://hdl.handle.net/1993/279 .

Aso, J., Martínez-Quiñones, J.V., Aso-Vizán, J., Pons, J., Arregui, R., Baena, S. (2005). Virtopsia. Aplicaciones de un nuevo método de inspección corporal no invasiva en ciencias forenses. *Cuadernos de Medicina Forense* 11(40):95-106.

Ball, R., Shu, C., Xi, P., Rioux, M., Luximon, Y., Molenbroek, J. (2010). A comparison between Chinese and Caucasian head shapes. *Applied Ergonomics* 41:832-839.

Baum, T., Carballido-Gamio, J., Huber, M.B., Müller, D., Monetti, R., Räth, C., Eckstein, F., Lochmüller, E.M., Majumdar, S., Rummeny, E.J., Link, T.M., Bauer, J.S. (2010). Automated 3D trabecular bone structure analysis of the proximal femur—prediction of biomechanical strength by CT and DXA. *Osteoporosis International* 21:1553-1564.

Benazzi, S., Fantini, M., De Crescenzio, F., Mallegni, G., Mallegni, F., Persiani, F., Gruppioni, G. (2009). The face of the poet Dante Alighieri reconstructed by virtual modelling and forensic anthropology techniques. *Journal of Archaeological Science* 36:278-283.

Benazzi, S., Bertelli, P., Lippi, B., Bedini, E., Caudana, R., Gruppioni, G., Mallegni, F

(2010). Virtual anthropology and forensic arts: the facial reconstruction of Ferrante Gonzaga. *Journal of Archaeological Science* 37:1572-1578.

Bilfeld, M.F., Dedouit F., Rousseau, H., Sans, N., Braga, J., Rougé, D., Telmon, N. (2012). Human coxal bone sexual dimorphism and multislice computed tomography: geometric morphometric analysis of 65 adults. *Journal of Forensic Science* 57(3):578-588.

Boas, F. (1912). Changes in the bodily form of descendants of immigrants. *American Anthropologist* 14(3):530-562.

Boyd, S.K. (2009). Micro-computed tomography. In C.W. Sensen and B. Hallgrímsson (eds) *Advanced Imaging in Biology and Medicine: Technology, Software Environments, Applications*, Springer-Verlag Berlin Heidelberg.

Brenner, D.J. (2010). Should we be concerned about the rapid increase in CT usage? *Reviews on Environmental Health*, *25*(1), 63-68.

Brooks, S., Suchey, J.M. (1990) Skeletal age determination based on the os pubis: A comparison of the Acsádi-Nemeskéri and Suchey-Brooks methods. *Human Evolution* 5:227-238.

Buikstra, J. E. and D. H. Ubelaker (eds.) (1994) *Standards for Data Collection from Human Skeletal Remains*. Arkansas Archeological Survey Research Series No. 44, Fayetteville, AR.

Decker, S.J., Davy-Jow, S.L., Ford, J.M., Hilbelink, D.R. (2011). Virtual determination of sex: metric and nonmetric traits of the adult pelvis from 3D computed tomography models. *Journal of Forensic Sciences* 56(5):1107-1114.

Dedouit, F., Telmon, N., Costagliola, R., Otal, P., Joffre, F., Rougé, D. (2007). Virtual anthropology and forensic identification: Report of one case. *Forensic Science International* 173:182-187.

DeSilva, J. M., Holt, K. G., Churchill, S. E., Carlson, K. J., Walker, C. S., Zipfel, B., Berger, L. R. (2013). The lower limb and mechanics of walking in Australopithecus sediba. *Science* 340(6129).

De Angelis, D., Sala, R., Cantatore, A., Grandi, M., Cattaneo, C. (2009). A new computer-assisted technique to aid personal identification. *International Journal of Legal Medicine* 123:351-356.

De Greef, S., Willems, G. (2005). Three-dimensional cranio-facial reconstruction in forensic identification: latest progress and new tendencies in the 21st century. *Journal of Forensic Sciences* 50(1):1-6.

Elton, S., Cardini, A. (2008). Anthropology from the desk? The challenges of the emerging era of data sharing. *Journal of Anthropological Sciences* 86:209-212.

Enciso, R., Shaw, A., Neumann, U., Mah, J. (2003). 3D head anthropometric analysis. *Medical Imaging* 5029:590-597.

Enciso, R., Alexandroni, E.S., Benyamein, K., Keim, R.A., Mah, J. (2004). Precision, repeatability and validation of indirect 3D anthropometric measurements with light-based imaging techniques. Biomedical Imaging: Nano to Macro. *IEEE International Symposium* Vol. 2:1119-1122.

Ferrant, O., Rougé-Maillart, C., Guittet, L., Papin, F., Clin, B., Fau, G., Telmon, N. (2009). Age at death estimation of adult males using coxal bone and CT scan: A preliminary study. *Forensic Science International* 186:14-21.

