Testing the Reliability and Reproducibility of the V.E.R.A Method: Application on

Entheses of the Proximal and Distal Humerus

A Thesis submitted to the faculty of
San Francisco State University
In partial fulfillment of
the requirements for
the Degree

Master of Arts

In

Anthropology

by

Jonathan Ieng

San Francisco, California

December   2023

Certification of Approval

I certify that I have read *Testing the Reliability and Reproducibility of the V.E.R.A Method: Application on Entheses of the Proximal and Distal Humerus* by Jonathan Puthdarin Ieng, and that in my opinion this work meets the criteria for approving a thesis submitted in partial fulfillment of the requirement for the degree Master of Arts in Anthropology at San Francisco State University.

_____

Cynthia Wilczak, Ph.D.
Professor of Anthropology,
Thesis Committee Chair

_____

Mark Griffin, Ph.D.
Associate Professor of Anthropology

Abstract

The purpose of this thesis is to test the reliability and reproducibility of the V.E.R.A. method. Additionally, it looks to examine whether the V.E.R.A. method's application goes beyond that of hand entheses. The V.E.R.A. method was applied to six entheses of the humerus. This included the subscapularis, supraspinatus, infraspinatus, teres minor, common extensor origin, and common flexor origin. 23 humeri of unknown origin and background were examined and measured using a three-dimensional scanner. Inter and intra-observer errors were calculated based off measurements taken and correlation tests were used to examine relationships among entheses of the humerus. Results showed that intra-observer error rates were between 1.83% and 3.23%. Intraclass correlation coefficient was used to evaluate inter-observer error and reported excellent values between 0.87 and 0.99, As expected, strong correlations were found among all entheses. The sample utilized had no demographic information which limited the scope of this study. Further research may look to apply the V.E.R.A method to humeri with known demographic   information.

Acknowledgements

I would like to thank my advisor, Dr. Wilczak, for assisting me throughout this whole process and making this thesis possible. Thank you for putting in so much work into getting the scanner and helping me push my thesis along. I want to also thank Dr. Griffin and Dr. Reifschneider for the multiple courses I took with each of them throughout the years. The three of them inspired me to always think critically and consider multiple perspectives. I applied this both in my thesis and with life in general. Additionally, thank you Dr. Karakostis for taking the time to help train me with using the V.E.R.A. method as well as looking over my data. I would also like to thank my friends for supporting me through this journey and lending me their ears when I needed someone to talk to. To my family, I know this took a lot longer than expected but I did it. Thank you for all the positive encouragement and for keeping me accountable. To my older brother Kelvin, I wish you could have been here to see me complete my thesis, and I know you would have been proud.

Table of Contents

List of Tables

List of Figures

Introduction

In bioarchaeology the study of entheseal changes has been used to reconstruct past activity patterns. These studies rely on the notion that there is a relationship between entheses and physical activity. Entheseal changes are thought to be a result of constant or heavy strain due to physical activities, therefore bioarchaeologists can hypothetically use them to infer a past individual's activity patterns (Jurmain et al. 2012; Villotte and Knusel 2013). Entheses are also easily accessible on dry bones and as a result they are more appealing to study in bioarchaeology (Villotte and Knusel 2013).

Contemporary quantitative methods have utilized the advancement in scanning technology to analyze entheseal changes through the use of three-dimensional scanners. The advantage to using three-dimensional scanners is that it allows the researcher to directly measure the complex structure of an enthesis in ways that were not accessible before. This includes analyzing its size, structure, shape, and surface topography in more detail. This in turn has produced data that is more flexible. That is to say that researchers can approach the data using different methods of statistical analyses. Additionally, the adoption of scanning technologies has broadened the scope of research possible for entheseal changes, allowing researchers to explore different approaches. This could include incorporating different techniques that other disciplines often use in their own research. For example, applying fractal analysis or micro-tomodensitometry (Zumwalt 2005, 2006; Berthon et al. 2015).

The most recent method developed for analyzing entheseal changes was created by Karakostis and Lorenzo (2016). This method involves using three-dimensional scanning in

conjunction with multivariate statistical analysis. In contrast with previous methods, Karakostis and Lorenzo's method examines multiple entheses and multiple features of entheses concurrently with one another. The purpose this serves is that the analysis produces results that show an amalgamation of similarities and differences across multiple entheses. Similar to al-Oumaoui and colleages (2004) and Milella and co-authors (2015), the statistical approach used in this method does not form groups prior to analysis. As a result, groups can more easily be discerned based on the similarities and differences of the entheses without *a priori* assumptions about entheseal correlation patterns. Karakostis and Lorenzo's method illustrated that it is possible to assess the influence of physical activity on entheseal shape and size, while at the same time addressing the multifactorial etiology of entheses. This method has proven to be successful in multiple studies, which have shown evidence of a correlation between physical activity and entheseal change (e.g. Karakostis et al 2018a, 2018b; Karakostis, Jeffrey, and Harvati 2019; Karakostis et al. 2019). However, it has only been applied to hand entheses. With this in mind, the aim of my research will be to apply the Karakostis and colleagues (2017) method to entheses of the rotator cuff along with the origin for the common extensor and common flexor. In doing so, my research will examine whether correlations exist between entheseal sites on the rotator cuff and if differences exist in the patterns of correlations. At the same time, this research will test the efficiency and reliability of this method and investigate whether its scope can go beyond hand entheses.

Chapter 1: Literature Review

Background

Studies on entheseal changes in bioarchaeology mainly look to reconstruct activity patterns of populations and individuals (Hawkey and Merbs 1995; Jurmain et al. 2012; Lai and Lovell 1992). The underlying assumption in studies of entheseal changes is that biomechanical stress caused by habitual activity is directly related to an enthesis formation and remodeling. Therefore, entheseal changes are a direct indicator of physical activity. Some studies have had success with linking entheseal changes to specific activities (e.g. Eshed et al. 2004; Godde and Taylor 2011; Hawkey and Merbs 1995; Lai and Lovell 1992; Molnar 2006, 2010; Niinimaki and Sotos 2013; Steen and Lane 1998), while other studies have not been able to link entheseal changes directly to specific activity patterns (e.g. al-Oumaoui, , Jiménez-Brobeil, and du Souich 2004; Wilczak 1998; Zumwalt 2005, 2006). The varying degrees of success has led researchers to critique the methodology involved in studying entheseal changes. The methodology going into the analysis and interpretation of entheseal changes has proven to be an issue in the field of bioarchaeology. Researchers have indicated that entheses have a multifactorial etiology and cannot be used as a direct link to physical activity (Davis et al. 2013; Jurmain et al. 2012; Weiss 2010; Zumwalt 2006; Wilczak 1998). There are factors such as age, sex, body size, genetics, and pathology that can affect the expression of an enthesis. Furthermore, entheses are complex structures that come in different types classified by the tissues present and the structure at the bone-tendon interface (Apostolakos et al. 2014; Benjamin et al. 2002, 2006; Benjamin and Ralphs 1998; Henderson et al. 2017; Villotte et al. 2016). Critiques have specifically addressed

these issues noting that some methodologies ignore the multifactorial etiology or do not consider the anatomy of entheses (Henderson and Cardoso 2013; Henderson et al. 2013; Henderson et al. 2017; Henderson and Nikita 2016; Jurmain et al 2012; Villotte et al. 2016; Wilczak 1998).

The purpose of this review is to examine the methodology used to analyze entheseal changes across the skeleton. By discussing the different methodologies used, one will be able to see the limitations that each of these methodologies have. Additionally, it will review the issues with the analyses of entheseal changes and how these issues hinder the research. Furthermore, reviewing the methodologies will reveal how they have shifted and modified, leading up to the emergence of a new multivariate methodology. The introduction of this new methodology has provided researchers with some of the best evidence for the effects of activity pattern on entheseal changes (e.g. Karakostis et al. 2017, 2018a, 2018b, 2019; Karakostis, Jeffrey, and Harvati 2019).

This review will cover the history of methodologies involving the study and analysis of entheseal changes. It will begin with an introduction to terminology and the background involving entheses. Methodology will be discussed in terms of qualitative and quantitative methods. Additionally, the limitations of these methodologies will be outlined as well as how contemporary methods look to address the limitations.

Entheseal changes: terminology and background

Entheses can be defined as "the region where a tendon, a joint capsule, or a ligament attaches to bone" (Benjamin et al. 2002: 931) and osteological changes to these areas are known as "entheseal changes" following the recommendation of Jurmain and Villotte (2010). The first use of enthesis dates back to 1959 and was first used by G. La Cava "for creating the word

"enthesitis" to designate inflammation of tendon attachments into bone" (Jurmain and Villotte

2010:1). There are and have been many different terms that were used before the adoption of

"entheseal change". These include, but are not limited to: enthesopathies, enthesophytes, muscle

crests, musculoskeletal stress markers, activity-induced pathology, evidence for occupation,

skeletal stress markers of occupational stress, activity-induced stress markers, or muscle

markings (Jurmain and Villotte 2010; Villotte and Knusel 2013). The term "entheseal change"

was adopted because it is neutral, as it avoids indicating the etiology of change. This change is

especially pertinent to studies that examine activity and occupation as using any of the

previously mentioned terms indicates a main etiology. Whereas "entheseal change" reflects a

multifactorial etiology (Jurmain and Villotte 2010). Another advantage is that these terms can be

used for both pathological and non-pathological changes (Jurmain and Villotte 2010).

There are two main types of enthesis: fibrous and fibrocartilaginous. The distinction of

these types of entheses is important as there are clear differences in their morphology,

development, the way they react to mechanical loads, and where they are located on the skeleton

(Apostolakos et al. 2014; Benjamin et al.2002; Benjamin and Ralphs 1998; Jurmain and Villotte

2010; Schelcht 2012). Fibrocartilaginous entheses are more ubiquitous on the skeleton as they

attach at the epiphysis of long bones and are also in the foot and hand (Apostolakos et al. 2014;

Benjamin and Ralphs 1998). They can also show up on some areas of vertebrae (Jurmain and

Villotte 2010). Fibrocartilaginous entheses have four zones of tissue. These include pure dense

fibrous connective tissue, uncalcified fibrocartilage, calcified fibrocartilage, and bone (Benjamin

et al. 1986; Benjamin et al. 2002). This transition from tendon to cartilage allows for more

mobility of the bone and reduces the mechanical stresses placed on the bone. However, this also

contributes to fibrocartilaginous entheses being more prone to injuries (Apostolakos et al. 2014; Benjamin et al. 2002; Benjamin and Ralphs 1998).

Fibrous entheses attach to the diaphysis in limb bones but can also be present in the skull and vertebrae (Benjamin and Ralphs 1998; Jurmain and Villotte 2010). They are split into two sub-categories based on how the tendon attaches. The tendon can either insert directly into the bone, known as "bony" fibrous entheses, or insert into the periosteum, known as "periosteal" fibrous entheses (Apostolakos et al 2014; Jurmain and Villotte 2010). Fibrous entheses are also associated with large muscles in the body, like the deltoid and muscles of the femoral shaft (Benjamin et al. 2002). However, fibrous entheses are not explored as much in the literature in comparison to fibrocartilaginous entheses, due to the latter being more prone to overuse and thus easier to discern biomechanical stress (Apostolakos et al. 2014). Additionally, there is no explicit definition for what a "normal" fibrous enthesis should be. That is to say that there is no "normal" expression of bony surface for fibrous entheses, whereas a normal fibrocartilaginous enthesis "is smooth, well, circumscribed and devoid of vascular foramina." (Villotte et al. 2016:52). Therefore, deviations away from normal fibrocartilaginous entheses are easier to define.

The application of entheseal changes

Entheseal changes are used in a variety of ways. Predominantly, entheseal changes have been used in research to reconstruct past activities (Jurmain et al. 2012; Villotte and Knusel 2013). However, their application has seen use beyond inferring about occupation and activity. They can also be used to infer the possible disabilities that individuals faced while living (Lai and Lovell 1992; Papathanasiou, Zachou, and Richards 2009). In addition, entheseal changes can help elucidate intra-population variation. Differences in sex, body shape, and size are among

some of the variations that can be explored (Godde and Taylor 2011; Hawkey and Merbs 1995; Molnar 2010; Myszka, Piontek, and Vancata 2012; Steen and Lane 1998; Weiss, Corona, and Schultz 2012). Depending on the area of the skeleton that is examined, analyses on entheseal changes will paint a different picture of life history. For example, studies conducted on the upper limb could illustrate activity patterns that involved a heavy use and reliance on the muscles of the upper limb. Eshed et al. (2004) study on entheses of the upper limb in Natufian hunter-gatherers and Neolithic farmers found that the entheses were a result of a transition towards an agricultural lifestyle. This transition resulted in higher stress loads on the upper limb in the Neolithic population compared to the Natufian population (Eshed et al. 2004). Studies on the lower limb may look to investigate the effects that different terrain exert on entheses of the lower limb. Acosta and Cunha (2017) found differences in frequency of entheseal changes when comparing individuals who resided on rugged terrain vs flat terrain. They found that those who inhabited flat terrain environments exhibited a higher frequency of entheseal changes. This is most likely due to exposing these entheses to high levels of biomechanical stress which they are otherwise not adapted to. This contradictory result illustrates the complexities of analyzing entheses and entheseal changes. Furthermore, the methodology used to analyze entheseal changes also plays a role in the depiction of life history.

Methodology plays an important role in understanding entheses because it directly affects the interpretation and outcome of a study. A method involves having accuracy, precision, and appropriate statistical testing. It also involves collecting the appropriate type of data (i.e. qualitative vs quantitative) which can influence the statistical method used. The consequence of using unsuitable data or statistical analysis can result in misleading, misinformed, or

oversimplified results. Therefore, having an appropriate method for studying entheseal changes

is crucial when trying to derive meaningful information regarding past behavior. There have

been a wide variety of methodologies developed for the analysis of entheses and entheseal

changes, and all of them aim to infer past behavior (al-Oumaoui, Jiménez-Brobeil, and Souich

2004; Berthon et al. 2015; Hawkey and Merbs 1995; Karakostis et al. 2017; Mariotti, Facchini,

and Belcastro 2007; Molnar 2010; Noldner and Edgar 2013; Nolte and Wilczak 2013; Villotte et

al. 2009; Santos et al. 2011; Wilczak 1998; Zumwalt 2005) This range of methodology stems

from the fact that there is no agreed upon way to analyze entheseal changes. It also stems from

the recent critiques on the validity of major methods used (Santos et al. 2011). These critiques

focus on the intra- and interobserver errors that may occur during analysis, as well as

confounding factors such as age and body size that are not considered in the analysis (Davis et al.

2013; Wilczak et al. 2017). Moreover, some methods try to account for other methods'

limitations, but in doing so they in turn have their own limitations. As a result, trying to create a

standardized methodology has proven to be difficult.

Qualitative methodologies

Early methods relied on qualitative scoring to analyze entheseal changes. These methods

rely on macroscopically analyzing the morphological changes occurring at ligament and tendon

attachment sites. This involves visual and meticulous inspection by the researcher who may (or

may not) rely on photographs or a magnification tool to discern characteristics of entheseal

changes. These changes are then scored into discrete categories.

Kennedy (1983) provides an early example of discussing methodology to analyze

entheses (what he referred to as markers of occupational stress) to infer past activity behavior.

His discussion involved the macroscopic examination of the proximal end of the ulna. Kennedy notes that hypertrophy of the supinator crests paired with deep supinator fossae can be an indicator of an individual's handedness. Kennedy also notes that these characteristics can aid in the identification of individuals who engaged in specific activities using one arm during life.

Angel and colleagues (1987), along with Kelley and Angel (1987), expand on Kennedy's (1983) use of muscle crests and markers of stress. Their study looked to reconstruct the lifestyle of the people in an early 19th century free Black community in Philadelphia. An aspect of lifestyle they sought to reconstruct was occupation, and they did it through the interpretation of what they referred to as "muscle crests". The method developed by Angel and colleagues was based on an ordinal system. This rating system scored "muscle crests" as being absent or present and used pluses (+) to rank severity (Angel et al. 1987). They also took note of the shape of the "muscle crests", ridging, and furrowing but did not score these characteristics. As the authors note, accuracy relies mainly on the experience of the observer. The method was applied to the deltoid and pectoral site of the humerus, the supinator site of the ulna, and the adductor site of the femur. They compared the results of individuals from the First African Baptist Church (FABC) with the "muscle crests" of slaves from Catoctin, Maryland. Both papers were able to link activities in life to the muscle crests on the individuals.

The two aforementioned methods reflect an early understanding of entheses and entheseal change. The authors recognize the characteristics of changes that can occur such as a build up of bone or furrowing. However, they did not have a system in which they could evaluate all these characteristics. Regardless, these early methods highlighted the characteristics that later methods adopted in their scoring.

Hawkey and Merbs

Hawkey and Merbs' (1995) method is widely used in the study of entheseal changes. The method was introduced in their study on two ancient Thule Eskimo groups from northwest Hudson Bay in Canada (Hawkey and Merbs 1995). This study looked to clarify habitual activity patterns among Thule adult males and females. The method was first applied to bones of the upper body (Hawkey and Merbs 1995). Hawkey and Merbs (1995) method consisted of a visual scoring system that rated entheses on three main categories of expression with each having their own numerical scale. The three main categories of expression consisted of robusticity: defined as rugged ridges or crests, including any type of new bone formation and prominence of cortical bone; stress lesions; defined as microtrauma in the form of putting or furrows, and ossification: defined as abrupt macrotrauma in the form of exostosis (Hawkey and Merbs 1995). Their scale consisted of four specific grades that ranged from 0-3. A score of 0 meant absence of expression, 1 meant faint or slight expression, 2 meant a moderate expression, and a 3 meant strong expression. These scores are taken separately and then converted into a single category that compromises of a scale from 0-6. This is done for the purpose of statistical analysis (Hawkey and Merbs 1995). To further illustrate their method, the authors scored both fibrous and fibrocartilaginous entheses in the same system. Hawkey and Merbs also viewed robusticity and stress lesions as accumulations of microtrauma where the scores represented a continuum from robusticity to stress lesions. Ossification was understood as being caused by abrupt microtrauma and therefore was not included in the same continuum scoring system as robusticity and stress lesions. In comparison to the methods before, Hawkey and Merbs (1995) were the first to introduce a numerical scoring system for the purpose of statistical analyses. They also expanded

upon the characteristics of entheses put forth by Kennedy (1989), Angel and colleagues (1987),

and Kelley and Angel (1987) by providing specific definitions.

This method introduced what they believed to be, a "simple, consistent way to

score…while using categories that were broad enough to account for individual variation…"

(Hawkey and Merbs 1995: 327). In the decades that have passed, Hawkey and Merbs (1995) has

proven to be a useful method as many researchers have applied the methods to other parts of the

skeleton with success (e.g. Eshed et al. 2004; Godde and Taylor 2011; Niinimaki and Sotos

2013; Molnar 2010; Steen and Lane 1998; Weiss 2004; Weiss 2010; Weiss, Corona, and Schultz

2012). Some studies have applied the method to analyze entheses on the skull and mandible

(Steen and Lane 1998; Weiss 2010). Others have applied this method to study entheses on the

innominate (Steen and Lane 1998; Niinimaki and Sotos 2013). Researchers have also applied the

method to entheses on the lower limb (i.e. femur, tibia, and fibula) (Steen and Lane 1998; Weiss

2004; Molnar 2010; Niinimaki and Sotos 2013).

