A three-dimensional digital microscopic investigation of entheseal changes as skeletal activity markers

Efthymia Nikita[1], Panagiota Xanthopoulou[1,2], Andreas Bertsatos[3], Maria-Eleni Chovalopoulou[3], Iosif Hafez[1]

[1] Science and Technology in Archaeology and Culture Research Center, The Cyprus Institute, Nicosia, Cyprus
[2] Department of Biology, Aristotle University of Thessaloniki, Greece
[3] Department of Biology, Faculty of Animal and Human Physiology, National and Kapodistrian University at Athens, Greece

Corresponding author contact details:
Efthymia Nikita
Science and Technology in Archaeology and Culture Research Center, The Cyprus Institute, Nicosia, Cyprus
tel. +357 22 208 669
e.nikita@cyi.ac.cy

Abstract

Objectives
The current paper explores the effectiveness of entheseal changes as skeletal activity markers by testing the correlation between such changes and cross-sectional geometric (CSG) properties while controlling for the effect of age and body size.

Materials and Methods
The originality of the paper lies in capturing entheseal changes in a continuous quantitative manner using 3D microscopy. Roughness and bone resorption were recorded on zone 1 and 2 of three humeral entheses (subscapularis, supraspinatus, infraspinatus) in a documented sample of 29 male skeletons.

Results
Our analysis found that merely 5.91% of the partial correlations between entheseal changes and CSG properties were statistically significant. In addition, two unexpected patterns were identified, namely a higher number of significant correlations on the left side entheses compared to the right side ones and a higher number of correlations between minimum roughness and CSG properties compared to mean and maximum roughness.

Discussion
These patterns are the inverse of what we would expect if activity had exerted an important effect on entheseal change expression. Therefore, they support the lack of association between entheseal changes and habitual activity, even though various factors potentially affecting the above results are discussed.

Keywords: activity markers; entheses; digital microscopy

Introduction

The assessment of activity patterns based on skeletal evidence has been such a pressing issue that it has received the characterisation 'Bioarchaeology's Holy Grail' (Jurmain, Alves Cardoso, Henderson, & Villotte, 2011). Osteoarthritis, long-bone diaphyseal cross-sectional geometric properties and entheseal changes are among the most frequently encountered methods in bioarchaeological studies for addressing issues of past habitual activities (Jurmain, 1999; Molnar, Ahlstrom, & Leden, 2011; Palmer, Hoogland, & Waters-Rist, 2016; Ruff, 2008).

Among these, entheseal changes (EC) have received great attention over the past decade regarding their actual potential to act as skeletal activity markers with this issue largely remaining unresolved. Clinical and biomechanical data support that bone reacts to mechanical stress by increasing blood flow in the affected areas and subsequently there is elevated bone growth, thus more pronounced EC among individuals exposed to increased levels of mechanical stress (e.g. Lieberman, Pearson, Polk, Demes, & Crompton, 2003; Parfitt, 2004; Petit et al., 2004; Woo et al., 1981). Based on this principle, a number of bioarchaeological studies have adopted EC as a means of reconstructing the activity patterns of past populations (e.g. Eshed, Gopher, Galili, & Hershkovitz, 2004; Hawkey & Merbs, 1995; Lieverse, Bazaliiskii, Goriunova, & Weber, 2009; Molnar, 2008; Schrader, 2012; Stefanović & Porčić, 2013; Villotte et al., 2010; Weiss, 2007).

Despite the broad use of EC as activity markers in the bioarchaeological literature, studies conducted largely over the past decade have highlighted the multifactorial aetiology of EC with age, body size, sex, metabolic, genetic, pathological and other factors exerting important influence on their expression (e.g. Colao, Ferone, Marzullo, & Lombardi, 2004; Villotte et al., 2010; Wilczak, 1998). In this direction, several studies have suggested that the primary factor affecting EC is age (Alves Cardoso & Henderson, 2010; Benjamin, Toumi, Suzuki, Hayashi, & McGonagle 2009; Milella, Belcastro, Zollikofer, & Mariotti, 2012; Niinimäki, 2011; Weiss, 2007; Wilczak, 1998). Even though it could be argued that the increased EC with age may be the outcome of cumulative activity effects, for fibrous entheses, age-related changes have been attributed to reduced rates of bone formation resulting in thinner cortical bone with rough external surfaces (Chapman, 1997; Mays, 2000; Robb, 1998; Wilczak, 1998), while for fibrocartilaginous entheses, to increasing tendon stiffness (Jurmain et al., 2011). Another factor that affects EC expression and may be more prominent than activity is body size, with larger individuals exhibiting more pronounced EC (Niinimäki, 2011; Weiss, Corona, & Schultz, 2012). Finally, with respect to the impact of sex, various studies have found that males exhibit higher levels of EC (Milella et al., 2012; Villotte et al., 2010; Wilczak, 1998). This pattern may reflect sexual division of labor, differences in body size (al-Oumaoui, Jiménez-Brobeil, & du Souich, 2004; Molnar, 2006), or hormonal differences (Mariotti, Facchini, & Belcastro, 2007; Niinimäki, 2011; Wilczak, 1998).