Freelon, D.G. (2010). ReCal: Intercoder reliability calculation as a web service. *International Journal of Internet Science* 5:20-33.

Friess, M. (2012). Scratching the surface? The use of surface scanning in physical and paleoanthropology. *Journal of Anthropological Sciences* 90:1-25.

Grabherr, S., Cooper, C., Ulrich-Bochsler, S., Uldin, T., Ross, S., Oesterhelweg, L., Bolliger, S., Christe, A., Schnyder, P., Mangin, P., Thali, M.J. (2009). Estimation of sex and age of “virtual skeletons”— A feasibility study. *European Radiology* 10:419-429.

Guyomarc’h, P., Santos, F., Dutailly, B., Desbarats, P., Bou, C., Coqueugniot, H. (2012). Three-dimensional computer-assisted craniometrics: A comparison of the uncertainty in measurement induced by surface reconstruction performed by two computer programs. *Forensic Science International* 219:221-227.

Hallgrímsson, B., Zelditch, M.L., Parsons, T.E., Kristensen, E., Young, N.M., Boyd, S.K. (2011). Morphometrics and biological anthropology in the postgenomic age. In

M.A. Katzengerg and S.R. Saunders (eds) *Biological Anthropology of the Human Skeleton* (2nd ed.), pp. 207-235, Wiley & Sons.

Hammond, P., Hutton, T.J., Allanson, J.E., Campbell, L.E., Hennekam, R.C.M., Holden, S., Patton, M.A., Shaw, A., Temple, I.K., Trotter, M., Murphy, K.C., Winter, R.M. (2004). 3D Analysis of Facial Morphology. *American Journal of Medical Genetics* 126A:339-348.

Hayes, A.F., Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures* 1(1):77-89.

Hoppa, R.D. (2000). Population variation in osteological aging criteria: An example from the pubic symphysis. *American Journal of Physical Anthropology* 111:185-191.

Jabir, S. (2012). Simulation in surgery: A proposal for quality control in the use of simulators in surgical training. *The Internet Journal of Medical Simulation* 3(2): 1-15.

Jantz, R. L., Meadows Jantz, L. (2000). Secular change in craniofacial morphology. *American Journal of Human Biology* 12(3):327-338.

Jones, P.R.M., Rioux, M. (1997). Three-dimensional Surface Anthropometry: Applications to the Human Body. *Optics and Lasers in Engineering* 28:89-117.

Kettner, M., Schmidt, P., Potente, S., Ramsthaler, F., Schrodt, M. (2011). Reverse engineering— Rapid prototyping of the skull in forensic trauma analysis. *Journal of Forensic Science* 56(4):1015-1017.

Kullmer, O. (2008). Benefits and risks in virtual anthropology. *Journal of Anthropological Sciences* 86:205-207.

Kuzminksy, S.C., Gardiner, M.S. (2012). Three-dimensional laser scanning: potential uses for museum conservation and scientific research. *Journal of Archaeological Science* 39:2744-2751.

Lottering, N., MacGregor D.M., Barry, M.D., Reynolds, M.S., Gregory, L.S. (2014). Introducing standardized protocols for anthropological measurement of virtual subadult crania using computed tomography. *Journal of Forensic Radiology and Imaging* 2:34-38.

Lovell, N.C. (1989). Test of Phenice's technique for determining sex from the os pubis. *American Journal of Physical Anthropology* 79:117-120.

Lu, J.M., Wang, M.J.J. (2008). Automated anthropometric data collection using 3D whole body scanners. *Expert Systems with Applications* 35:407-414.

Malina, R.M., Peña Reyes, M.E., Kheng Tan, S., Buschang, P.H., Little, B.B., Koziel, S. (2004). Secular change in height, sitting height and leg length in rural Oaxaca, southern Mexico: 1968-2000. *Annals of Human Biology* 31(6):615-633.

Martin, C.M. (2012). Exploration of three-dimensional morphometrics of the hip joint and reconstructive technologies. *University of Western Ontario - Electronic Thesis and Dissertation Repository.* Dissertation, University of Western Ontario, program in Kinesiology.

Meadows Jantz, L., Jantz, R. L. (1999). Secular change in long bone length and proportion in the United States, 1800–1970. *American Journal of Physical Anthropology* 110(1):57-67.

Nagamine, T., Uemura, T., Masuda, I. (1992). 3D facial image analysis for human identification. *Pattern Recognition* Vol. I. Conference A: Computer Vision and Applications, Proceedings., 11th IAPR International Conference on. IEEE, 1992.

Jantz, R., S. Ousley (2005) *FORDISC 3.0*. University of Tennessee Press, Knoxville, TN.