Other qualitative methods

Though Hawkey and Merbs (1995) method is widely used, researchers have noted that

the categories used do not accurately depict the range of variation in entheseal expression

(Molnar 2006, 2010; Mariotti, Facchini, and Belcastro 2004, 2007; Weiss 2010). To address this

issue some researchers have modified the Hawkey and Merbs (1995) method to include different

or intermediate categories.

Hawkey and Merbs (1995) method made a distinction between robusticity markers, stress

lesions, and ossification by scoring them separately in the initial observation. Molnar (2006,

2010) uses a modification of this method, developed by Nagy (2000), that treats robusticity

markers and stress lesions as one category during the initial scoring. The author notes that "the different appearances represent a continuum rather than different types of expressions." (Molnar 2006:15). Her method consists of scoring skeletal markers from 0 to 3.5 with 0.5 intervals and 3.5 representing the most distinct (Molnar 2006). A score of 0 represented that there was "no expression or there was faint roundedness" (Molnar 2006:15). Entheses with a score of 1 showed "notable unevenness, pitting, or furrowing" (Molnar 2006:15). A 2 meant that there was a distinct skeletal change "with furrowing and pitting with a rugged appearance" (Molnar 2006:15). A score of 3 demonstrated that there was an "extensive and clearly defined irregular surface, crest, pitting, ridge, or a combination of both" (Molnar 2006:15). Molnar's method incorporates a different understanding of how entheseal changes manifest in comparison with Hawkey and Merbs (1995) method. They used the similar descriptions to explain stress lesions but applied it differently. This demonstrates how entheses and the development of entheseal changes was not yet fully understood.

Mariotti and colleagues (2004, 2007) provide a modified version of Hawkey and Merbs (1995) method. They recognize two aspects of entheses that can be considered separate. One is robusticity, what refers to "the normal osseous marking at site of attachment of a muscle or ligament" (Mariotti, Facchini, and Belcastro 2004:148) and can have variable degree of expression. The other is enthesopathies, which refers to the nonspecific pathological status of the enthesis and can be absent or present with different degrees of development (Mariotti, Facchini, and Belcastro 2004). They split enthesopathies into two types: osteophytic (OF) and osteolytic (OL). OF is a proliferative osteophytic form that is characterized by the appearance of enthesophytes. OL is an erosive osteolytic form that is "characterized by eroded or pitting areas"

(Mariotti, Facchini, and Belcastro 2004:148). In comparison to Hawkey and Merbs (1995) method, Mariotti and colleagues separated Hawkey and Merbs robusticity into robusticity (what they considered normal osseous markings) and osteophytic markings. They also used osteolytic lesions as an equivalent to Hawkey and Merbs stress lesions.

Mariotti and colleagues (2007) used 3 differing levels for the degree of robusticity. These levels included: degree 1 reflected a weak to moderate expression, degree 2 represented strong development of the enthesis, and degree 3 was used for extreme or strong expression of robusticity without the presence of an enthesopathy (Mariotti, Facchini, and Belcastro 2007). Additionally, "the first degree is further broken down into three subcategories. Degree 1a, a slight impression or an extremely low development. Degree 1b, low development. Degree 1c, medium development" (Mariotti, Facchini, and Belcastro 2007:300). The inclusion of subcategories allows more flexibility in the scoring process that can accommodate to the experience of the scorer. They applied this method to 23 postcranial skeleton entheses and used different criteria for the stages of development on each of the individual entheses, reflecting differences in the underlying cortical bone form at each enthesis. The authors proposed a different method for measuring enthesopathies. They created a scale that could be applicable to any enthesis and used a scale from 0-3 that reflected the development of the enthesopathy. Mariotti and colleagues' method distinguishes between the variation of cortical bone architecture at entheses and enthesopathies. They also use a distinct scoring system that examines entheseal surface changes separately rather than treating it as a continuum between bone formation and bone lysis. This contrasts with how Molnar's (2006, 2010) method treated robusticity and stress lesions as a continuum. Mariotti and colleagues (2004, 2007) method is also one of the few

methods to directly address enthesopathies and incorporate a standardized method to evaluate them.

Havelková and Villotte (2007) provide a detailed discussion of Villotte's proposed visual scoring method from 2006. What differentiates Villotte's method from its predecessors is that it utilizes medical insights on the histological and anatomical structure of entheseal insertion sites provide by Benjamin and colleagues (1986; 2002). As a result, Villotte's method recognizes that different areas of the enthesis undergo different types of remodeling. The Villotte method consists of two parts, a scoring system and assessment of taphonomy (Havelková and Villotte 2007). The scoring system consists of four independent scoring systems based on the type of remodeling that occurs in four muscle groups. Taphonomy is evaluated through the assessment of enthesopathies that are given a grade of A, B, or C (Havelková and Villotte 2007). This method considers the difference in how each type of enthesis responds to mechanical and biological factors. This is further extended to the different zones created for the purpose of analyzing fibrocartilaginous entheses. Each zone acts differently, where one zone is more akin to respond like a fibrous enthesis. The scoring system for each zone also reflects how biomechanical stress manifests itself in these areas. The outer zone scores are based on the degree of irregularity or "osseous" formation and the inner zone scores represent different types of changes that include erosions, foramina, and calcifications or bony deposits (Havelková and Villotte 2007).

In 2010, Villotte and colleagues revised the method combining stage B and C to create a simplified system of presence and absence of enthesopathy. When comparing Mariotti and Villotte's methods to Hawkey and Merbs, they provided a more descriptive visual scoring

method. Villotte's method only considers surface changes caused by new bone formation instead of including overall prominence like the previous mentioned methods do. Additionally, Villotte does not distinguish between osteolytic and osteophyic changes like Mariotti. Their scoring is based on the histological structure of entheses and how different entheses attach. Villotte also considers the differences within entheses as well. Villotte's method does address enthesopathies in a similar way to Mariotti's. That is to say that both aimed to create a standardized way of evaluating them.

Aspects of Mariotti and Villotte's methods were combined in the conception of the Coimbra method. The Coimbra method was a result of a 2009 workshop held in Coimbra, Portgual that addressed issues of studying entheseal changes (Henderson et al. 2013). The Coimbra method is specific to fibrocartilaginous entheses. Following the Villotte method, it divides the enthesis into zone 1 the margin, and zone 2 the surface. The original method records five different features– bone formation, erosion, fine porosity, macroporosity, and cavitation— for two or three degrees of expression (Henderson et al. 2013). A revision to the method, the "New" Coimbra method, added textural change as a sixth feature and reduced the degrees of expression to two for all features except textural change with one degree of expression (Henderson et al; Henderson, Wilczak, and Mariotti 2017). Zone 1 is only scored for bone formation and erosion, while zone 2 is scored for all features. Each feature is scored with varying degrees of expression and interpreted separately. Moreover, this method accounts for anatomical knowledge of entheseal changes as well as variability in the types of changes. Out of the methods talked about so far, the Coimbra method has proven to have higher reproducibility, but it can still be improved (Henderson et al. 2013). The advantage of the Coimbra method is that it

scores features separately. This allows features to be analyzed separately as some features may be more sensitive to biomechanical factors (Michopoulou, Nikita, and Henderson 2017). The Coimbra method is the most detailed out of the qualitative methods and this level of detail allows a more in-depth analysis of entheseal changes on fibrocartilaginous entheses (Palmer et al 2018).

Most qualitative analyses use univariate statistics that compare single entheses at a time. However, al-Oumaoui and co-authors (2004) along with Milella and colleagues (2015) took a different approach utilizing multivariate methods. al-Oumaoui and co-authors (2004) created a simplified method that only examines the presence or absence of entheses. The authors applied their method to five populations from the Iberian Peninsula to inquire about physical activity and differences between sexes within each culture. They noted that the Hawkey and Merbs (1995) method was "developed in non-Iberian populations that were genetically different from those under study here." (al-Oumaoui, Jiménez-Brobeil, and du Souich 2004:346) The populations under study were all morphologically gracile so the presence/absence scale was more appropriate and confounding factors like body size did not have to be accounted for. al-Oumaoui and colleagues (2004) relied on the frequencies of absence and presence of entheses in conjunction with historical data to make interpretations. They performed cluster analyses on the data in order to identify similar groups within the population. The advantage of cluster analyses is that groups are formed during the process rather than being known *a priori* (al-Oumaoui, Jiménez-Brobeil, and du Souich 2004). This type of analysis deviates from the typical univariate analyses used in qualitative methods and aligns more with the multivariate analysis that will be discussed later on. The authors acknowledge the limitations of their method. Specifically, that it limits what can be compared among populations and that it cannot produce information regarding specific physical

activities carried out (al-Oumaoui, , Jiménez-Brobeil, and du Souich 2004). Nevertheless, the results of their method, used with historical and archaeological data, produced meaningful information regarding general patterns of activity. al-Oumaoui and colleagues provided a novel approach to studying entheseal changes through the use of multivariate statistics.

Milella and colleagues (2015) took a similar approach to al-Oumaoui and colleagues (2004) use of a multivariate method. A multivariate approach seeks to "explore patterns of similarity/dissimilarity in a sample with regard to several variables…" (Milella et al. 2015:216) The method was applied to a skeletal sample of 372 male individuals with known occupations from Italy and Portugal. The authors utilized previously collected data from Milella and colleagues (2012) and Cardoso (2006). The Italian sample data used Mariotti and colleagues (2007) method to analyze the data. The Portuguese sample of Cardoso utilized Hawkey and Merbs method. As a result, the scores for Mariotti's robuscitiy and Hawkey and Merbs were converted to a simple presence and absence scoring (Milella et al. 2015). A nonlinear principal component analysis was then applied to these binary data as well as cluster analyses. Age was also controlled for. The results of the analysis revealed three main classes. Class 1 reflected occupations related to farming, Class 2 related to physically demanding occupations not related to farming, and Class 3 represented physically undemanding occupations (Milella et al. 2015). However, the authors created different datasets that related to scoring sides separately, sides pooled, and asymmetry and these differences were not always found across the datasets. For example, differences between certain classes are only found in some, but not all, datasets (Milella et al. 2015). This inconsistency was most likely due to the use of nonmetric data.

Nonetheless, Milella and colleagues' method highlights the use of alternative statistical methods when analyzing entheseal changes.

The evolution of qualitative methods was influenced by medical data that brought new insights into bioarchaeologists understanding of entheses. This in turn led to methodologies that looked to address the nuances that surrounded entheseal changes. At the same time, this led to methodologies improving on areas that previous methodologies failed to address. However, certain methodologies could only be used in specific or niche situations and still suffered from lacking statistical power.

Issues and limitations with qualitative methods

There have been critiques on the limitation of qualitative methods. Wilczak (1998) provides the drawbacks of using categorical data. One of the drawbacks on this scoring method is that it assigns an ordinal value that eliminates the variation. That is to say that an individual or a population may show a greater degree of entheseal change than another but that would not be expressed through the scoring system due to it being assigned to an absolute value (Wilczak 1998). Another issue with many of these qualitative methods is that they do not account for the influence of the "robustness or gracility of the underlying bone," on scoring (Wilczak 1998, 312) as this can skew results with populations that have large differences in skeletal size or between sexes. This is especially crucial in the methods that focus on the examination of robusticity. Similarly, some qualitative methods fail to consider confounding factors that may affect the expression of entheses like age, body size, sex, genetics, and pathologies. Or ignore that entheseal changes have a multifactorial etiology (Henderson and Cardoso 2013; Henderson et al. 2013; Henderson et al. 2017; Henderson and Nikita 2016; Jurmain et al 2012; Wilczak 1998). As

a result, methods can either oversimply the etiology or assume the etiology to be linked to one primary factor (Jurmain et al. 2012; Villotte et al. 2016). For example, with the Hawkey and Merbs method the assumption is that the expression of robusticity and stress lesions were a continuum of increased microtrauma. In addition, the method assumes that ossification exostoses represented abrupt microtrauma. Similarly, Molnar's method treated robusticity and stress lesions as a continuum as well, but unlike Hawkey and Merbs who scored these separately, Molnar scored all features in one category. With these methods there was no consideration of clinical or histological literature (Henderson et al. 2013; Jurmain et al. 2012). Mariotti's method was also constructed without reference to medical literature and a discussion of anatomy or etiology of changes of entheses (Henderson et al 2013.) These are issues that involve understanding the anatomy and etiology of entheses.

There are also logistical issues regarding qualitative methods. Davis and colleagues (2013) point out that qualitative methods such as the Hawkey and Merbs (1995) have very poor repeatability. Repeatability impacts the validity of interpretations and affects how data can be compared. The are a myriad of factors that can cause issues of repeatability and high rates of intra- and interobserver error. This may stem from reducing variability of enthesis morphology into discrete categories, creating categories with too many criteria, using vague terminology when describing features, conditions of the environment of observation, and the observer's ability in identifying features (Davis et al. 2013; Wilczak et al. 2017). Furthermore, the application of qualitative methods involves a relatively high degree of subjectivity. For example, in Mariotti's method the standard description for each of the 23 entheses described in their study differs (Mariotti et al 2004, 2007). Each entheseal site was given its own set of criteria for

determining the stages of development. Although the method includes subcategories for the purpose of flexibility it also opens it up to a great amount of subjectivity which then leads to a higher rate of intra- and interobserver error. Villotte's method also had issues with the use of a three-stage scale and found that revising the method to have a simple dichotomy of presence and absence decreased error.

Further subjectivity arises from having to visually score entheses. Qualitative methods rely on the naked eye to discern entheses features and expressions. As a result, there can be differing opinions between researchers when deciding the degree of expression or what constitutes as a feature. Even when methods provide a more detailed explanation for the criteria required in scoring there are still difficulties. As Wilczak and colleagues (2017) point out, it is challenging to communicate the nuances of a method to those who were not trained by the developers of the method. Instead, training will involve the use of publications, digital photographs, or models (Wilczak et al. 2017). These modes of second-hand training create an obstacle. For example, Wilczak and colleagues (2017) held an online training session for the Coimbra method using digital photos. The Coimbra method was designed to score macroscopic features that can clearly be seen with the naked eye. The use of digital photos allows observers to manipulate the image and see features that might not have been seen if the observer were to view the enthesis in person. Other discrepancies that can occur result from the conditions of observation. Lighting plays a big role in being able to properly analyze an enthesis (Henderson et al 2016; Wilczak et al. 2017). Observers who are in environments that do not provide sufficient lighting may come to different conclusions in interpretations. Observers can also suffer from

fatigue if they do not take breaks when collecting data and this can lead to poor evaluations (Wilczak et al. 2017).

Another limitation of qualitative methods relates to the statistical analyses involved. Qualitative data restricts the type of statistical analyses that can be performed and therefore can limit the type of information derived. Many of the qualitative methods above depended on "parametric or nonparametric univariate and bivariate statistical protocols." (Milella et al. 2015:216) One issue with these types of statistical analyses is that they have prior assumptions about the data. Other issues with these types of statistical analyses are that they often have low statistical power and only allow for one or two variables to be looked at or compared simultaneously. This is problematic for studying entheses and entheseal changes because they are a result of muscles acting in groups and not independently (Henderson and Cardoso 2013; Stefanović and Porčić 2013). When variables are analyzed in these statistical tests, they are treated separately rather than all together. As a result, the complexities involved with the process of entheseal changes are oversimplified and results can be misleading. For example, the Hawkey and Merbs (1995) method uses ranking statistics. This method compares the order of ranked means of entheses scores between two groups. If the order of score means is different than the implication is that the groups engaged in different activities (Stefanović and Porčić 2013). However, that is not the case, as the differences can be a result of confounding factors like body size. The Coimbra method used ordinal regressions which restricted analyses to use only one dependent variable at a time. Out of the qualitative methods mentioned, al-Oumaoui's method and Milella and colleagues' method attempted to use a multivariate statistical method (see later

discussion on quantitative methods). In both cases the authors were hindered by the lack of data and/or use of binary categorical data.

Despite these limitations, qualitative methods are still widely used because they are easy to apply, they do not require specialized equipment and they can produce meaningful results under the right circumstances. However, there are researchers who have recognized that these limitations keep the field from moving forward and have attempted to break away from qualitative methodologies. Instead, they have worked towards creating completely new methodologies that focused on quantitative scoring.

Quantitative methodologies

Quantitative methods have looked to overcome the limitations of qualitative methods. Entheses are complex three-dimensional structures, and the use of qualitative methods may not sufficiently capture its morphology. As a result, researchers have adopted the use quantitative methods that look to directly measure the features of entheses. More contemporary methods have made use of three-dimensional technology to quantify entheseal surfaces through three-dimensional scanning. These methods give more precise representations of entheses morphology because they account for its size (Noldner and Edgar 2013).

Early quantitative methods

Wilczak (1998) provided a critique on the limitations of qualitative methods and instead proposed a quantitative method for scoring entheses. The author's method focuses on the direct measurement of the enthesis using two-dimensional areas, allometric-size standardization, and the identification of lateral asymmetry. Wilczak's method involved quantifying the muscle

insertion sites through the use of videotaping them and transferring them into an imaging software. The insertion sites were delineated with white chalk in order to make it stand out in the videotaping for measurements. Wilczak defined the insertion area as "both the area covered by the tendious attachment and areas immediately adjacent to this attachment which show the formation of new bone on the smooth cortical surface." (Wilczak 1998:314). Wilczak's method is an early example of how quantitative methods differ from qualitative methods. Instead of using arbitrary numbers to evaluate an enthesis, Wilczak directly measured the entheseal site itself. Another difference is that this method also considers intrinsic factors that could potentially influence entheseal changes. Wilczak's method demonstrated how a quantitative method can directly address confounding factors such as age and body size.

Three-dimensional scanning

The advancement of scanning technology has proven to be useful in the study of entheseal changes. Quantitative methods have begun incorporating the use of three-dimensional scanning technology to quantify entheses. In comparison to two-dimensional methods, three-dimensional methods allow researchers to quantify an enthesis size and its surface topography. The use of three-dimensional scanning has allowed researchers to examine the complexity, structure, and shape of entheses.

Zumwalt (2005) proposed a new method for analyzing entheses where the surface of enthesis was measured using a laser scanner and analyzed using fractal analysis. However, this method was tested on sheep remains rather than human remains. Laser scans were viewed using Geographic Information Software (GIS) which gave the researcher the ability to view the scan in three dimensions and freely rotate it to examine all aspects of morphology (Zumwalt 2005).

Fractal analysis was applied to examine the entheseal surfaces. Fractal analysis examines the surface complexity of the entheseal area of interest. Through fractal analysis the difference between entheseal surfaces as well as difference within an entheseal surface can be examined simultaneously (Zumwalt 2005). The results of the study found that the exercised sheep had no significant differences when compared to the sedentary animals. Despite the results indicating that there was no relationship between entheseal morphology and activity, Zumwalt provided a method that allowed variation in the entheseal surface to be measured and assessed. Zumwalt also showcased the potential applications of 3D scanning technology and how it allows researchers to examine details not easily discernable to the naked eye.