A renewed interest in EC as potential activity markers has emerged upon the acknowledgement that entheses are anatomically distinguished in fibrocartilaginous and fibrous (Benjamin et al., 2002). This acknowledgement has led to the support that fibrocartilaginous entheses are better

indicators of activity compared to fibrous ones, particularly before the age of 50 years when the effect of age is not predominant (Villotte et al., 2010). The better performance of fibrocartilaginous entheses as activity markers has been attributed to the fact that loadings are more evenly dispersed across the bone surface in fibrous entheses, thus their effect is dissipated (Benjamin et al., 2002; Zumwalt, 2006).

An issue that has received increasing attention in the past years relates to the recording schemes for EC. Traditionally, both fibrous and fibrocartilaginous EC were recorded using the Hawkey and Merbs (1995) method, which focuses on new bone formation and bone resorption. More recently, Villotte et al. (2010) recommended a simpler presence/absence recording scheme, applicable exclusively to fibrocartilaginous entheses. Even more recently, the (new) Coimbra method (Henderson, Mariotti, Pany-Kucera, Villotte, & Wilczak, 2013, 2016) suggested the division of each fibrocartilaginous enthesis into two zones and the recording of multiple variables that capture bone formation and resorption in each zone. Michopoulou and colleagues (Michopoulou, Nikita, & Valakos, 2015; Michopoulou, Nikita, & Henderson, 2017) tested the association between upper limb fibrocartilaginous EC and cross-sectional geometric properties in a modern collection with documented age, sex and occupation and found that current recording schemes for EC cannot identify a consistent association between these markers and activity (as attested through cross-sectional geometric properties).

Other recent studies have adopted continuous quantitative recording methods. Simple approaches involve the use of sliding calipers for measuring the size of entheses and profile gauges for quantifying their shape (Henderson 2013). More advanced methods employ three-dimensional scanning technologies (Karakostis & Lorenzo, 2016; Karakostis, Hotz, Scherf, Wahl, & Harvati, 2017, 2018; Noldner & Edgar, 2013; Nolte & Wilczak, 2013). A limitation of the three-dimensional quantitative studies is that, with the exception of the Karakostis et al. (2018) paper, they focus only on the surface area of the entheses, failing to account for shape differences. Two notable zooarchaeological studies in the direction of quantifying entheseal changes are those by Zumwalt (2005), who combined 3D laser scanning with fractal analysis in female sheep, and Wallace et al. (2017), who adopted microcomputed tomography and morphological topographic analysis in female turkeys. With regard to human skeletal studies, even in the Karakostis et al. (2018) paper, which represents an important development in the direction of employing geometric morphometric methods in EC studies, the quantification approach does not effectively take into account prior morphological methods for EC scoring, in the sense that it does not divide the (fibrocartilaginous) entheses in zones and it does not account for bone resorption such as micro- and macro-porosity or for variables such as textural change.

The current paper aims at complementing previous studies by adopting a 3D microscopic approach, focused on quantifying bone formation and resorption in the different zones of humeral fibrocartilaginous entheses. The focus of this paper is to test whether EC captured using 3D digital microscopy exhibit a significant correlation with cross-sectional geometric properties, while controlling for the effect of age and body size in a modern male sample.

Materials and Methods

Materials

The current study included 29 male individuals from the University of Athens Human Skeletal Reference Collection or in short, The Athens Collection. This collection is curated at the Department of Animal and Human Physiology, National and Kapodistrian University of Athens, Greece. The collection comprises the skeletal remains of individuals who lived primarily in the second half of the twentieth century and originate from cemeteries in the area of Athens. The age, sex, occupation, and cause of death is known for most of them (Eliopoulos, Lagia, Manolis, 2007).

The 29 male skeletons were selected on the basis of representing in a balanced manner different adult age groups and within each age group, individuals of different stature, thus different body size. These 29 individuals had diverse documented occupations/professions, such as plumber, dentist, driver, electrician, translator, military personnel, painter, farmer, sailor, bank clerk, and chef, among others. This suggests that our sample included individuals with different levels of physical activity; however, it is unclear how much actual physical labour is involved in each profession, as well as for how long the individuals practiced the reported professions (see also Cardoso and Henderson, 2013 for bias caused by the categorisation of occupation in skeletal collections). For this reason, we avoided using 'occupation' as a variable in our analyses and we opted instead for cross-sectional geometric properties as an activity proxy (see following sections).