Ozsoy, U., Demirel, B.M., Yildirim, F.B., Tosun, O., Sarikcioglu, L. (2009). Method selection in craniofacial measurements: Advantages and disadvantages of 3D digitization method. *Journal of Cranio-Maxillofacial Surgery* 37:285-290.

Phenice, T. W. (1969) A newly developed visual method of sexing the os pubis. *American Journal of Physical Anthropology* 30:297-301.

Robinson, C., Eisma, R., Morgan, B., Jeffery, A., Graham, E.A.M., Black, S., Rutty, G.N. (2008). Anthropological measurement of lower limb and foot bones using multi-detector computed tomography. *Journal of Forensic Sciences* 53(6):1289-1295.

Salih, Q.A., Ramli, A.R., Mahmud, R., Wirza, R. (2004). 3D visualization for blood cells analysis versus edge detection. *The Internet Journal of Medical Technology* 1(2).

Sansoni, G., Docchio, F., Trebeschi, M., Scalvenzi, M., Cavagnini, G. (2007). Application of three-dimensional optical acquisition to the documentation and the analysis of crime scenes and legal medicine inspection. *Advances in Sensors and Interface, 2007. IWASI 2007. 2nd International Workshop on*. IEEE, 2007.

Sansoni, G., Cattaneo, C., Trebeschi, M., Gibelli, D., Porta, D., Picozzi, M. (2009). Feasibility of Contactless 3D Optical Measurement for the Analysis of Bone and Soft Tissue Lesions: New Technologies and Perspectives in Forensic Sciences. *Journal of Forensic Sciences* 54(3):540-545.

Sansoni, G., Cattaneo, C., Trebeschi, M., Gibelli, D., Poppa, P., Porta, D., Maldarella, M., Picozzi, M. (2011). Scene-of-crime analysis by a 3-dimensional optical digitizer: a useful perspective for forensic science. *American Journal of Forensic Medicine and Pathology* 32(3):280-286.

Shetti, R.C. (2006). Augmented reality training kit can offer novel teaching solution for future cardiac surgery and other medical professionals. *The Internet Journal of Medical Simulation* 2(1).

Scheuer, L., Black, S. (2007). Osteology. In T. Thompson and S. Black (eds) *Forensic Human Identification: An Introduction.* Boca Raton, FL: CRC Press, pp. 199-219.

Stock, S. R. (2008). *Microcomputed tomography: methodology and applications*. Boca Raton, FL: CRC press, 348 p.

Telmon, N., Gaston, A., Chemla, P., Blanc, A., Joffre, F., Rougé, D. (2005). Application of the Suchey-Brooks method to three-dimensional imaging of the pubic symphysis. *Journal of Forensic Sciences* 50(3):1-5.

Thali, M.J., Braun, M., Brueschweiler, W., Dirnhofer, R. (2003a). Morphological imprint: Determination of the injury-causing weapon from the wound morphology using forensic 3D/CAD-supported photogrammetry. *Forensic Science International* 132:177-181.

Thali, M.J., Braun, M., Dirnhofer, R. (2003b). Optical 3D surface digitizing in forensic medicine: 3D documentation of skin and bone injuries. *Forensic Science International* 137:203-208.

Thali, M.J., Taubenreuther, U., Karolczak, M., Braun, M., Brueschweiler, W., Kalender, W.A., Dirnhofer, R. (2003c). Forensic microradiology: micro-computed tomography (micro CT) and analysis of patterned injuries inside bone. *Journal of Forensic Sciences* Vol. 48(6).

Tocheri, M. W., Razdan, A., Dupras, T. L., Bae, M., Liu, D. (2002). Three dimensional quantitative analyses of human pubic symphyseal morphology: Can current limitations of skeletal aging methods be resolved. *American Journal of Physical Anthropology* (Supplement 34) 155.

Ubelaker, D.H., Volk, C.G. (2002). A test of the Phenice method for the estimation of sex. *Journal of Forensic Sciences* 47(1):19-24.

Uldin, T., De Froidmont, S., Vaucher, P., Mangin, P., Grabherr, S. (2012). Virtual anthropology: a comparative study of real bone vs. virtual bone surface.

Proceedings: American Academy of Forensic Sciences. Annual Scientific Meeting, Atlanta, GA., February 20-25 2012, pp. 421-422.

University of Granada (2013). New imaging technique for identifying the age and sex of a corpse. *ScienceDaily*. Published January 23, 2013. Retrieved March 23, 2014 from www.sciencedaily.com/releases/2013/01/130123094303.htm

Villa, C., Normand Hansen, M., Buckberry, J., Cattaneo, C., Lynnerup, N. (2013a). Forensic age estimation based on the trabecular bone changes of the pelvic bone using post-mortem CT. *Forensic Science International* 233:393-402.