Nolte and Wilczak (2013) provide a three-dimensional method that was applied to biceps brachii enthesis of a human skeletal sample. Their method makes use of the NextEngine™ 3D scanner. This scanner provides high-resolution surface topography and allows surface area to be measured in length, width, and elevation. Before scanning the enthesis site, the authors first delineated the enthesis area using chalk artist's pencils. Delineation of entheseal areas were determined with the assistance of the macroscopic grain of bone. Two independent outlines were made. One that represents the rugose area and the other that traces the tuberosity area. The rugose area included the osteophytes and crests of the bone, distinctions in bone density, and grain and texture. The tuberosity area included the raised biceps tuberosity (Nolte and Wilczak 2013). Scans were taken from multiple angles and fused together using a processing tool provided by the scanner to create the three-dimensional image. In addition to measuring the entheseal area, measurements of the humerus and radius were taken for body size proxies. The results of their study found that there was a strong correlation between the body size proxies and

entheseal areas, sex and entheseal areas, and age and entheseal areas. Furthermore, the two measures of entheseal areas were strongly correlated with one another (Nolte and Wilczak 2013). Nolte and Wilczak's method expand on Wilczak's (1998) method. Instead of measuring a two-dimensional area, they make use of 3D-scanning technology to measure a three-dimensional area, which is more representative of the enthesis. Like Wilczak's method, their method also assesses confounding factors. More specifically body size is evaluated, and the authors take it a step further by demonstrating the relationship between body size and entheseal change. Nolte and Wilczak's method underline the importance of considering confounding factors in the interpretation of entheseal changes. Their method also demonstrates some of the statistical advantages of quantifying entheses.

Berthon and colleagues (2015) present a different three-dimensional method that uses micro-tomodensitometry to study the microarchitecture of entheseal changes. Micro-tomodensitometry involves using micro-computed tomography (micro-CT) in conjunction with 3D reconstruction imaging software. This allows the researcher to not only create a 3D reconstruction of the enthesis, but it also allows them to see the trabecular bone microarchitecture. The authors utilized micro-tomodensitometry for the purpose of comparing the microarchitecture of normal diaphyseal bone with the entheseal change. Results of the analysis demonstrated that the microarchitecture of the entheseal change regions were different, most notably the organization of bone and canals. Although this method is primary a qualitative method, it serves as a preliminary method for the incorporation of a quantitative aspect in studying the microarchitecture of entheseal changes. Further exploration of this method may have potential in helping researchers delineate entheseal borders by elucidating where an

enthesis may start through its microarchitecture. Similar to Zumwalt, Berthon and colleagues showcase the potential of 3D scanning technology and how it can enhance our understanding of entheseal changes. They presented an experimental way to examine entheses through comparing the microarchitecture of an enthesis with the normal surrounding bone.

Three-dimensional multivariate analysis

Karakostis and Lorenzo (2016) introduced a three-dimensional scanning method that focuses on using multivariate statistical analysis. Previous studies and methods involving the analysis of entheses focused on singular areas. These methods also treated them separately from one another. The method proposed by Karakostis and Lorenzo examines multiple entheses and multiple aspects of entheses in conjunction with one another. The advantage of this is that the result of the analysis provides an amalgamation of similarities and differences across multiple entheses. The main statistical procedure used in this method is principal component analysis (PCA). Unlike other statistical approaches there is no *a priori* group categorization. This means that groups can more easily be discerned based on the similarities and differences. Another important aspect of their method is the use of three-dimensional scanning and 3D imaging software. The use of these tools enabled the authors to more easily delineate and measure entheses. This is carried out through the assistance of 3D imaging software that reduces the subjectivity in delineation by providing filtering tools. These filtering tools help highlight the entheseal area by emphasizing the characteristics of an enthesis. This includes coloration, surface topography, and elevation (Karakostis and Lorenzo 2016; Karakostis et al. 2017). Furthermore, measurement of the area is done through the software as well which eliminates having to rely on taking any type of manual measurement. This differs from Nolte and Wilczak (2013) method

which delineated the enthesis by using chalk outlines. In 2017, Karakostis and colleagues applied this method to the Spitalfriedhof Saint Johann collection. The benefit of this collection is that it possesses extensive documentation on each individual in regard to social and demographic information. This allowed the authors to compare the results with known information. The results of the analyses showed two morphological patterns. One group that reflected a power-grasping pattern and another showing precision grasping pattern. The documentation revealed that those in the power-grasping pattern were part of a group of life-long construction workers, whereas the other group was part of long-term precision workers like a seamstress or tailor (Karakostis et al 2017). The results of their study showed a close association between the multivariate patterns of the entheseal morphology and occupational activities. The results also demonstrated that the overall size of entheses were significantly correlated with body weight as well as the hand bones lengths. Additionally, the results were not associated with confounding factors like age, sex, pathological conditions, and genetic variation (Karakostis et al. 2017). Karakostis and colleagues' method has demonstrated that it is possible to assess the impact of physical activity on entheseal size and shape, while at the same time controlling for confounding factors. Compared to Zumwalt (2005, 2006) and Berthon et al. (2015), Karakostis and Lorenzo uses multivariate statistics and focuses on macroscopic characteristics of entheses. They also consider the confounding factors that can affect entheseal change. Karakostis and Lorenzo's (2016) method has proven to be useful as it has been successfully applied to a variety of other studies (e.g. Karakostis et al 2018a, 2018b; Karakostis, Jeffrey, and Harvati 2019; Karakostis et al. 2019).

Quantitative methods succeeded in addressing some of the limitations and issues innate to qualitative methods. The incorporation of metric data and the adoption of three-dimensional scanning technology has allowed researchers to approach the study of entheseal changes in different ways. Researchers can apply different types of statistical methods and explore other avenues of analyzing entheseal changes that would otherwise not be capable through older methods. However, no method is perfect and quantitative methods are also subject to their own limitations and issues.

Issues and limitations of quantitative methodologies

Although quantitative methods have avoided many of the issues that arise with qualitative methods, they are far from perfect. Similar to qualitative methods, there is still a degree of subjectivity to delineating and measuring entheses. Observers still have to decide where to chalk the margins before using the 3D method and these boundary lines can be somewhat arbitrary (Noldner and Edgar 2013; Nolte and Wilczak 2013). Measurements errors can also still occur. Another drawback to quantitative methods is that the whole insertion site is required to be intact in order for a value to be assigned, whereas with qualitative methods can still assign a score based on portions that are visible (Noldner and Edgar 2013). Other issues with quantitative methods relate to logistics and feasibility, especially with 3D methods. 3D methods can be quite time consuming, and collection of data is not as expedient as using visual scoring. Additionally, it costs more to use 3D scanning in comparison to visual scoring. 3D scanning is also not as readily available as visual scoring (Noldner and Edgar 2013). There is also training involved in how to properly use the technology used and how to properly scan the entheses.

Noldner and Edgar (2013) did a comparison between different methodologies to assess the benefits and drawbacks of each. They examined a three-dimensional scanning method, an ordinal method, and two-dimensional area measurement method. The authors tested enthesis asymmetry then compared the results of each method to see if there was any agreement with one another. The three methods utilized were Nolte and Wilczak (2013), Hawkey and Merbs (1995), and Wilczak (1998). The results of the analysis demonstrated that ordinal and three-dimensional results agreed with one another. Moreover, when comparing the data between the three-dimensional and two-dimensional methods there was significant correlation between the two. This correlation reflects that in this situation both methods were comparable in representing entheseal development. The authors also compared intraobserver rates between the methods and found that the three-dimensional method yielded better error rates. However, the error rates for the three-dimensional method were still fairly high and could be further refined. Karakostis and Lorenzo (2016) method does address this issue. Their method utilizes computer calculations to delineate entheseal borders and has proven to have low error rates. Noldner and Edgar (2013) demonstrate that in certain situations it is beneficial to utilize ordinal methodology, but three-dimensional methods offer more in terms of data inquiry and data flexibility.

Research moving forward

Methodology is crucial in the study of entheseal changes. Without the proper use of scoring methods (i.e. qualitative vs quantitative) and methods of statistical analysis it can lead to inconsistent results that can either oversimplify the complexities of entheses or provide misinformation. Despite the limitations in quantitative methodologies, it has proven to be a more advantageous and suitable avenue to take in the study of entheseal changes. Three-dimensional

multivariate analysis has produced some of the best supporting evidence for the affects of physical activity on entheses thus far. Researchers who have applied the use of a multivariate approach feel that it is the most appropriate way to analyze entheseal changes (Karakostis et al. 2017, 2018a, 2018b, 2019; Karakostis, Jeffrey, and Harvati 2019; Milella et al. 2015; Stefanović and Porčić 2013; Yonemoto 2016). This approach accounts for the multifactorial etiology of entheses and deals with confounding factors with ease. The multifactorial approach also eliminates any type of a priori assumptions and lets the data form patterns. Further refinement of these methods will prove to be a beneficial tool for bioarchaeologists whose interest is in the study of entheseal changes. Currently, there is a lack of data on the application of these methods to the human skeleton. In so far, the only cases have been provided by Karakostis and colleagues (2016; 2017; 2018a; 2018b; 2019; 2021) studies on hand entheses. Therefore, the objective of this study is to see the feasibility of Karakostis and colleagues (2016) method to entheses outside the hand.

Chapter 2: Materials and Method

Materials

The teaching collection housed in the osteology lab of San Francisco State University was used for analysis. A total of twenty-three humeri from this collection were included. The individuals in the collection are well preserved, with most having humeral heads and distal ends fully intact. There is no documented background info on this collection and what is known is that the remains in this collection were purchased from biological and medical supply companies. It should be prefaced that any individual that exhibited any gross pathological change that might affect the entheses such as bone fractures or osteoporosis was not included in this study. Additionally, individuals that have any damage to the entheseal sites were not considered. Initially there were twenty-four humeri examined and labeled a respective number from 1 to 24. However, humerus number 11 turned out to possess extreme amounts of osteophytes and was therefore not included in the analysis. Six fibrocartilaginous entheses were examined on each individual humerus. These include the common extensor origin, common flexor origin, infraspinatus, supraspinatus, subscapularis insertion, and teres minor. The infrapinatous, supraspinatus, and teres minor were analyzed together due to difficulties in accurately delineating each of the entheses. Therefore, a total of 4 entheseal groups were analyzed in this study.

Sex Estimation

Sex was estimated from the humeri since they were the only element available to assess. A combination of both visual and metric techniques was utilized to estimate sex. The visual method was the primary method for estimating sex as it is independent of size and robusticity. Metric methods were used to supplement the results. The visual method utilized in this study looks at five features of the distal end of the humerus. These features include the orientation of the medial aspect of the trochlea relative to the shaft, the trochlear constriction, the trochlear symmetry, the shape and depth of the olecranon fossa, and the angle of the medial epicondyle (Rogers 1998). Each of these features were scored either male or female based on the criteria. Roger's (1998) method has been successful in multiple populations including European, African, Native American, South Asian, and Hispanic.

Metric methods involve taking specific measurement from the humerus. These include taking maximum length, vertical head diameter, and epicondylar width (Mall et al. 2001; Stewart 1979). Two metric methods were applied. One method was developed by Mall and colleagues (2001) who conducted a discriminant analysis on these three measurements and came up with a discriminant function that has a reported accuracy of 93 percent in a contemporary middle European population. The discriminant score is calculated as follows: $D = 0.196$ maximum length (cm) $+ 1.962$ head diameter (cm) $+ 1.160$ epicondylar width (cm) $- 22.608$. If D is less than or equal to 0.30 the individual is most likely female, if it is greater than 0.30 the individual is most likely male. The other method was developed by Stewart (1979) who uses only the measurement of the vertical head diameter to estimate sex. Stewart's method provides ranges for the exact measurements taken of the vertical head diameter. If the diameter is less than 44mm the

individual is female, if it is greater than 47mm the individual is male, anything in between 44mm

and 46mm are indeterminate (Stewart 1979). This method was conducted on the Terry collection

and therefore has only been successful with African and European populations.

Body Size and Age-at-death

Body size was also estimated to see the degree of influence it has on the morphometrics

of entheses. In a similar way to estimating sex, body size was estimated using metric

measurements of the humerus. More specifically the biepicondylar breadth, maximum length,

and vertical head diameter can be used as body size proxies (Mall et al. 2001; Nolte and Wilczak

2013). These proxies were then analyzed in conjunction with the entheseal measurements to

assess the correlation between them. This was done for each individual fibrocartilaginous

enthesis measured as well as the rotator cuff as a whole. Age-at-death estimates were not

available due to only having the humeri. Only adults were included in the sample based on full

fusion of the epiphyses.

3D Scanning and Inspection

Three-dimensional scans were taken using the Shining 3D Transcan C scanner and the

accompanying program EXScan C. The Shining 3D Transcan C uses white LED light and

structured-light scanning. Scans were taken at the 150mm x 96mm scanning range using the

highest resolution setting (0.0375mm). Scans were also taken in color. A turntable was used to

take scans of 16 different aspects of the bone. Due to the scan range, the proximal and distal ends

of the humerus were scanned separately resulting in two separate files. Additional snapshot scans

were taken to fill any gaps. The scans were cleaned up in EXScan C and exported as PLY files.

Meshlab was used for the 3D inspection, delineation of the enthesis, and measurement of the

entheseal surface.

Statistical Analysis and Error Testing

Statistical analysis was conducted using SPSS version 28 and Microsoft excel. Statistical

analysis followed the protocol laid forth by Karakostis (2022). This publication provided step-

by-step guidance on conducting proper statistical analysis using the V.E.R.A. method. A method

that is reliable and repeatable will have low inter and intra observer error. Therefore, precision

tests were employed, and multiple statistical tests were run on the measurements taken. The same

bone scan was used for each corresponding bone. Only delineations and measurements were

repeated. Two repetitions were taken to use for intra-observer error testing. Each repetition was

taken two days apart. Inter-observer measurements were compared with those taken by Dr.

Karakostis and Dr. Wilczak. Inter-observer error was briefly tested using seven random bones

for the subscapularis and five bones for the supraspinatus, infraspinatus, teres minor, common

extensor origin, and common flexor origin. This included bones 1, 8, 9, 10, 15, 17, and 21 used

for the subscapularis. Bones 1, 7, 8, 18, and 22 were used for the other three entheseal sites

measured. To ensure reliability and repeatability, paired T-tests, Lin's concordance correlation

coefficient, intraclass correlation coefficient (ICC), mean deviations, and percent deviations were

analyzed. The ICC estimates and their 95% confident intervals were calculated using SPSS

statistical package version 28 (SPSS Inc, Chicago, IL) based on single-rater measurement, 2-way

mixed model, and absolute agreement type. Additionally, both univariate and multivariate

statistical analysis were employed to examine the correlation between body size and entheseal

size. This included the use of multivariate analysis of variance (MANOVA), and bivariate

statistical tests.

The V.E.R.A Method

The central method used in this research was the three-dimensional multivariate

statistical analysis method developed by Karakostis and Lorenzo (2016), [coined] as the

Tübingen University Validated Entheses-based Reconstruction of Activity (Abbreviated as the

V.E.R.A. method) (Karakostis and Harvati 2021). The V.E.R.A method was first applied to

examine hand entheses of individuals from a late-medieval San Pablo collection (Karakostis and

Lorenzo 2016). It was further refined and used in another study that looked at hand entheses of

individuals from the Spitalfriedhof Saint Johann collection (Karakostis et al. 2017). This method

has proven to be successful as other studies have produced results that show evidence of a

correlation between physical activity and entheseal changes (e.g. Karakostis et al 2018a, 2018b;

Karakostis, Jeffrey, and Harvati 2019; Karakostis et al. 2019). However, it has only been applied

to human hand entheses.

The V.E.R.A. method analyzes multiple entheses at once and emphasizes the statistical

relationship among different entheseal patterns. This method is done in three stages. The first

stage consists of scanning the bone using a 3D scanner and importing the images into imaging

software that creates a high-resolution 3D model of the bone. The second stage involves

delineating the entheseal surface within the imaging software using three criteria: distinct

elevation, surface complexity, and darker coloration (Karakostis and Lorenzo 2016). However,

Karakostis and Harvati (2021) advise that coloration should be used more as a supplemental

criterion as its expression depends on taphonomic factors. Additionally, once the entheseal

surface is delineated it is then measured. The third stage is conducting statistical analyses using the measurements taken. More specifically, multiple MANOVA and bivariate analyses were performed on the data using SPSS.

The First Stage

The first stage in this method involves creating three-dimensional scan models of the bone. The Shining 3D Transcan C scanner and EXScan C were used to scan and create a total of 46 humerii models. This included 23 scan models of the proximal end of the humerus and 23 scan models of the distal end of the humerus. Scans were taken by the Shining 3D Transcan C scanner and immediately uploaded to EXScan C. The EXScan C program provides the user with tools to create a clean three-dimensional model of the scan. The three-dimensional model is created in two separate steps provided in the program. These steps are "Scan" and "Post Processing" (See figure 2.1 and 2.2).
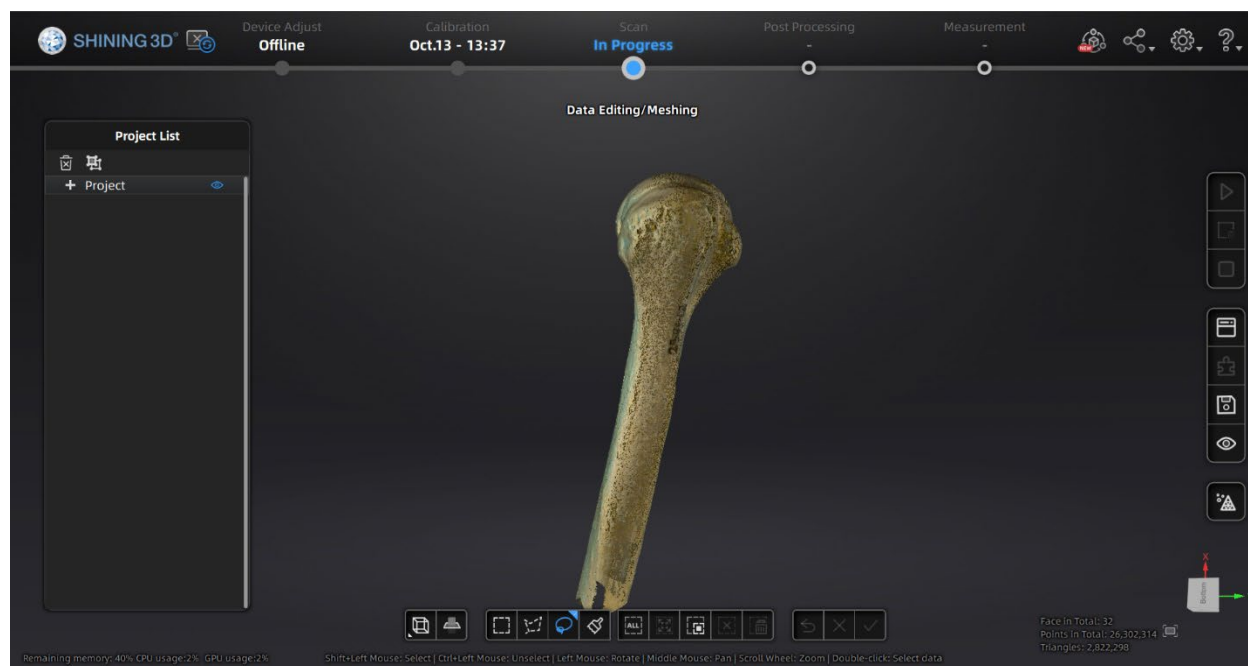
Figure 2.1 Scan step in EXScan C program.
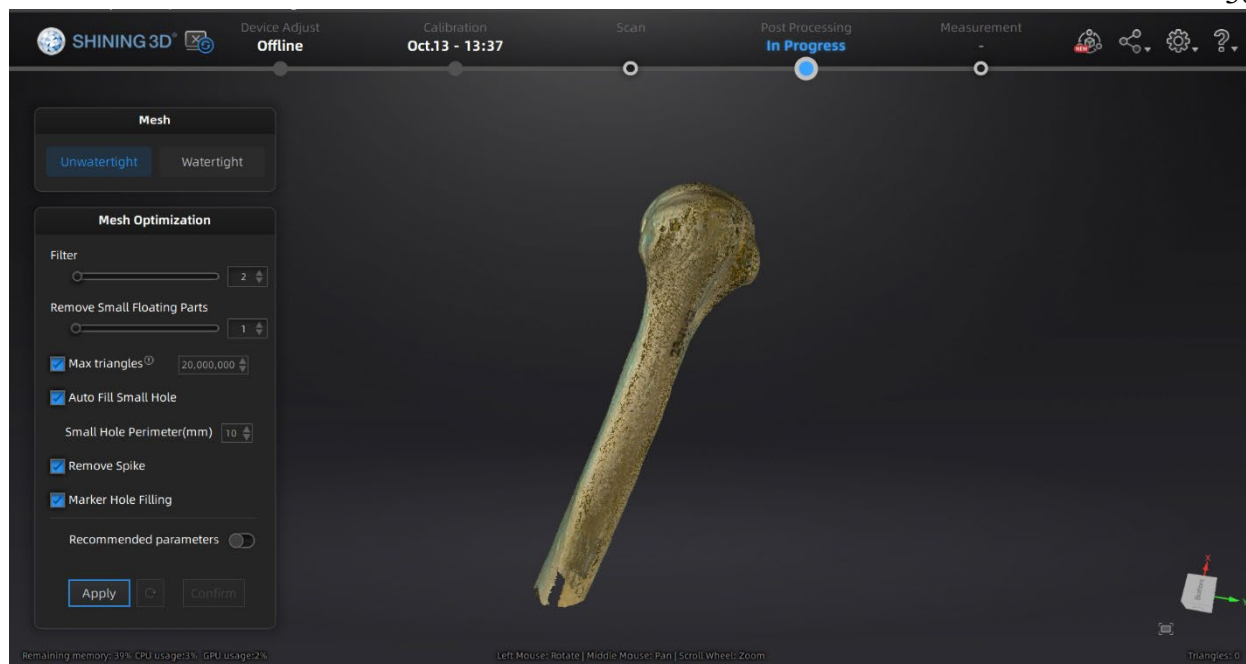
Figure 2.2 Post Processing step in EXScan C program.