Individuals with insufficient age and sex documentation and individuals with pathological lesions or post-mortem damage that could inhibit the correct recording of entheseal changes were excluded. Regarding pathological lesions, individuals suffering from seronegative spondyloarthropathy, DISH, or acromegaly were also excluded since these conditions have been found to influence entheseal morphology (e.g., Henderson, 2008). Table 1 presents the composition of the study sample per age and related descriptive statistics.

Casting method

The humeral entheses under examination included the subscapularis, supraspinatus and infraspinatus. These fibrocartilaginous entheses were selected because, as mentioned in the Introduction, it has been suggested that fibrocartilaginous entheses exhibit a higher correlation with activity than fibrous entheses (Villotte et al. 2010). In addition, for these three entheses the zones, as defined in the (new) Coimbra method (Henderson et al. 2013, 2016) are easy to identify. For the 29 individuals included in the sample, the right and left elements were examined separately.

The 3D digital microscopic analysis was performed using a Hirox KH 8700 digital microscope, housed at the Science and Technology in Archaeology and Culture Research Center of The

Cyprus Institute in Nicosia. Due to restrictions in the portability of the microscope, high-resolution casts of the entheses were created and used in all analysis. Moulds were made using a polyvinylsiloxane impression material (Coltène President Jet Light Body). Casts of the entheses were created by setting the silicone moulds in putty (Provil Novo Putty) and pouring resin and hardener (Araldite 2020) into them. The resin and hardener were placed in a vacuum chamber for ten minutes in order to remove any microscopic bubbles prior to pouring into the silicone moulds. Subsequently, the casts were left overnight in a high pressure chamber (~3.4 atm) so that the resin would penetrate any porosity and adhere more firmly on the mould's surface, and also to minimize the size of any remaining bubbles that might interfere with the analysis. Creating casts with polyvinylsiloxane impression material and resin and hardener has been a standard approach in dental microwear analysis and has been found to reproduce microscopic features with a resolution of a fraction of a micron (Goodall, Darras, & Purnell, 2015). More details on the casting method and photographic documentation of the different steps can be found in Supplement 1.

Microscopic analysis

Since EC are expressed in two main forms, new bone formation and bone resorption, we collected data on the roughness of the surface of each enthesis as well as estimated the sum of all bone resorption. Supplement 1 provides step by step instructions on the capturing of porosity and roughness using the Hirox KH 8700 digital microscope.

For the recording of bone resorption, the casts were viewed under ×35 magnification and the surface area of each pore or area of bone resorption (cavitation, erosion etc.) was measured using built-in measuring tools of the Hirox microscope software (circle and free area selection) (Figure 1). Subsequently, all areas of bone resorption were added separately for zone 1 and zone 2 per enthesis (see Henderson et al., 2016 for definition of these zones). This mode of collecting data for bone resorption has the limitation that it does not differentiate between different types of resorption (e.g. microporosity, macroporosity, cavitation etc.), instead it pools all types together. However, in the current paper we are interested in exploring the potential effect of activity on EC and using broader categories of variables may lead to clearer patterns in this direction, rather than using multiple variables in an already rather small sample of skeletons.

To measure new bone formation, the casts were viewed under ×50 magnification and the roughness tool of the built-in Hirox software was employed. This tool is based on the three-dimensional recreation of the surface of each enthesis and the subsequent vertical 'cut' of this surface by artificial planes placed at selected points/chords (Figure 2). The arithmetic mean roughness (Ra) was then automatically calculated by removing the standard length from the roughness curve in the direction of the mean line, totalling absolute values of deviations between the removed mean line and the measurement curve, and averaging them. In this approach, the roughness values depend on the cut-off wavelength applied, which separates roughness from waviness before calculating roughness. As a cut-off value between roughness