Villa, C., Buckberry, J., Cattaneo, C., Lynnerup, N. (2013b). Reliability of Suchey-Brooks and Buckberry-Chamberlain methods on 3D visualizations from CT and laser scans. *American Journal of Physical Anthropology* 151:158-163.

Weber, G.W., Bookstein, F.L. (2011). *Virtual Anthropology: A Guide to a New Interdisciplinary Field.* New York: Springer.

Willis, A., Anderson, D., Thomas, T., Brown, T., Lawrence Marsh, J. (2007). 3D Reconstruction of Highly Fragmented Bone Fractures. *Proceedings of SPIE* Vol. 6512:65121P1-65121P10.

Wink, A. E. (2011). *Application of Anthropological Aging Methods to Three-dimensional Reconstructions of Clinical CT Scans of the Adult Pelvis*. Master's thesis, Boston University, School of Medicine.

Wink, A. E. (2014). Pubic symphyseal age estimation from three−dimensional reconstructions of pelvic CT scans of live individuals. *Journal of Forensic Sciences* 59(3):696-702.

Wong, J.Y., Oh, A.K., Ohta, E., Hunt, A.T., Rogers, G.F., Mulliken, J.B., Deutsch, C.K. (2008). Validity and reliability of craniofacial anthropometric measurement of 3D digital photogrammetric images. *The Cleft Palate-Craniofacial Journal* 45(3):232-239.

Wozniak, K., Rzepecka-Wozniak, E., Moskala, A., Pohl, J., Latacz, K., Dybala, B. (2012). Weapon identification using antemortem computed tomography with virtual 3D and rapid prototype modeling: A report in a case of blunt force head injury. *Forensic Science International* 222:e29-e32.

Zech, W.D., Hatch, G., Siegenthaler, L., Thali, M.J., Lösch, S. (2012). Sex determination from os sacrum by postmortem CT. *Forensic Science International* 221:39-43.

Zhou, B., Willis, A., Sui, Y., Anderson, D., Thomas, T., Brown, T. (2009). Improving inter- fragmentary alignment for virtual 3D reconstruction of highly fragmented bone fractures. *SPIE medical imaging* pp. 725934-725934. International Society for Optics and Photonics.

Zollikofer, C. P. E., Ponce De León, M. S., Martin, R. D. (1998). Computer−assisted paleoanthropology. *Evolutionary Anthropology: Issues, News, and Reviews* 6(2): 41-54.

CURRICULUM VITAE

Name: Maria Julia García de León Valenzuela

Year of birth: 1986

E-mail address: julesgdl@bu.edu

Profile

Reading University archaeology graduate with a focus on human osteology, trauma and palaeopathology. Research interests include crime scene investigation, the application of 3D technologies to forensic anthropology, bone trauma, the study of cancer and other pathological conditions in past societies, food archaeology and anthropology, burial archaeology and the attitude towards death and dying. Has experience speaking to a crowd in an academic setting, excels at researching and summarising large quantities of information and is very proficient with technology and software (three semesters of IT Engineering on a university level completed before starting a career in archaeology).

Education

- **2007 to 2010** - BSc Honours in Archaeology - First Class Degree, University of Reading, U.K.
- **2012 -** MSc in Forensic Anthropology - Boston University, U.S., class of 2014.

Field Schools, Achievements & Presentations

- **2008** - Silchester Field School, Silchester, U.K. (Romano-British archaeology).
- **2010 -** Award, Best Undergraduate Dissertation for the Archaeology Department of the University of Reading. Dissertation title: A Study of Trauma in the Anglo-Saxon Agricultural Community of Bishopstone.

- **2011 -** Conference, “Set in Stone: The Evolution of Sculpture in Ancient Rome from the Republic to Early Christianity”, Technological Institute for Higher Studies of Monterrey in Hermosillo, Sonora.
- **2011 -** Short lecture, “Antropologia Forense: Cuerpos que Hablan”.
- **2012 -** Scholarship, Becas CONACyT al Extranjero.
- **2013 -** Short lecture and casework-based training in the study of skeletal trauma at the Human Identification Center in Medellín, Colombia.
    - **Instructor**: Elizabeth DiGangi, Ph.D. (edigangi@yahoo.com)
- **2014 -** Poster presentation, 66th Annual Scientific Meeting of the American Academy of Forensic Science, Seattle, WA (Feb 2014).
    - Toon, C., García de León, J. (2014). A Comparison of the Klales *et al*. (2012) and Phenice (1969) Methods of Sex Estimation on a Modern Colombian Sample. *Proceedings of the 66th Annual Scientific Meeting* Vol. 20:463.

Specific Skills & Training

- Trained in the mapping of sites with a Leica total station combined with TDS Recon data logging equipment, using TDS Survey Pro software.
- OsiriX Medical Imaging Software

Languages

Spanish (fluent), English (fluent), Japanese (beginner), French (beginner).