In the Scan step each humerii were scanned from 16 different aspects and these 16

separate scans were then stitched together using the alignment feature. Once the scans were

stitched together, they were further cleaned up using the Global Optimization. After the scans are

optimized, the program allows the user to move onto the Post Processing step. In the Post

Processing step a mesh model of the stitched scan is created. The mesh model is further

optimized resulting in a smooth three-dimensional model of the bone. These models were saved

as ply files to use for the second stage.

The Second Stage

The second stage of the V.E.R.A. method involves delineating the entheseal site of

interest. The three-dimensional models created in stage one was imported to MeshLab. Each

entheseal site was delineated individually, meaning that a new scan file was used each time. The

three-dimensional renders of the bones ended up being very detailed. They contained vertices

and faces in the millions. The issue with this is that there are too many data points that interfere with the filters and algorithms used to delineate the entheseal site. To counteract this issue, simplification steps were taken that still maintained the integrity and shape of the original model but made it easier for the algorithms to delineate the entheseal site. Figures 2.3 and 2.4 show the difference between the non-simplified and simplified meshes.
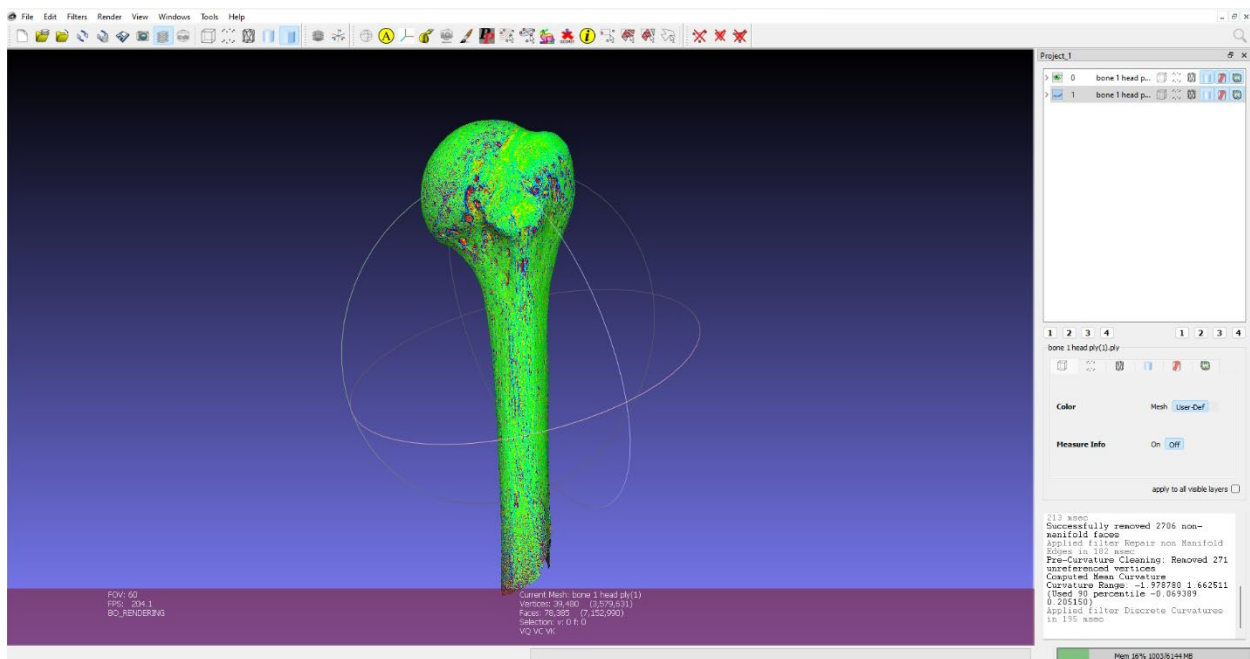
Figure 2.3. Non-simplified mesh of humerus in Meshlab.
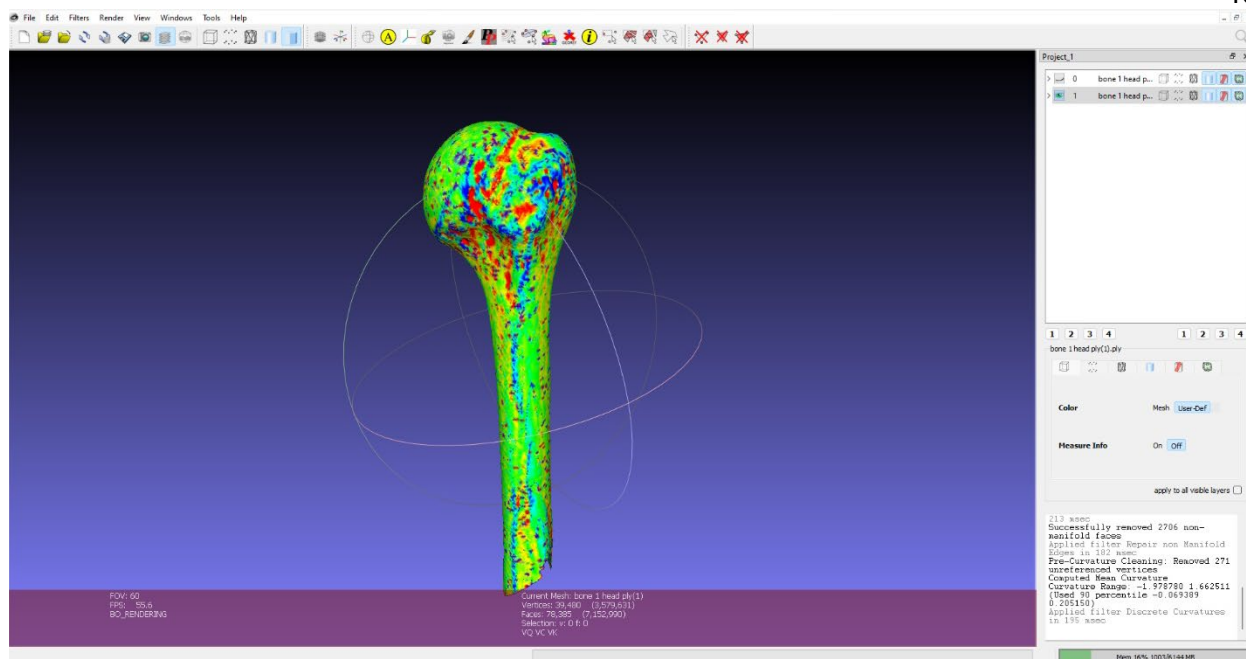
Figure 2.4. Simplified mesh of humerus in Meshlab.

The same simplification method was used for all delineations. Once the ply file was loaded into MeshLab, the first step was to simplify the render. To do this the "Filters" tab was used. The first step in simplifying the render was to "Filters" -> "Remeshing, Simplification and Reconstruction" -> "Simplification: Clustering Decimation". This step reduces the number of vertices and faces of the original model. A cell size of 0.6 was used for simplification. The next step was to smooth over any jagged edges created by the simplification. To do this, the "Filters" tab was used again, specifically "Filters" -> "Smoothing, Fairing and Deformation" -> "Laplacian Smooth". The last step in the sequence is to fix any non-manifold edges created through the smoothing process. If non-manifold edges are present, MeshLab does not allow the user to perform any tasks with the render. To fix this "Filters" -> "Cleaning and Repairing" -> "Repair non-Manifold Edges" was used. With the render properly simplified, the delineation of the entheseal surface can begin.
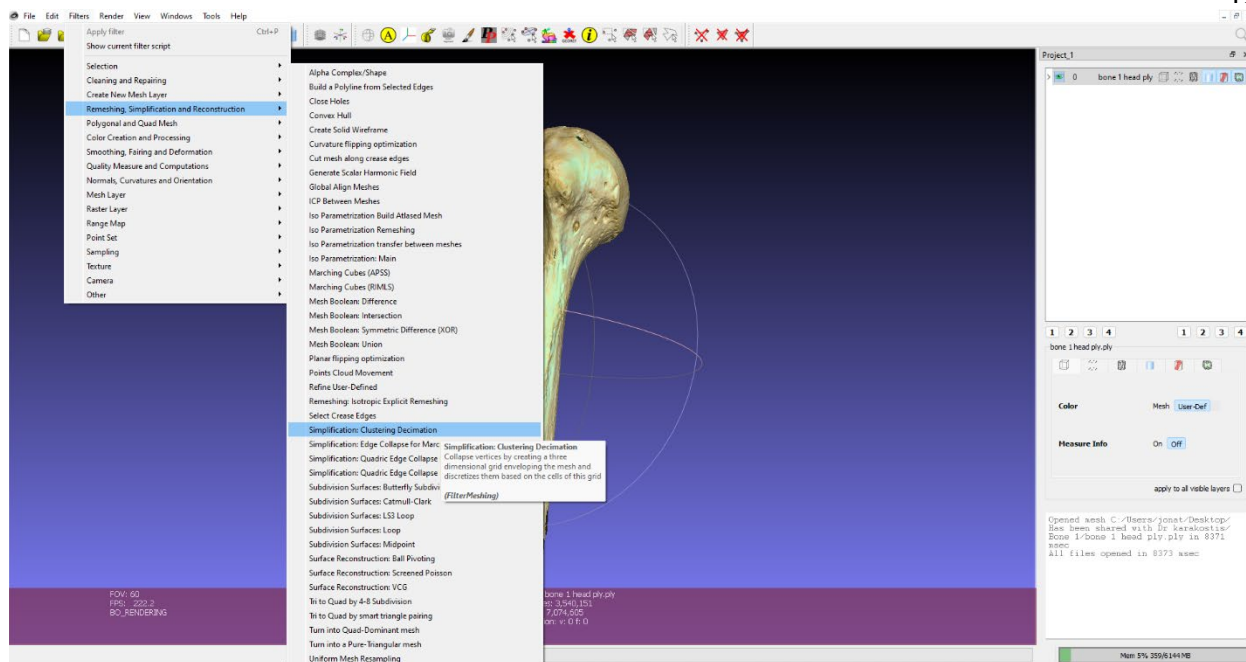
Figure 2.5. Simplification selection step in Meshlab.
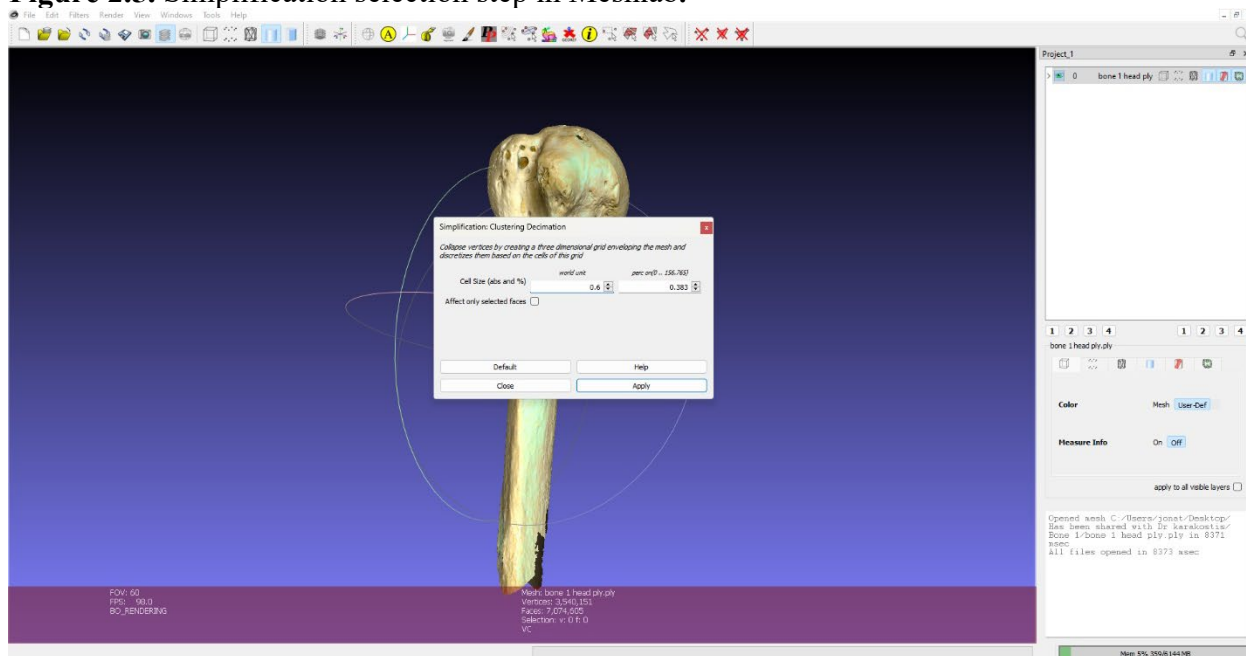
Figure 2.6. Inputting cell size in Meshlab.
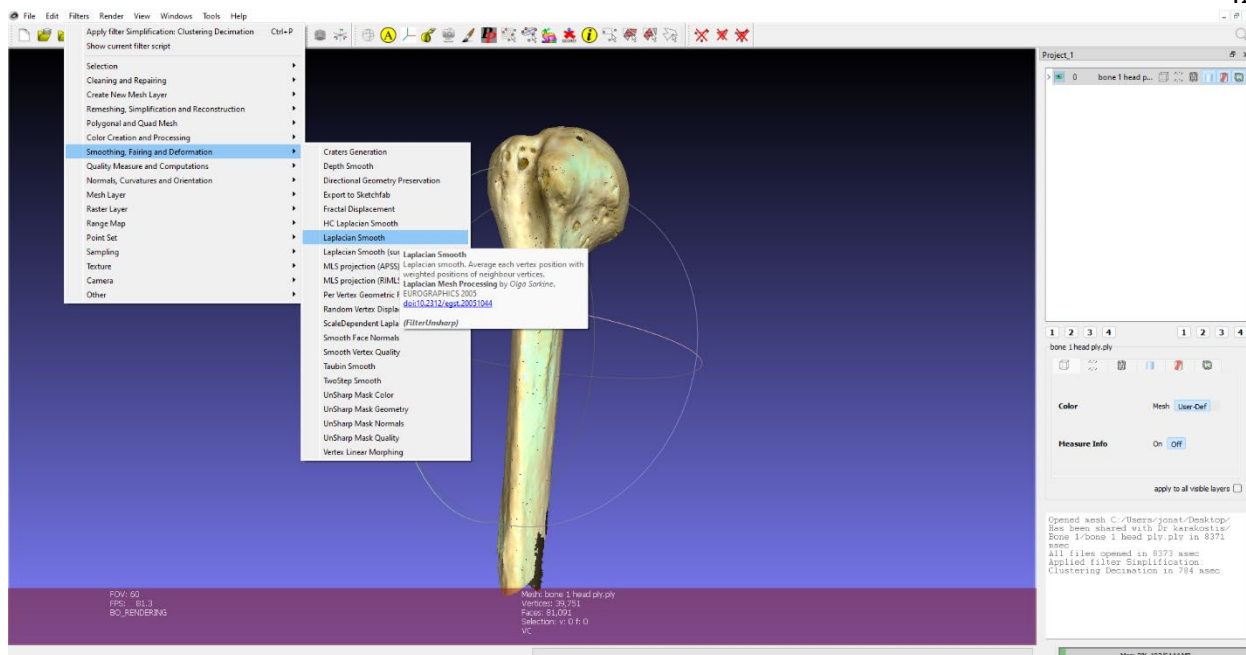
Figure 2.7. Selection of Laplacian smooth step in Meshlab.
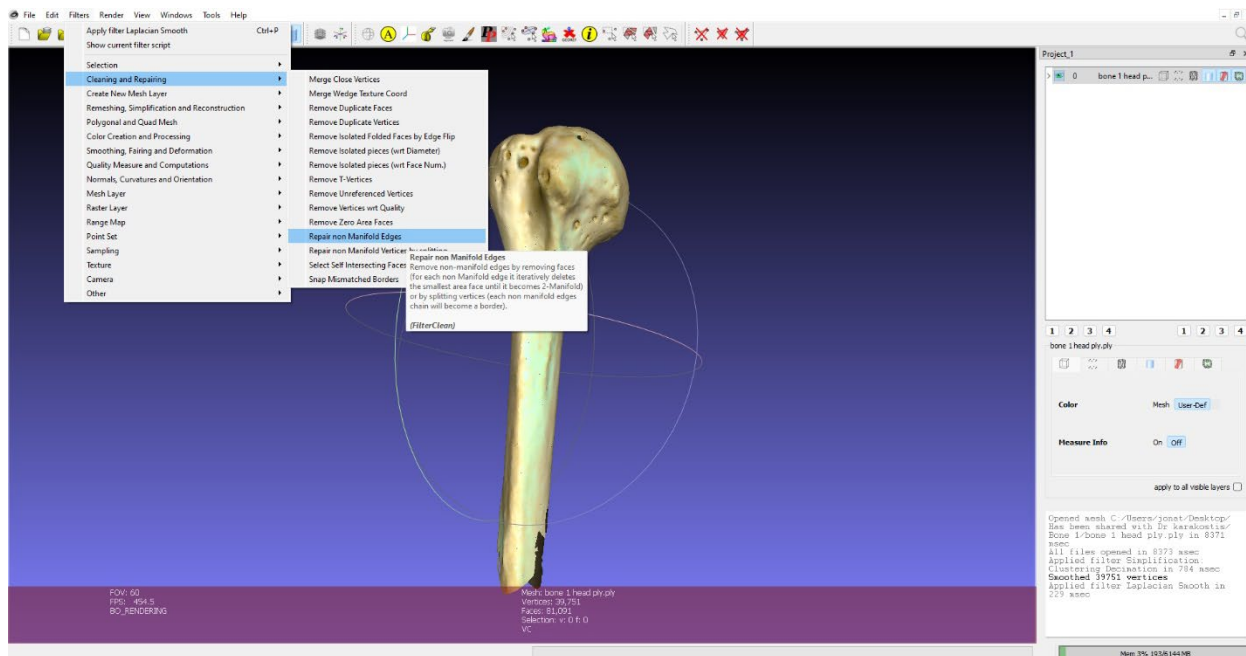
Figure 2.8. Selection of Repair Manifold edges step in Meshlab.

The process of delineating the enthesis involved using filters and selection tools provided by MeshLab. "Filters" -> "Color Creation and Processing" -> "Discrete Curvatures" was used to delineate distinct borders of the enthesis. Then the Z-painting tool was used to select the entheseal area. The selection of the entheseal area is informed by anatomical knowledge and the discrete borders created by the discrete curvatures filter. The enthesis of interest was selected in parts. Selection was taken from different angles to ensure that no parts were missed and then a final check was done to make sure that the correct areas were selected (see figures 2.9-2.12).
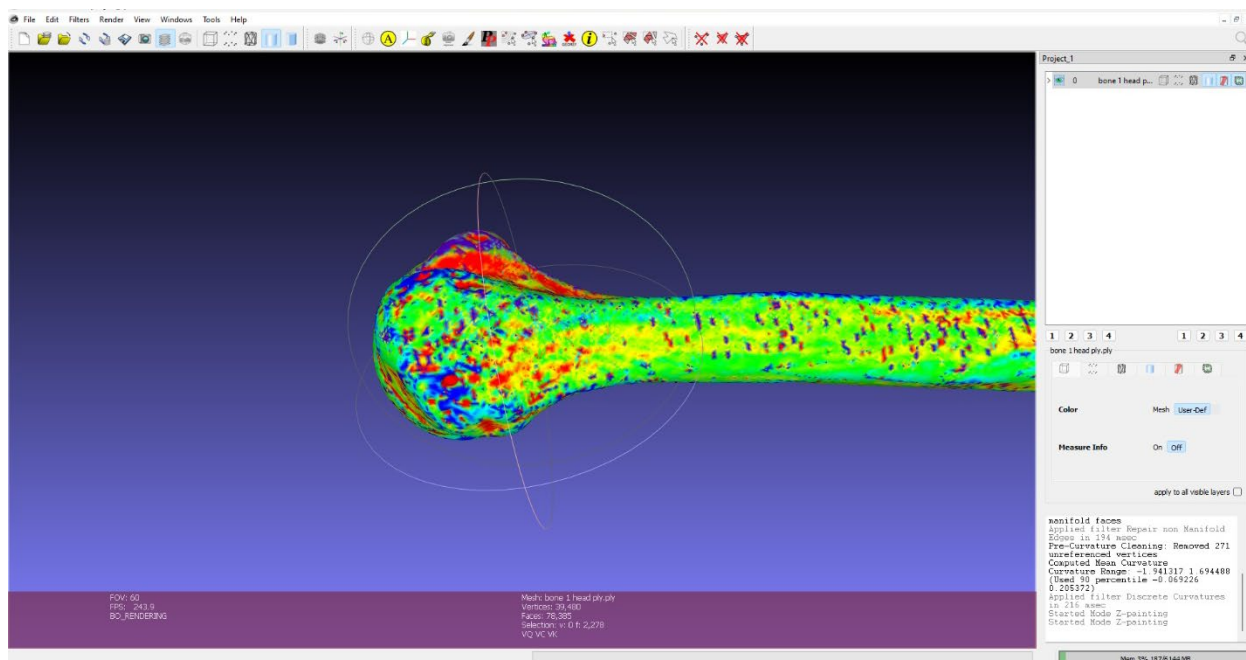
Figure 1.9. Example of how selection was taken of the subscapularis from a lateral view.
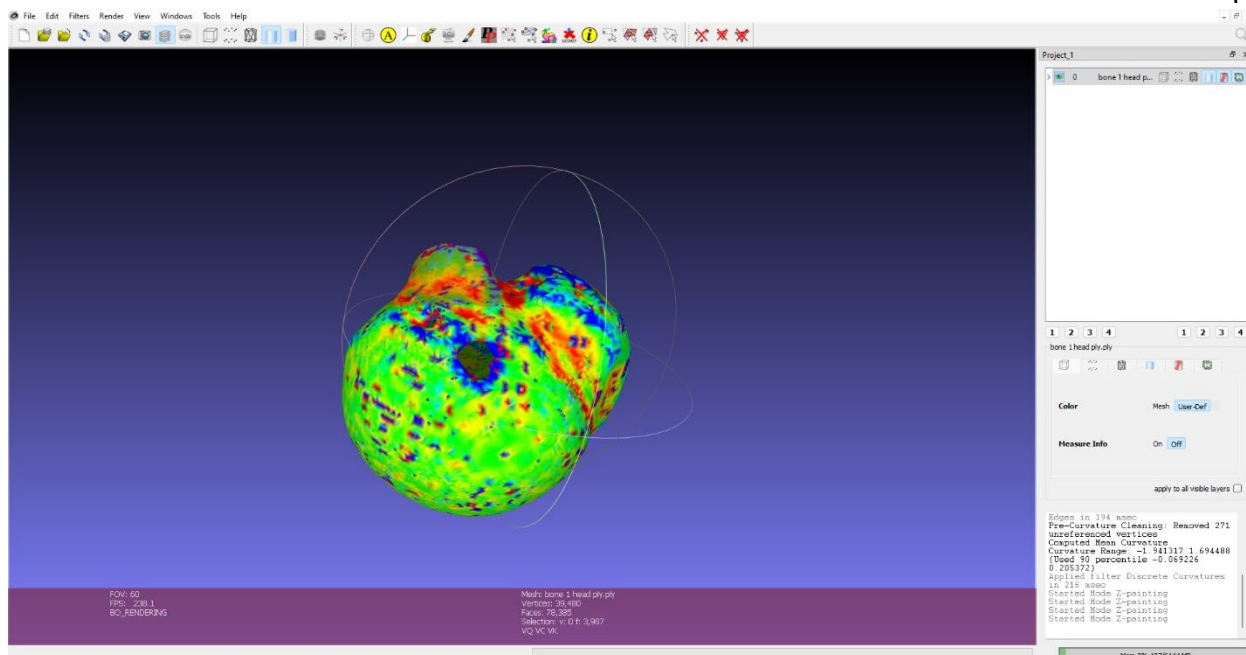
Figure 2.10. Example of how selection was taken of the subscapularis from a superior view.
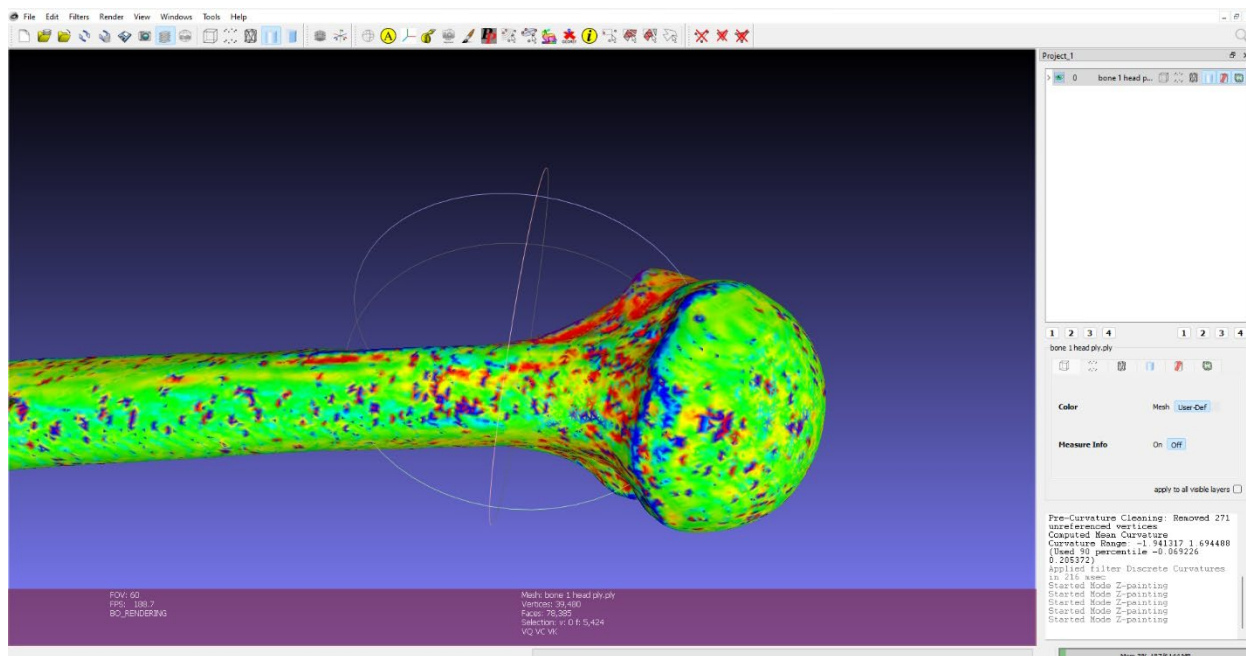
Figure 2.11. Example of how selection was taken of the subscapularis from a medial view.
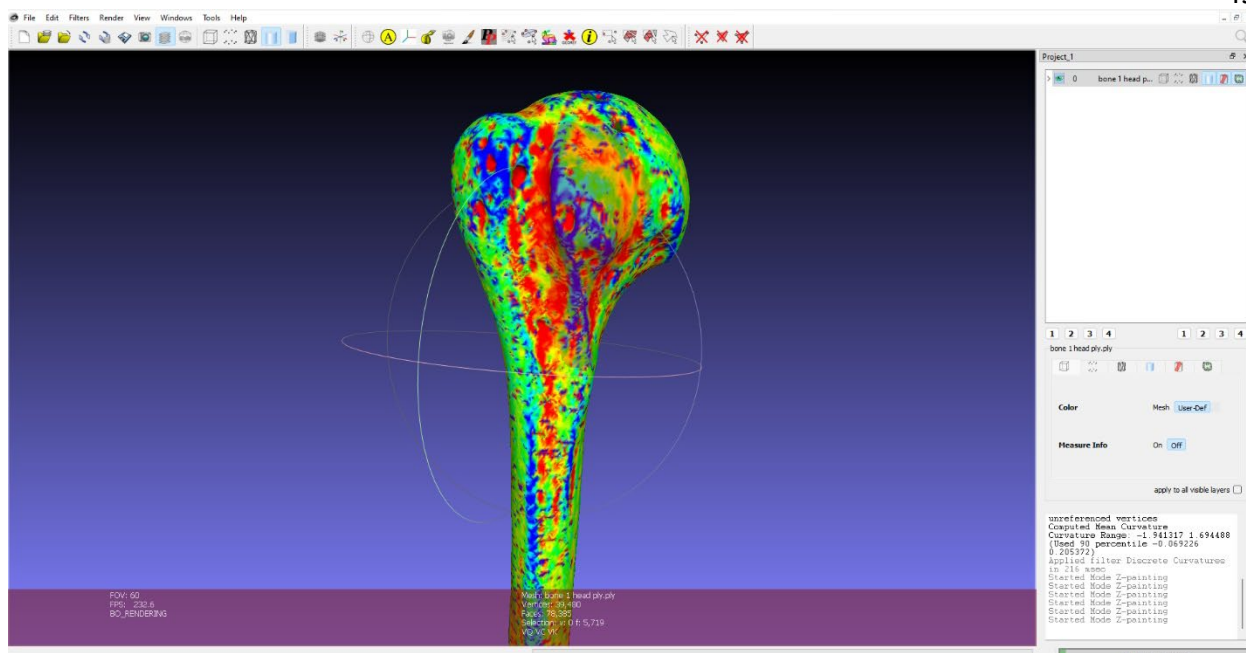
Figure 2.12. Example of what the final selection of the subscapularis looks like.

Another important step in the selection process was selecting a small area around the

enthesis. The importance of this step will be explained when that filter is applied. The scan is

then further cleaned up by removing the un-selected parts by inverting the selected area and

going to "Filters" -> "Selection"-> "Delete Selected Faces and Vertices". Sometimes in the

selection process unnecessary areas may have been selected and it will manifest itself in this step

as floating isolated areas. To fix this, these areas were selected and deleted. Other issues that

sometimes arose during the selection process are when areas are not fully selected resulting in

small holes showing up in the enthesis scan. If there are multiple, noticeably large holes the

enthesis should be re-selected. However, if there are few, small holes MeshLab can fill them in

using "Filters" -> "Remeshing, Simplification and Reconstruction" -> "Close Holes". The next

step after cleaning up any issues is to delineate the enthesis. This is done through applying

"Filters" -> "Color Creation and Processing" -> "Compute curvature principal directions" and

under "Method" selecting "Principal Component Analysis". The result is the general entheseal

area scan having different colorations based off elevation (See Figures 2.13-2.15). The most

outer border (the blue) is the area that reflects the regular bone, hence why the small extra area

was selected in the earlier step. This was used as a base comparison for the filter to compare

elevation. The outer blue area is selected using the Z-painting tool and removed. What is left is

the delineated enthesis. The surface area of the enthesis was then measured using the tools in

MeshLab. "Filters" -> "Quality Measures and Computations" -> "Compute Geometric

Measures" was used to take the surface area measurement in mm^2.
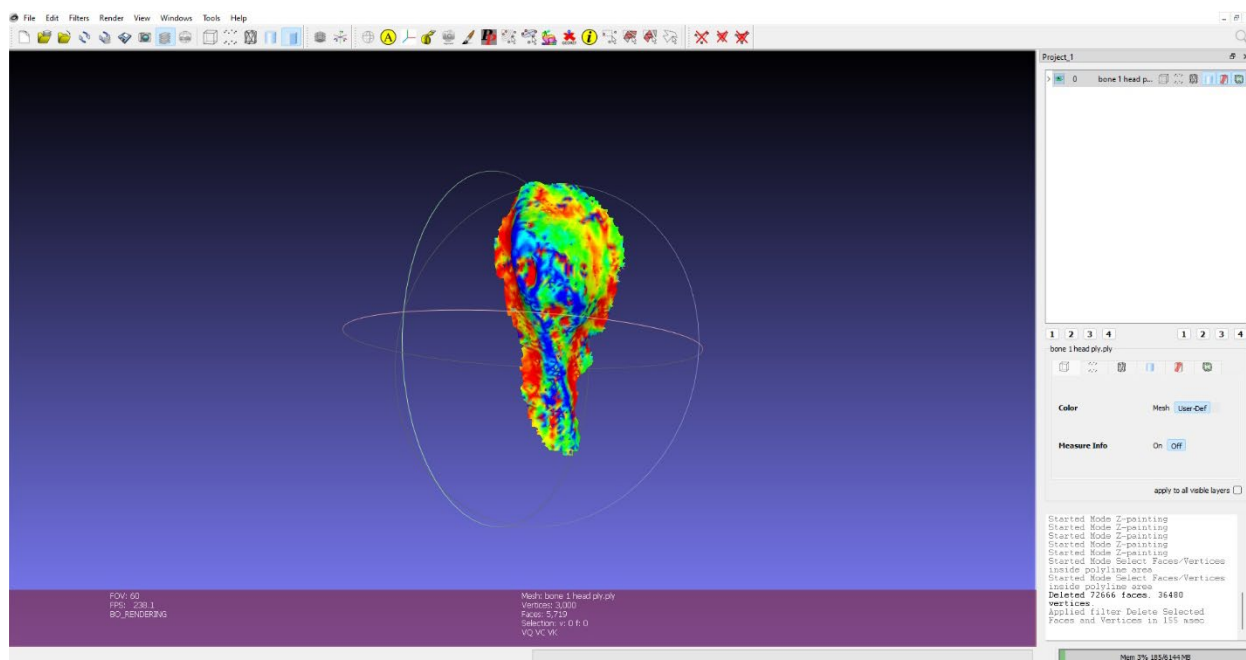
Figure 2.13. Example of what the selected entheseal area looks like after removing unselected areas.

Subscapularis used as an example.
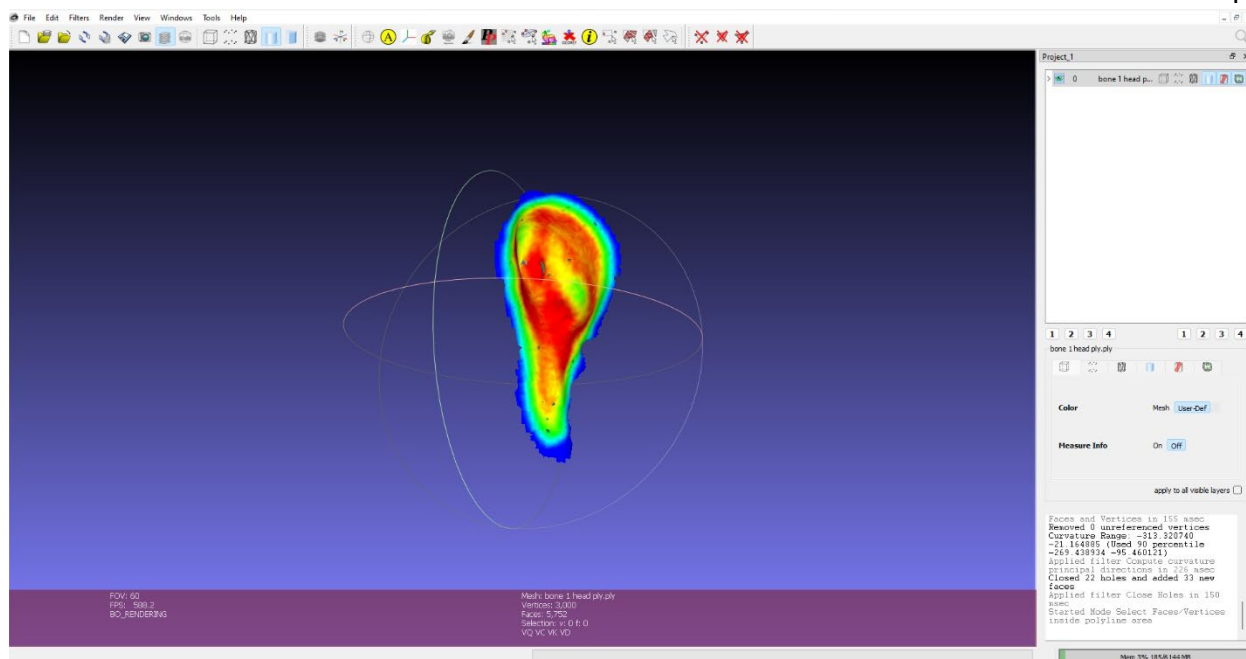
Figure 2.14. Example of what the selected entheseal area looks like after running the compute curvature

principal directions step. Subscapularis used as an example.
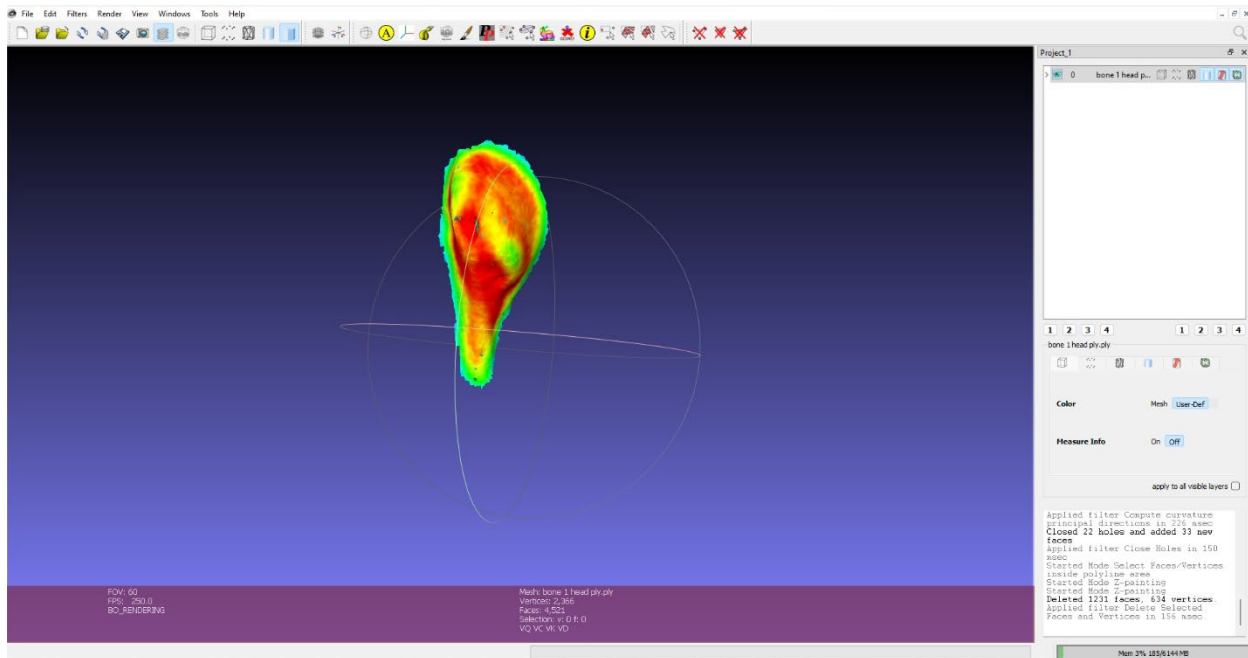
Figure 2.15 Example of what the final selected enthesis used for measurement looks like. Subscapularis used as an example.

The Third Stage

The third stage in the V.E.R.A method compromises of conducting the proper statistical analyses. This is deciding whether it is appropriate to conduct univariate analyses, multivariate analyses, or both on the data set worked with. Univariate analyses were the only relevant statistical protocol to utilize with the data in this thesis. As a result, multiple correlation tests were conducted on the data and Pearson correlation coefficient was used.

Chapter 3: Results

This chapter will produce the results of the skeletal observations, three-dimensional inspection, error-testing, and statistical analyses conducted above. Originally principal component analysis (PCA) was planned to be conducted on the data collected. However, due to the lack of demographic information for the teaching collection along with the small sample size, PCA tests were not able to be conducted. Instead, univariate, bivariate, and MANOVA testing were used.

Skeletal Observations

Table 3.1 displays the size and sex estimations recorded from the humeri in the collection. These descriptive statistics were planned to be used in the analysis but due to a lack of data available it did not seem appropriate. Nevertheless, it is included to demonstrate the process of how this data was collected for the thesis. The distribution of left to right humeri are almost even, with left humeri being slightly more represented. In terms of sex, the distribution heavily leans towards males while there are two females and one indeterminate (See appendix for sexing estimation). Of course, it is always cautioned to take these sex estimations with a grain of salt as they are not based on using the skull or pelvis. Table 3.2 shows the body size proxies taken for each humeri. Figure 3.1 shows the distribution of the entheses measurements. The subscapularis and group of supraspinatus, infraspinatus, teres minor have a greater range of entheseal area in comparison to the common extensor origin and common flexor origin.

Table 3.1. Teaching collection data

Number of humeri	23
Number of left humeri	13
Number of right humeri	10
Number of Males	18
Number of Females	2
Number of indeterminate	1

Table 3.2. Measurements of body size proxies for each bone. Measurements taken in mm.

Bone Number	Max Length	Bi-epicondylar breadth	Vertical head diameter
1	302	55	35
2	298	57	40
3	298	50	38
4	279	51	36
5	282	53	35
6	288	53	36
7	258	47	33
8	294	54	36
9	280	53	37
10	269	46	33
12	313	64	42
13	332	65	49
14	334	66	48
15	332	64	44
16	320	62.5	46
17	326	60.5	44
18	335	64.5	49
19	316	62.5	43

Bone Number	Max Length	Bi-epicondylar breadth	Vertical head diameter
20	319	58	40
21	275	52	36
22	271	50	35
23	270	50	33
24	272	48	34
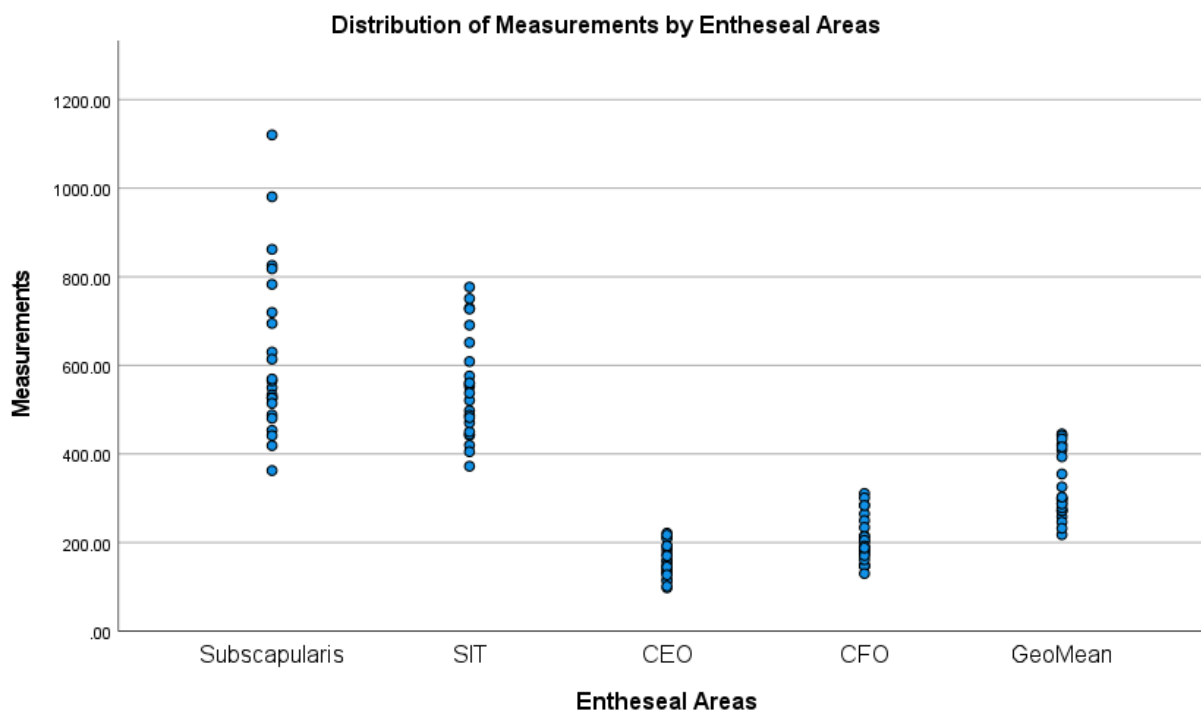
Figure 3.1 The distribution of measurements in relation to each entheseal site and geometric mean. SIT = supraspinatus, infraspinatus, teres minor. CEO = common extensor origin. CFO = common flexor origin. GeoMean = geometric mean.

Error Testing

Testing for inter- and intra- observer error is principle to gauging whether a method is both reliable and reproduceable. To assess inter- and intra- observer error multiple statistical protocols were utilized. Table 3.3-3.6 show the measurements taken for each repetition on each entheseal site. Additionally, it shows the mean between the two repetitions, the mean deviation,

and percent deviation. Most fell within the ≤ 5% error percentage range with a few outliers being

between the 5%-10% range. The highest percent deviation came from bone 5 common flexor

origin, which resulted in a percent deviation of 11.19%.

Table 3.3. Intra-observer measurements for Supraspinatus, Infraspinatus, Teres minor (SIT). Measurements taken in mm^2

Bone Number	Repetition 1	Repetition 2	Mean	Mean Deviation	Percent Deviation
Bone 1 SIT	516.97	525.06	521.02	4.04	0.78
Bone 2 SIT	591.92	625.85	608.89	16.96	2.79
Bone 3 SIT	604.97	547.28	576.13	28.84	5.01
Bone 4 SIT	498.17	497.43	497.80	0.37	0.07
Bone 5 SIT	418.00	421.82	419.91	1.91	0.46
Bone 6 SIT	458.11	426.44	442.28	15.83	3.58
Bone 7 SIT	365.24	379.56	372.40	7.16	1.92
Bone 8 SIT	446.51	441.58	444.04	2.47	0.56
Bone 9 SIT	556.44	548.37	552.41	4.04	0.73
Bone 10 SIT	403.45	406.90	405.18	1.73	0.43
Bone 12 SIT	524.08	593.33	558.71	34.63	6.20
Bone 13 SIT	639.43	663.54	651.48	12.06	1.85
Bone 14 SIT	738.07	720.76	729.42	8.65	1.19
Bone 15 SIT	562.19	558.96	560.57	1.62	0.29
Bone 16 SIT	796.16	757.58	776.87	19.29	2.48
Bone 17 SIT	750.29	704.43	727.36	22.93	3.15
Bone 18 SIT	682.89	698.99	690.94	8.05	1.17
Bone 19 SIT	767.46	734.42	750.94	16.52	2.20
Bone 20 SIT	541.87	533.68	537.78	4.10	0.76
Bone 21 SIT	478.37	497.13	487.75	9.38	1.92
Bone 22 SIT	441.21	459.35	450.28	9.07	2.014
Bone 23 SIT	468.05	472.81	470.43	2.38	0.51
Bone 24 SIT	491.54	472.90	482.22	9.31	1.93

Table 3.4. Intra-observer measurements for Subscapularis (SC). Measurements taken in mm^2

Bone Number	Repetition 1	Repetition 2	Mean	Mean Deviation	Percent Deviation
Bone 1 SC	486.84	490.10	488.47	1.63	0.33
Bone 2 SC	652.84	607.50	630.17	22.67	3.60
Bone 3 SC	526.24	573.07	549.66	23.42	4.26
Bone 4 SC	544.61	522.53	533.57	11.04	2.07
Bone 5 SC	424.92	481.92	453.42	28.50	6.29
Bone 6 SC	506.62	623.56	565.09	58.47	10.35

Bone Number	Repetition 1	Repetition 2	Mean	Mean Deviation	Percent Deviation
Bone 7 SC	352.22	373.08	362.65	10.43	2.88
Bone 8 SC	512.28	541.88	527.08	14.80	2.81
Bone 9 SC	588.23	550.47	569.35	18.88	3.32
Bone 10 SC	411.48	426.49	418.99	7.51	1.79
Bone 12 SC	695.06	744.30	719.68	24.62	3.42
Bone 13 SC	1099.73	1141.18	1120.46	20.72	1.85
Bone 14 SC	844.58	879.68	862.13	17.55	2.04
Bone 15 SC	1014.22	947.43	980.83	33.40	3.41
Bone 16 SC	829.28	822.53	825.91	3.38	0.41
Bone 17 SC	710.38	679.30	694.84	15.54	2.23
Bone 18 SC	793.19	843.17	818.18	24.99	3.05
Bone 19 SC	768.08	798.07	783.07	14.99	1.92
Bone 20 SC	488.47	564.27	526.37	37.90	7.20
Bone 21 SC	625.33	602.92	614.12	11.20	1.82
Bone 22 SC	532.88	496.10	514.49	18.39	3.57
Bone 23 SC	437.86	445.32	441.59	3.73	0.84
Bone 24 SC	504.00	457.35	480.67	23.33	4.85

Table 3.5. Intra-observer measurements for Common Flexor Origin (CFO). Measurements taken in mm^2

Bone Number	Repetition 1	Repetition 2	Mean	Mean Deviation	Percent Deviation
Bone 1 CFO	147.04	149.70	148.37	1.33	0.90
Bone 2 CFO	184.60	221.47	203.03	18.43	9.08
Bone 3 CFO	182.83	175.37	179.10	3.73	2.08
Bone 4 CFO	181.11	182.36	181.74	0.63	0.34
Bone 5 CFO	130.55	163.44	147.00	16.44	11.19
Bone 6 CFO	174.41	171.90	173.15	1.26	0.73
Bone 7 CFO	149.95	144.28	147.12	2.83	1.93
Bone 8 CFO	167.17	155.32	161.24	5.92	3.67
Bone 9 CFO	169.81	171.96	170.88	1.08	0.63
Bone 10 CFO	131.03	129.31	130.17	0.86	0.66
Bone 12 CFO	205.22	225.42	215.32	10.10	4.69
Bone 13 CFO	317.01	304.96	310.99	6.02	1.94
Bone 14 CFO	292.80	276.13	284.47	8.33	2.93
Bone 15 CFO	304.57	297.27	300.92	3.65	1.21
Bone 16 CFO	252.20	277.72	264.96	12.76	4.82
Bone 17 CFO	246.91	252.60	249.75	2.84	1.14
Bone 18 CFO	277.99	289.58	283.79	5.80	2.04
Bone 19 CFO	229.48	238.86	234.17	4.69	2.00
Bone 20 CFO	226.71	200.15	213.43	13.28	6.22
Bone 21 CFO	186.39	190.50	188.45	2.05	1.09
Bone 22 CFO	207.31	202.88	205.09	2.22	1.08
Bone 23 CFO	195.87	188.99	192.43	3.44	1.79
Bone 24 CFO	185.99	190.36	188.17	2.19	1.16

Table 3.6. Intra-observer measurements for Common Extensor Origin (CEO). Measurements taken in mm^2

Bone Number	Repetition 1	Repetition 2	Mean	Mean Deviation	Percent Deviation
Bone 1 CEO	142.71	142.16	142.43	0.28	0.20
Bone 2 CEO	135.04	154.17	144.60	9.57	6.62
Bone 3 CEO	136.58	135.86	136.22	0.36	0.27
Bone 4 CEO	111.30	120.86	116.08	4.78	4.12
Bone 5 CEO	168.83	148.53	158.68	10.15	6.40
Bone 6 CEO	165.55	153.15	159.35	6.20	3.89
Bone 7 CEO	113.52	113.35	113.43	0.09	0.08
Bone 8 CEO	105.62	90.57	98.10	7.52	7.67
Bone 9 CEO	104.65	98.46	101.56	3.10	3.05
Bone 10 CEO	134.68	129.06	131.87	2.81	2.13
Bone 12 CEO	193.51	172.87	183.19	10.32	5.63
Bone 13 CEO	186.15	160.71	173.43	12.72	7.33
Bone 14 CEO	218.56	206.35	212.45	6.11	2.87
Bone 15 CEO	191.97	191.22	191.60	0.38	0.20
Bone 16 CEO	199.36	219.00	209.18	9.82	4.70
Bone 17 CEO	233.82	208.89	221.35	12.47	5.63
Bone 18 CEO	148.09	151.59	149.84	1.75	1.17
Bone 19 CEO	217.37	217.57	217.47	0.10	0.05
Bone 20 CEO	124.44	138.13	131.28	6.85	5.21
Bone 21 CEO	146.03	144.39	145.21	0.82	0.57
Bone 22 CEO	128.88	125.34	127.11	1.77	1.39
Bone 23 CEO	165.87	174.35	170.11	4.24	2.49
Bone 24 CEO	187.59	197.83	192.71	5.12	2.66

Another way that inter- and intra-observer error was assessed was through looking at Lin's concordance correlation coefficient (Lin's CCC). Table 3.7 displays the Lin's CCC value between intra-observer measurements, while Table 3.8 shows Lin's CCC value between inter-observer measurements. Lin's CCC for the measurement repetitions for intra-observer error were all above 0.9, with the lowest being 0.94 and the highest being 0.97. Lin's CCC for the inter-observer measurements were reported by Dr. Karakostis and Dr. Wilczak. Dr. Karakostis provided the inter-observer measurement for the subscapularis while Dr. Wilczak provided it for the remaining areas of interest. Dr. Karakostis reported the Lin's CCC to be above 0.9 for the

subscapularis. Dr. Wilczak reported Lin's CCC to be above 0.9 except for the common extensor

origin which was reported to be 0.87.

Table 3.7 Lin's concordance correlation coefficient for the intra-observer measurements of each
entheseal site.

	Subscapularis	Supraspinatus, Infraspinatus, Teres Minor	Common Extensor Origin	Common Flexor Origin
LinCCC	0.970426	0.972034	0.937435	0.95756

Table 3.8 Lin's concordance correlation coefficient for the inter-observer measurements of each
entheseal site. Only a subset of the sample was reported. N = 7

	Subscapularis	Supraspinatus, Infraspinatus, Teres Minor	Common Extensor Origin	Common Flexor Origin
LinCCC	0.99	0.98	0.87	0.99

To ensure reliability and reproducibility, intraclass correlation coefficient (ICC) and

ANOVA were used. Table 3.9 shows the values for ICC for each entheseal site and Table 3.10

shows the p-values. With ICC all entheseal sites reported numbers back above 0.9 with the

lowest being 0.97. As for the ANOVA tests ran, there were no significant differences found

between the repetitions taken for each individual entheseal site.

Table 3.9. Intraclass Correlation Coefficient of the measurements for each entheseal site.

	Subscapularis	Supraspinatus, Infraspinatus, Teres Minor	Common Extensor Origin	Common Flexor Origin
ICC	0.99	0.99	0.97	0.98

Table 3.10. ANOVA P-values for measurements of each entheseal site.

	Subscapularis	Supraspinatus, Infraspinatus, Teres Minor	Common Extensor Origin	Common Flexor Origin
Sig.	0.24	0.70	0.31	0.47

Correlations

In addition to testing for reliability and reproducibility, this thesis aimed to look at the relationship between entheses of the humerus using the V.E.R.A. method. To evaluate the relationship between entheses of the humerus, as well as its dimensions, correlation tests were performed. Table 3.11 shows the relationship between each of the entheses measured. All four groups of entheses measured showed a moderate to strong positive correlation with one another. Notably there was an especially strong positive correlation between the subscapularis and the common flexor origin which gave a Pearson's r coefficient of .925. In contrast, the lowest r coefficient was between the subscapularis and common extensor origin, resulting in a r value of .571.

Table 3.11. Correlations between entheseal areas raw sizes.

Enthesis	Subscapularis	Supraspinatus, Infraspinatus, Teres minor	Common Extensor Origin	Common Flexor Origin
Subscapularis	—			
Supraspinatus, Infraspinatus, Teres minor	.745	—		
Common Extensor Origin	.571	.672	—	
Common Flexor Origin	.925	.772	.630	—

One relationship looked at was the correlation between entheses and body size. Table 3.12-3.14 displays the correlations between body size proxies and the entheses of interest. Figures 3.2-.3.5 are a visual representation of these correlations in the form of scatterplots. Table 3.12 shows the relationship between bi-epicondylar breadth and the distal entheses. Both the common extensor origin and common flexor origin showed a strong positive correlation, with the common flexor origin having a stronger correlation. Table 3.13 shows the relationship between vertical head diameter and the proximal entheses of the humeral head. The subscapularis and the group of the supraspinatus, infraspinatus, and teres minor showed very strong correlations with r

values of .914 and .871 respectively. Table 3.14 gives the r values between humeral max length

and each entheses. The subscapularis, common flexor origin, and group of supraspinatus,

infraspinatus, and teres minor all showed relatively similar strong correlations. Their r values all

fell above .8, whereas the common extensor origin had a moderate correlation, having the lowest

r value of .568. Table 3.15 presents the p-values for the entheseal measurements and body size

proxies in relation to the geometric mean. Adjusting overall size results in stronger correlations

between measurements with one another as this is reflected in the very high p-values.

Table 3.12. Correlations between Bi-epicondylar breadth and entheses of distal humerus.

	Common Extensor Origin	Common Flexor Origin
Bi-epicondylar breadth	.607	.856

Table 3.13. Correlations between Vertical head diameter and entheses of humeral head.

	Subscapularis	Supraspinatus, Infraspinatus, Teres minor
Vertical head diameter	.914	.871

Table 3.14. Correlation between Max bone length and raw sizes entheseal areas.

	Subscapularis	Supraspinatus, Infraspinatus, Teres minor	Common Extensor Origin	Common Flexor Origin
Max bone length	.839	.816	.568	.836

Table 3.15. P-values showing the correlation between Geometric Mean, raw sizes of entheseal areas, and
bone measurements.

	Subscapularis	Supraspinatus, Infraspinatus, Teres minor	Common Extensor Origin	Common Flexor Origin	Max Length	Bi-epicondylar breadth	vertical head diameter
Geometric mean	.923	.891	.797	.941	.864	.894	.928
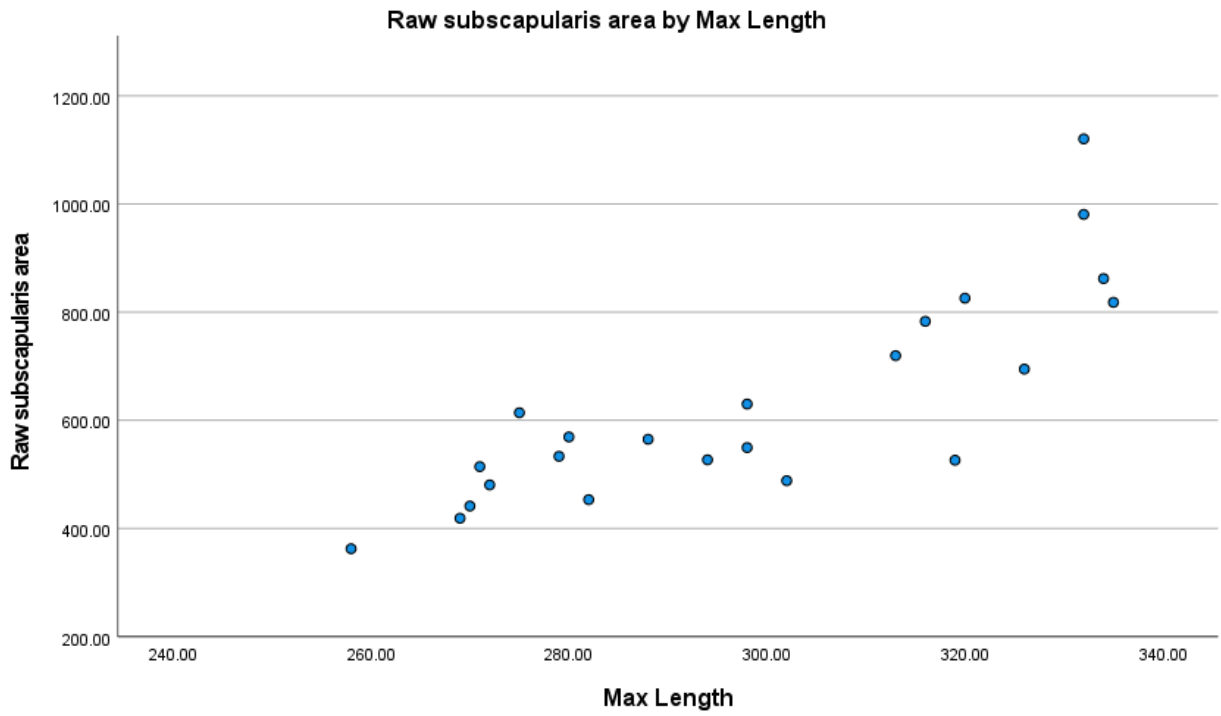
Figure 3.2 Scatterplot showing the relationship of raw subscapularis area measurements by max bone

length.
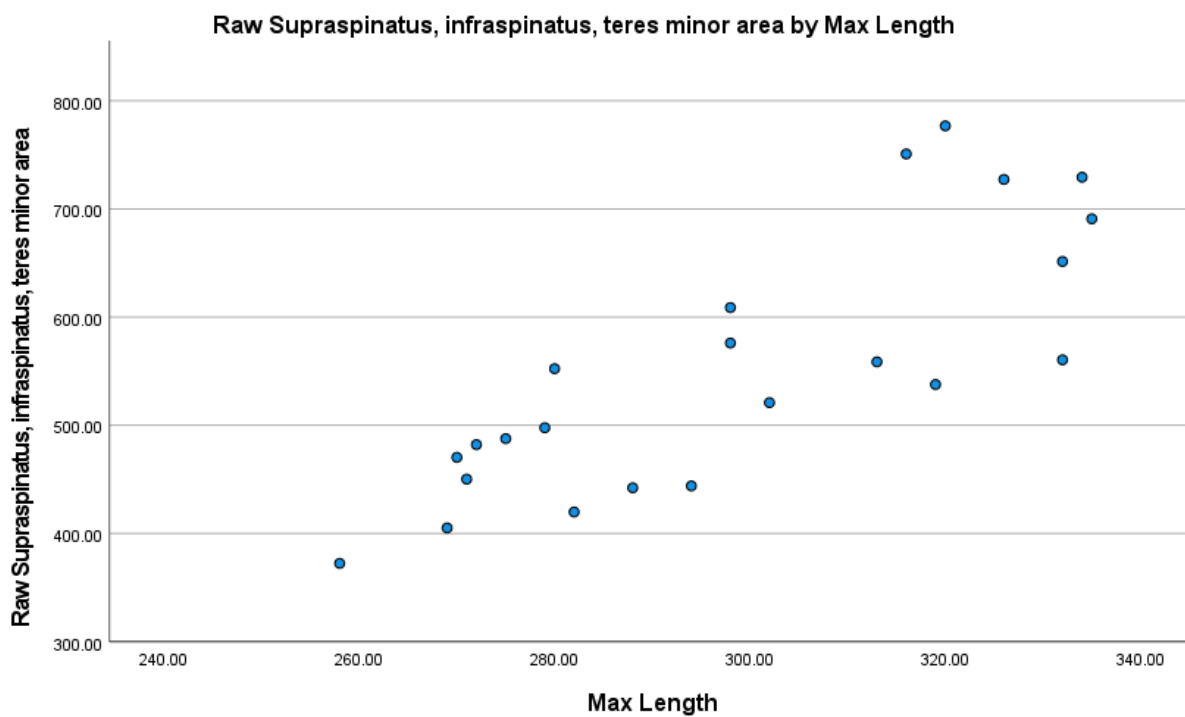
Figure 3.3 Scatterplot showing the relationship of the supraspinatus, infraspinatus, and teres minor with max bone length.
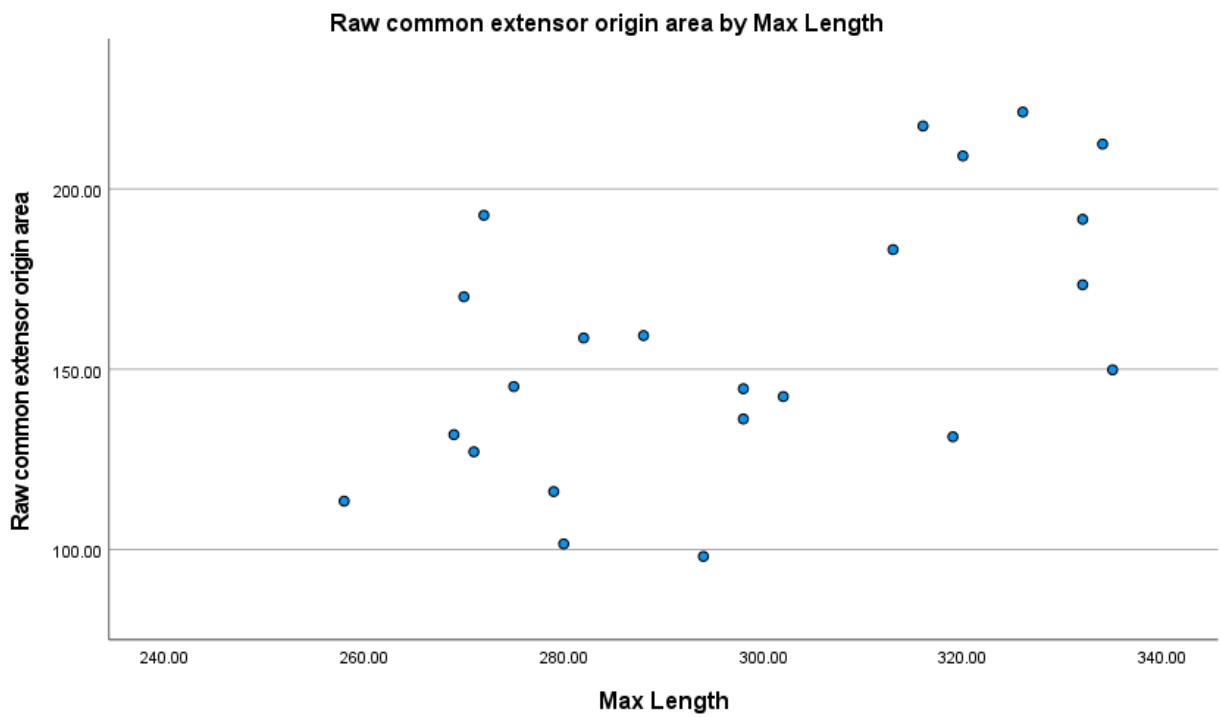
Figure 3.4 Scatterplot showing the relationship between the common extensor origin are measurements and max bone length.
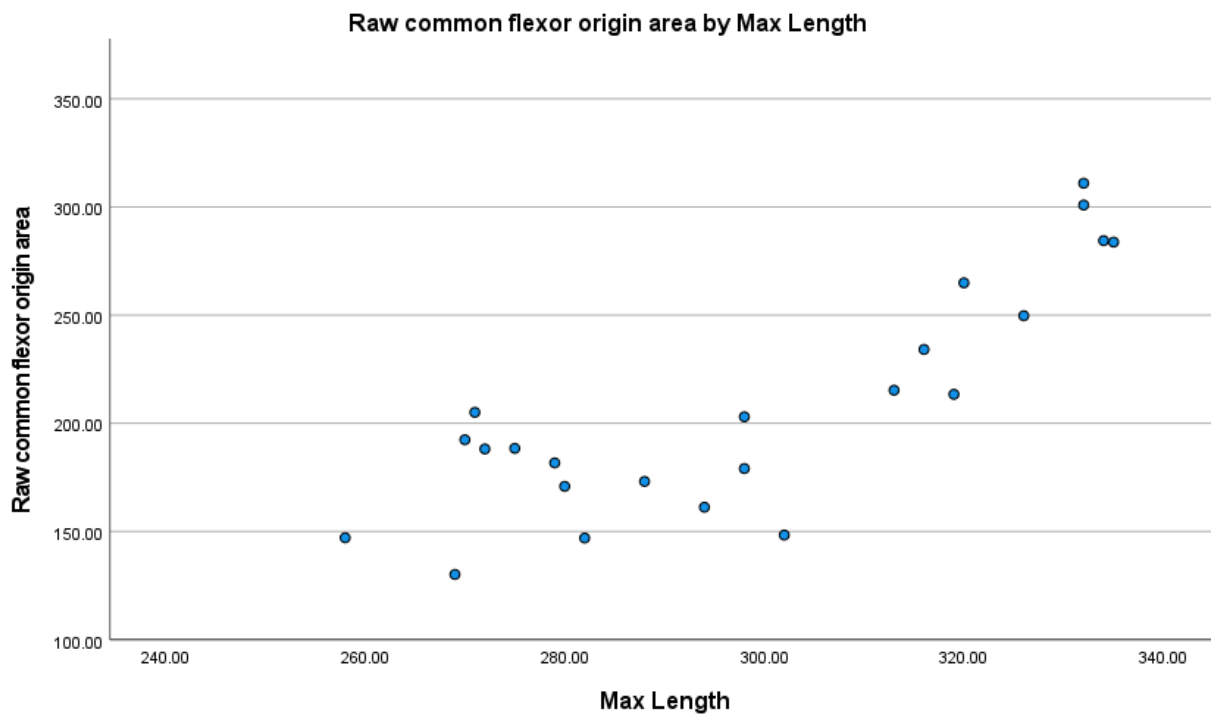
Figure 3.5 Scatter plot showing the relationship of the common flexor origin area measurements with max bone length.

Chapter 4: Discussion

This chapter will discuss and interpret the results of the previous chapter. It will cover the reliability and reproducibility of the V.E.R.A method from the perspective of a first-time user. The following section will then discuss the application of the method and how results agree with previous studies conducted. It will also discuss the benefits and drawbacks of using the V.E.R.A method.

The V.E.R.A method is still new and has not been extensively tested on human remains outside of the hand bones (Karakostis and Harvati 2021). A large part of this thesis was testing the reliability and reproducibility of the V.E.R.A method in relation to enthesis outside of the hand. This was done through extensive error testing. Reliability and reproducibility are essential to any good method. Overall, the results of error testing reported back good numbers that reflect both a reliable and reproducible method. The rule of thumb for a good percent deviation/percent error is $\pm 5\%$. Most percent deviations between the repetitions taken for each bone fell within this range. Outliers can be explained by user inexperience with the method and can be fixed by the inclusion of more repetitions as well as more practice. Additionally, more familiarity with the anatomy of muscle origins and insertions can aid in more accurate delineations.

The assessment of inter and intra observer error also returned good numbers that reflect a reliable and reproducible method. Lin's CCC measures both precision and accuracy. It is interpreted along the same lines as Pearson's correlation coefficient. This means that anything above 0.8 means good while using a stricter upper bound of 0.9 meaning excellent. All Lin's

CCC for the measurement repetitions for intra-observer error were in the excellent range.

Similarly, Lin's CCC for the inter-observer measurements were also all above 0.9.

ICC returned excellent numbers as well. ICC was utilized to measure the reliability and

consistency of measurements taken. More specifically, it aided in evaluating the intra-rater

reliability. Similar to Lin's CCC, ICC is a modified version of Pearson correlation coefficient,

and its values can be interpreted using the 95% confidence intervals (Ko and Li 2016). All ICC

values reported were between 0.97 and 0.99. These numbers reflect very reliable intra-observer

measurements.

These error values demonstrated that measurements had very high agreement with one

another and were consistent between multiple measurements taken by one observer as well as

measurements taken between observers. These results also agree with previous studies that

conducted error testing on the V.E.R.A. method (Karakostis and Lorenzo 2016; Kunze et al.

2020). Although Kunze and colleagues (2020) were able to produce error rates that were all

below 5%. The error numbers from this study can be improved upon, but nonetheless the results

illustrate that the V.E.R.A. method is indeed reliable and reproducible.

The other portion of this thesis looked at the applicability of the V.E.R.A method on

entheses of the humerus. The V.E.R.A method was used to look at the correlation between

entheseal area and body size proxies as well as entheseal areas with one another. As expected,

correlations were found between the entheseal area and body size proxies. All correlation tests

resulted in moderate to strong correlations. This falls in line with previous studies that have

demonstrated a relationship between body size and entheseal area (Niinimaki 2009; Nolte and

Wilczak 2013; Weiss 2004; Weiss et al 2010; Wilczak 1998b). In relation to specific body size

proxies, the strongest correlations were found between the subscapularis along with the grouping of the supraspinatus, infraspinatus, and teres minor with the vertical head diameter. This may be due to the nature of location of these two entheses. This would also explain the similar strong correlations between bi-epicondylar breadth with the common extensor origin and common flexor origin. These entheses are all located directly on or near the place where these measurements are taken, so it would make sense that as the size of these entheses increase so do the corresponding body size proxies. It would be interesting to see how results may change if the supraspinatus, infraspinatus, and teres minor were not grouped together.

Examining correlations between each of the entheseal areas showed fairly strong correlations. The largest correlation was between subscapularis and common flexor origin. This result is quite peculiar as these two muscle groups do not directly interact with one another. It would be expected that there would be a correlation between the two but not as high as a correlation found in this study. This correlation may be due to the fact the common flexor origin had a high percent deviation, which may have thrown some values off. It is not believed that reducing the error rates would change the overall outcome of the correlation results (i.e. resulting in no correlation), but instead may slightly decrease the correlation values for the common flexor origin. If percent deviations were within the proper range, the correlation between the subscapularis and common flexor origin would probably be weaker.

There are still some limitations to the V.E.R.A. method. Similar to other 3D methods like Noldner and Edgar (2013), as well as Nolte and Wilczak (2013), there is still a degree of subjectivity to delineating an enthesis. The user still needs to be able to identify the general area of the enthesis so that the program can calculate where the borders approximately are.

Encompassing too large of an area or too small of an area can result in misrepresenting the enthesis actual size. This emphasizes the importance that the user is well informed on the anatomy of the entheseal area of interest. Additionally, this method does require the user to have knowledge of how to use 3D imaging or have a training session (which is highly recommended) with the developers of the method to be applied reliably. Although this could be circumvented by a publishing of a detailed step by step of how to use the V.E.R.A method, similar to what Karakostis (2022) did with the publication of a step-by-step guide on how to properly apply statistical methods using the V.E.R.A. method. Subjectivity is highly unavoidable and a universal problem that all methodologies face. To avoid any subjectivity in a methodology, it would require the perfect specimen that has all points of interest intact. Although the V.E.R.A. method does not eliminate subjectivity entirely it does reduce some of the subjectivity in comparison to methods that rely on ordinal scoring. Moreover, in comparison to other 3D methods the V.E.R.A. method has significantly smaller error rates. Noldner and Edgar (2013) reported intra-observer error rates between 10% and 15 %. Nolte and Wilczak (2013) reported intra-observer error rates between 12.58% and 16.7%. Results from this study had intra-observer error rates between 1.83% and 3.23%. This reduction in error rates reflects the delineation methods used. Noldner and Edgar (2013), along with Nolte and Wilczak (2013), utilized the chalk outline method to delineate the entheses. The chalk outline method is imprecise due to the subjective nature of drawing a chalk outline. Whereas the V.E.R.A method utilizes computer assisted coloration to detect changes in curvature and elevation to help delineate entheses. Perhaps as more medical studies are conducted on entheseal morphology and anatomy it can help inform researchers in how to delineate the exact borders of entheses.

In a more ideal situation, a skeletal collection with known demographic history would allow for further analysis and application of the V.E.R.A. method. The commingled nature of this skeletal collection limited the lanes of inquiry and analysis that could be taken. Despite these limitations the V.E.R.A method was able to produce meaningful correlations that coincide with previous studies' results. Furthermore, the results of the error testing for the V.E.R.A. method conducted in this thesis have demonstrated that it is both reliable and reproducible.

Chapter 5: Conclusion

This thesis looked to test the reliability and reproducibility of the V.E.R.A. method, and whether its applicability goes beyond hand entheses. The results of this study have demonstrated that the V.E.R.A method is reliable and reproducible. Both results for intra- and inter-observer errors were faily low and highly agreeable with one another. This study also demonstrated that it has the capabilities to be applied to entheses of the humerus. Although this study alone may not be indicative of its applicability outside the hand it does show promising results.

The adoption and application of 3D technologies to analyzing entheses have brought upon a new wave of research in the field of entheseal change. These technologies have allowed researchers to explore complexities of entheses that were not possible before. Furthermore, 3D technologies have helped create 3D methods to research entheses. Methods like the V.E.R.A. method have been demonstrated to be a reliable and repeatable way to analyze them. Further research can expand on the reliability and reproducibility of the V.E.R.A method, and look to apply this method to other parts of the skeleton. The V.E.R.A. method seems very promising and with further research it may be able to become a standardized way of analyzing entheseal change.

Literature Cited

Acosta, M. A., and Eugénia Cunha. 2017. The Effect of Terrain on Entheseal Changes in the Lower Limbs. *International Journal of Osteoarchaeology* 14 343–11. http://doi.wiley.com/10.1002/oa.2597

al-Oumaoui, I., S. Jiménez-Brobeil, and P. du Souich. 2004. Markers of Activity Patterns in Some Populations of the Iberian Peninsula. *International Journal of Osteoarchaeology* 14 (5): 343–59. https://doi.org/10.1002/oa.719.

Cardoso HFV. 2006. Brief communication: The collection of identified human skeletons housed at the Bocage Museum (National Nuseum of Natural History), Lisbon, Portugal. *American Journal of Physical Anthropology* 129:173–176

Angel, J. L., J. O. Kelley, M. Parrington, and S. Pinter. 1987. Life Stresses of the Free Black Community as Represented by the First African Baptist Church, Philadelphia, 1823–1841. *American Journal of Physical Anthropology* 74(2): 213–229. doi: 10.1002/ajpa.1330740209

Apostolakos, J., T. JS Durant, C. R. Dwyer, R. P. Russell, J. H. Weinreb, F. Alaee, K. Beitzel, M. B. McCathy, M. P. Cote, and A. D. Mazzocca. 2014. The enthesis: a review of the tendon-to-bone insertion. *Muscles, Ligaments and Tendons Journal* 4(3):333-342.

Benjamin, M., Evans, E.J., and Copp, L., 1986. The histology of tendon attachments to bone in man. *Journal of Anatomy*. 149:89–100.

Benjamin, M., H. Toumi, J. R. Ralphs, G. Bydder, T. M. Best and S. Milz. 2006. Where Tendons and Ligaments Meet Bone: Attachment Sites ('Entheses') in Relation to Exercise and/or Mechanical Load. *Journal of Anatomy* 208 (4): 471–90.

Benjamin, M., T. Kumai, S. Milz, B. M. Boszczyk, A. A. Boszczyk, and J. R. Ralphs. 2002. The skeletal attachment of tendons - tendon "entheses". *Comparative Biochemistry and Physiology, Part A: Molecular & Integrative Physiology* 133:931-945.

Benjamin, M. and J. R. Ralphs. 1998. Fibrocartilage in tendons and ligaments - an adaptation to compressive load. *Journal of Anatomy* 193:481-494.

Berthon, W., C. Rittermard, B. Tihanyi, G. Pálfi, H. Coqueugniot, and O. Dutour. 2015. Three-dimensional microarchitecture of entheseal changes: preliminary study of human radial tuberosity. *Acta Biologica Szegediensis* 59(1): 79-90.

Davis, C. B., K. A. Shuler, M. E. Danforth, and K. E. Herndon. 2013. Patterns of Interobserver Error in the Scoring of Entheseal Changes. *International Journal of Osteoarchaeology* 23 (2): 147–51. https://doi.org/10.1002/oa.2277.

Eshed, V., A. Gopher, E. Galili, and I. Hershkovitz. 2004. Musculoskeletal Stress Markers in Natufian Hunter-Gatherers and Neolithic Farmers in the Levant: The Upper Limb. *American Journal of Physical Anthropology* 123 (4): 303–15. https://doi.org/10.1002/ajpa.10312.

Godde, K. and R. W. Taylor. 2011. Musculoskeletal Stress Marker (MSM) Differences in the Modern American Upper Limb and Pectoral Girdle in Relation to Activity Level and Body Mass Index (BMI). *Forensic Science International*, 6.

Havelková, P and S. Villotte. 2007. Enthesopathies: test of reproducibility of the new scoring system based on current medical data. *Slovenská antropológia* 10 (1): 51-57.

Hawkey, D. E., and C. F. Merbs. 1995. Activity-Induced Musculoskeletal Stress Markers (MSM) and Subsistence Strategy Changes among Ancient Hudson Bay Eskimos. *International Journal of Osteoarchaeology* 5: 324–38.

Henderson, C. Y., V. Mariotti, F. Santos, S. Villotte, and C. A. Wilczak. 2017. The New Coimbra Method for Recording Entheseal Changes and the Effect of Age–At–death. *Bulletins et Mémoires de la Société d'Anthropologie de Paris* 29 140–49.

Henderson, C. Y., V. Mariotti, D. Pany-Kucera, S. Villotte, and C. Wilczak. 2016. The New 'Coimbra Method': A Biologically Appropriate Method for Recording Specific Features of Fibrocartilaginous Entheseal Changes. *International Journal of Osteoarchaeology* 26 (5): 925–32.

Henderson, C. Y. and E. Nikita. 2016. Accounting for multiple effects and the problem of small sample sizes in osteology: a case study focusing on entheseal changes. *Archaeological and Anthropological Sciences* 8: 805-817.

Henderson, C. Y. and F. Alves Cardoso. 2013. Special Issue Entheseal Changes and Occupation: Technical and Theoretical Advances and Their Applications. *International Journal of Osteoarchaeology* 23: 127-134.

Henderson, C. Y., V. Mariotti, D. Pany-Kucera, S. Villotte, and C. Wilczak. 2013. Recording specific entheseal changes of fibrocartilaginous entheses: initial tests using the Coimbra method. *International Journal of Osteoarchaeology* 23:152-162.

Jurmain, R., F. Alves-Cardoso, C. Henderson, and S. Villotte. 2012. "Bioarchaeology's Holy Grail: The Reconstruction of Activity." In *A Companion to Paleopathology*, edited by Anne

L. Grauer, 531–52. Oxford, UK: Wiley-Blackwell.
https://doi.org/10.1002/9781444345940.ch29.

Jurmain, R.; Villotte, S. 2010. Terminology. Entheses in medical literature and physical anthropology: a brief review [Online]. Document published online in 4th February following the Workshop in Musculoskeletal Stress Markers (MSM): limitations and achievements in the reconstruction of past activity patterns, University of Coimbra, July 2-3, 2009. Coimbra, CIAS - Centro de Investigação em Antropologia e Saúde. [Consulted in 25th June 2010]. Available from: http://www.uc.pt/en/cia/msm/MSM_terminology3.

Karakostis, A. F. and C. Lorenzo. 2016. Morphometric patterns among the 3D surface areas of human hand entheses. *American Journal of Physical Anthropology* 160:694-707.

Karakostis, A. F., G. Hotz, H. Scherf, J. Wahl, and K. Harvati. 2017. Occupational manual activity is reflected on the patterns among hand entheses. *American Journal of Physical Anthropology* 164:30-40.
——2018a. A repeatable geometric morphometric approach to the analysis of hand entheseal three-dimensional form. *American Journal of Physical Anthropology* 166:246-260.

Karakostis, F. A., G. Hotz, V. Tourloukis, and K. Harvati. 2018b. Evidence for precision grasping in Neandertal daily activities. *Science Advances* 4 (9): eaat2369. https://doi.org/10.1126/sciadv.aat2369.

Karakostis, F. A., N. Jeffrey and K. Harvati. 2019. Experimental proof that multivariate patterns among muscle attachments (entheses) can reflect repetitive muscle use. *Scientific Reports* 9:16577. https://doi.org/10.1038/s41598-019-53021-8

Karakostis F. A., I. J. Wallace, N. Konow, and K. Harvati. 2019. Experimental evidence that physical activity affects the multivariate associations among muscle attachments (entheses). *Journal of Experimental Biology* 222: jeb213058. https://doi:10.1242/jeb.213058

Karakostis, F. A., and K. Harvati. 2021. New horizons in reconstructing past human behavior: introducing the "Tübingen University Validated Entheses-based Reconstruction of Activity" method. *Evolutionary Anthropology*, 30: 185–198

Karakostis, F. A. 2022. Statistical protocol for analyzing 3D muscle attachment sites based on the "Validated Entheses-based Reconstruction of Activity" (VERA) approach. *International Journal of Osteoarchaeology*, 1-14. https://doi.org/10.1002/oa.3196

Kelley, J. O. and J. L.Angel. 1987. Life Stresses of Slavery. *American Journal of Physical Anthropology* 74 (2): 199–211.

Kennedy, K. A. R. 1983. Morphological variations in ulnar supinator crests and fossae as identifying markers of occupational stress. *Journal of Forensic Sciences* 28:871-876.

Koo, Terry K. and Li, Mae Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15:155-163.

Kunze, J., Karakostis, F. A., Merker, S., Peresani, M., Tourloukis, V., & Harvati, K. 2022. Evidence for habitual tool use in early hominins. *PaleoAnthropology*, 2022: 195–210. https://doi.org/10.48738/2022. iss2.61

Lai, P., and N. C. Lovell. 1992. Skeletal Markers of Occupational Stress in the Fur Trade: A Case Study from a Hudson's Bay Company Fur Trade Post. *International Journal of Osteoarchaeology* 2 (3): 221–34. https://doi.org/10.1002/oa.1390020306.

Mall, G., M. Hubig, A. Buttner, J. Kuznik, R. Penning, and M. Graw. 2001. Sex determination and estimation of stature from the longbones of the arm. *Forensic Science International* 117: 23-30.

Mariotti, V., F. Facchini, and M. G. Belcastro. 2004. Enthesopathies – proposal of a standardized scoring method and applications. *Collegium Antropologicum* 28 (1): 145-159.
——2007. The Study of Entheses: Proposal of a Standardised Scoring Method for Twenty-Three Entheses of the Postcranial Skeleton. *Collegium Antropologicum* 31(1): 291-313.

Milella, M., F. Alves Cardoso, S. Assis, G. Perréard Lopreno, and N. Speith. 2015. Exploring the Relationship Between Entheseal Changes and Physical Activity: A Multivariate Study. *American Journal of Physical Anthropology* 156 (2): 215–23.

Milella, M., M. G. Belcastro, C. P.E. Zollikofer, and V. Mariotti. 2012. The effect of age, sex, and physical activity on entheseal morphology in a contemporary Italian skeletal collection. *American Journal of Physical Anthropology* 148: 379-388.

Michopoulou, E., E. Nikita, and C. Y. Henderson. 2017. A Test of the Effectiveness of the Coimbra Method in Capturing Activity–Induced Entheseal Changes. *International Journal of Osteoarchaeology* 1–9.

Molnar, P. 2006. Tracing Prehistoric Activities: Musculoskeletal Stress Marker Analysis of a Stone-Age Population on the Island of Gotland in the Baltic sea. *American Journal of Physical Anthropology* 129: 2-23. doi:10.1002/ajpa.2

Molnar, P. 2010. Patterns of physical activity and material culture on Gotland, Sweden, during the Middle Neolithic. *International Journal of Osteoarchaeology* 20 (1): 1–14. https://doi.org/10.1002/oa.1000.

Myszka, A., J. Piontek, and A.Vančata. 2012. Body Mass Reconstruction on the Basis of Selected Skeletal Traits. *Anthropologischer Anzeiger* 69 (3): 305–15. https://doi.org/10.1127/0003-5548/2012/0179.

Nagy, B. 2000. The life left in bones: evidence of habitual activity patterns in two prehistoric Kentucky populations. PhD thesis, Arizona State University.

Niinimäki, S., and L. Baiges Sotos. 2013. The Relationship Between Intensity of Physical Activity and Entheseal Changes on the Lower Limb. *International Journal of Osteoarchaeology* 23 (2): 221–28. https://doi.org/10.1002/oa.2295.

Noldner, L. K., and H. J. H. Edgar. 2013. 3d Representation and Analysis of Enthesis Morphology. *American Journal of Physical Anthropology* 152 (3): 417–24.

Nolte, M. and C. Wilczak. 2013. Three-dimensional surface area of the distal biceps enthesis, relationship to body size, sex, age, and secular changes in a 20th century American sample. *International Journal of Osteoarchaeology* 23: 163-174.

Palmer, J. L. A., K. Quintelier, S. Inskip, and A. L. Waters–rist. 2018. A Comparison of Two Methods for Recording Entheseal Change on a Post–Medieval Urban Skeletal Collection From Aalst (Belgium). *Archaeometry* 27 (S54): 828–15.

Papathanasiou, A., E. Zachou, and M. P. Richards. 2009. Bioarchaeological Analysis of the Human Osteological Material from Proskynas, Lokris. *Hesperia Supplements* 43: 223–35.

Rogers, T. L. 1998. A visual method of determining the sex of skeletal remains using the distal humerus. *Journal of Forensic Sciences* 44(1): 57-60.

Santos, A. L., F. Alves-Cardoso, S. Assis, and S. Villotte. 2011. The Coimbra Workshop in Musculoskeletal Stress Markers (MSM): an annotated review. *Antropologia Portuguesa* 28: 135–61. https://doi.org/10.14195/2182-7982_28_5.

Schlecht, S. H. 2012. Understanding entheses: bridging the gap between clinical and anthropological perspectives. *The Anatomical Record* 296: 1239-1251.

Steen, S. L. and R. W. Lane. 1998. Evaluation of Habitual Activities among Two Alaskan Eskimo Populations Based on Musculoskeletal Stress Markers. *International Journal of Osteoarchaeolgy* 8: 341–53.

Stefanović, S. and M. Porčić. 2013. Between-group differences in the patterning of musculo-skeletal stress markers: avoiding confounding factors by focusing on qualitative aspects of physical activity. *International Journal of Osteoarchaeology* 23:94–105.

Stewart, T. D. 1979. *Essentials of Forensic Anthropology: Especially as developed in the United States*. Springfield, Illinois: Charles C. Thomas.

Villotte, S., S.Assis, F. Alves Cardoso, C. Y. Henderson, V. Mariotti, M. Milella, D. Pany-Kucera, N. Speith, C. A. Wilczak, and R. Jurmain. 2016. In search of consensus: terminology for entheseal changes (EC). *International Journal of Paleopathology* 13: 49-55.

Villotte, S., and C. J. Knüsel. 2013. Understanding Entheseal Changes: Definition and Life Course Changes: Understanding Entheseal Changes. *International Journal of Osteoarchaeology* 23 (2): 135–46. https://doi.org/10.1002/oa.2289.

Villotte, S., D. Castex, V. Couallier, O. Dutour, C.r J. Knüsel, and D. Henry-Gambier. 2010. Enthesopathies as Occupational Stress Markers: Evidence from the Upper Limb. *American Journal of Physical Anthropology* 142: 224–34. https://doi.org/10.1002/ajpa.21217.

Weiss, E., L. Corona, and B. Schultz. 2012. Sex Differences in Musculoskeletal Stress Markers: Problems with Activity Pattern Reconstructions. *International Journal of Osteoarchaeology* 22 (1): 70–80. https://doi.org/10.1002/oa.1183.

Weiss, E.. 2004. Understanding Muscle Markers: Lower Limbs. *American Journal of Physical Anthropology* 125 (3): 232–38. https://doi.org/10.1002/ajpa.10397.
——2010. Cranial muscle markers: a preliminary examination of size, sex, and age effects.
HOMO—Journal of Comparative Human Biology 61:48-58.

Wilczak, C. A. 1998a. Consideration of sexual dimorphism, age, and asymmetry, in quantitative measurements of muscle insertion sites. *International Journal of Osteoarchaeology* 8:311-325.
——1998b. A new method for quantifying musculo skeletal stress markers (MSM): a test of the
relationship between enthesis size and habitual activity in archaeological populations. PhD
Thesis, Cornell University, Ithaca, NY

Wilczak, C. A, V. Mariotti, D. Pany–kucera, S. Villotte, and C. Y. Henderson. 2017. Training and Interobserver Reliability in Qualitative Scoring of Skeletal Samples. *Journal of Archaeological Science: Reports* 11: 69–79. http://dx.doi.org/10.1016/j.jasrep.2016.11.033.

Yonemoto, S. 2016. Difference in the effects of age on the development of entheseal changes among historical Japanese populations. *American Journal of Physical Anthropology* 159: 267-283.

Zumwalt, A. (2005). A New Method for Quantifying the Complexity of Muscle Attachment Sites. *The Anatomical Record* 286B: 21-28.
——2006. The effect of endurance exercise on the morphology of muscle attachment sties. *The
Journal of Experimental Biology* 209: 444-454.

Appendix: Supplementary Data

Table 1. Dr. Wilczak entheses measurements (in mm²)

	CEO	CFO	SIT		CEO	CFO	SIT
Bone 1	130.4299	157.899	520.9019		142.43	148.37	521.9
Bone 7	105.1088	145.065	354.849		113.42	147.12	372.4
Bone 8	105.825	160.5397	422.448		98.1	161.24	444.04
Bone 18	165.383	293.1779	658.277		149.84	283.79	690.94
Bone 22	120.961	220.49	441.0515		127.11	205.09	450.28

Table 2. Dr. Karakostis subscapularis measurements (in mm²)

Bone 1 Subscapularis	**499.42**
Bone 8 Subscapularis	511.81
Bone 9 Subscapularis	567.87
Bone 10 Subscapularis	461.34
Bone 15 Subscapularis	1058.53
Bone 17 Subscapularis	706.24
Bone 21 Subscapularis	647.13

Table 3. Roger (1998) method sex determination

Bone	Trochlear constriction	Trochlear symmetry	Shape and depth of olecranon fossa	Angle of medial epicondyle
1R	M	M	M	M
2R	M	M	M	M
3R	M	M	M	M
4R	M	M	F	M
5R	M	M	M	M
6R	F	M	M	M
7L	M	F	M	M
8L	M	M	M	M
9L	M	F	M	M
10L	M	F	M	F
11L	M	F	M	F

Bone	Trochlear constriction	Trochlear symmetry	Shape and depth of olecranon fossa	Angle of medial epicondyle
12L	M	M	M	M
13L	F	M	M	F
14L	M	F	M	M
15L	M	M	M	F
16L	M	F	F	F
17L	F	M	M	M
18L	M	F	M	F
19R	M	M	M	M
20R	M	M	M	M
21L	M	M	M	F
22L	M	M	F	F
23R	M	M	M	M
24R	M	M	M	M

Table 4. Stewart (1979) and Mall et al. (2001) sex determination

Bone	Stewart (1979)	Mall et al. (2001)
1R	F	F
2R	F	F
3R	F	F
4R	F	F
5R	F	F
6R	F	F
7L	F	F
8L	F	F
9L	F	F
10L	F	F
11L	F	F
12L	F	F
13L	M	M
14L	M	M
15L	I	F
16L	I	F
17L	I	F
18L	M	M
19R	F	F
20R	F	F

Bone	Stewart (1979)	Mall et al. (2001)
21L	F	F
22L	F	F
23R	F	F
24R	F	F