and waviness we used 1/5 of the length of the chord along which roughness was measured, following Wieland (1999) and Whitehouse (1994). Roughness was measured at three areas on zone 1 and two areas on zone 2 for each enthesis. If we visualise zone 1 as an arch, the first point was on one end, the second in the middle and the third at the other end. For zone 2 the two areas were chosen so that they are representative of the overall roughness in this zone; that is, if part of the enthesis exhibited particularly marked roughness/new bone formation at a certain area but was otherwise smooth, we obtained one measurement at the rough area and another measurement at the smooth part. In each area, five separate measurements in 5mm chords were obtained and averaged so that roughness is captured more accurately, given that if we shift the chord a little bit to the right or left, the roughness measurement will change, and we wanted to account for such non-directional error. Note that we could not use chords larger than 5mm because of limits in the field of view under magnification, while we opted for the ×50 magnification because the greater the magnification, the more accurate the roughness score is. Subsequently, three different measures of roughness were used per zone: the mean roughness score (average of the three scores for zone 1 and the two scores for zone 2), maximum roughness score (the highest value of the three scores for zone 1 and the two scores for zone 2) and minimum roughness score (the smallest value of the three scores for zone 1 and the two scores for zone 2). Areas of bone resorption (porosity, cavitation or other) were avoided during the measurement of roughness since they had already been captured in the bone resorption measurements, while the purpose of the roughness variable was primarily to capture new bone formation. We acknowledge that roughness is often a combination of bone formation and resorption. For this reason, even though this variable was used principally as a proxy for new bone formation and, as stated above, we made sure to avoid areas of clear bone resorption while measuring it, it will be treated and discussed throughout the paper as a variable that captures a combination of bone formation and resorption.

Cross-sectional geometric properties

Table 2 presents the cross-sectional geometric (CSG) properties used in the current study and their definition according to Ruff (2008). The CSG properties of the humeral samples were calculated with the "long-bone-diaphyseal-CSG-Toolkit" (Bertsatos, 2018), which is freely available at https://github.com/pr0m1th3as/long-bone-diaphyseal-CSG-Toolkit. The CSG Toolkit works directly on triangular mesh 3D models, which were created from the original bone samples used in the present study by 3D photogrammetry pipeline utilizing Photoscan Pro v1.4 (Agisoft LLC, Russia) software. The CSG Toolkit, the correct and reliable operation of which has been extensively validated (Bertsatos & Chovalopoulou, 2018), automatically optimizes the orientation of the long bone in order to virtually slice the diaphyseal shaft at 20%, 35%, 50%, 65% and 80% along the maximum bone length. Subsequently, a number of CSG properties are returned for each individual cross-section, such as those given in Table 2. The CSG Toolkit is an alternative to the latex-cast method for calculating cross-sectional geometric properties (O'Neill & Ruff, 2004), which minimizes observer error and digitization noise and its accuracy is only limited by the quality of the produced 3D models. The rated accuracy of the 3D models was estimated at ±0.12mm with respect to the bone's maximum length for the

present sample. Descriptive statistics for the CSG properties of the sample used in the current study are given in Table 3, where it is clear that this sample exhibited considerable variation in the CSG values, which aligns with the variation in documented occupations (see Materials section).

Statistical analysis

The primary aim of the statistical analysis was to measure the strength of the relationship between EC (roughness and bone resorption) and CSG properties (as a proxy for activity) while controlling for the effect of age and body size. Tests of normality (Shapiro-Wilk test) showed that many of our variables violated the normality assumption. For this reason, we could not apply Generalised Linear Models, as in Michopoulou et al. (2015, 2017) and instead we used partial Spearman correlations to test the correlation between roughness/bone resorption in each zone per enthesis and each CSG property in each section of the humeral diaphysis, while age and body size were used as control variables. To capture body size, all analyses were run separately with stature and body mass as covariates. Note that stature and body mass were calculated using the regression equations by Nikita and Chovalopoulou (2017), which have been produced using the Athens Collection, because relevant information was not available in the medical records of the individuals of this collection. Based on the literature for Pearson's correlation coefficients (Campbell, 2006), r-values between 0.40 and 0.60 indicate moderate correlation, whereas r-values over 0.60 demonstrate strong correlation. Cast accuracy and repeatability were tested using intra-class correlation coefficient. The ICC models of ICC (3, k) and ICC (3, 1) were computed using the ICC function of the psych library in R.

Results

Analysis of cast accuracy

Prior to any data analysis, we tested how accurate the casts were compared to the original bones. For this purpose, roughness and bone resorption were recorded on two bones and on their respective casts and the values obtained were compared. These variables were recorded and compared separately in zone 1 and zone 2 of the three entheses under study. Note that the same elements were used in the repeatability analysis (see below), thus data on the bones versus the casts were recorded on five separate occasions and all measurements were used in testing cast accuracy. In Table 4, it can be seen that the ICC values are above 0.7, especially the values of ICC(3, k) are always above 0.8, suggesting that intra-class correlation is excellent (Lo et al., 2017).

Analysis of repeatability

To test the precision of the methodology proposed, a repeatability analysis was conducted. The same bones and casts used in testing cast accuracy (see above) had roughness and bone resorption recorded on five separate occasions, each two weeks apart from the previous. The results are given in Table 4. It is seen that the ICC values are always above 0.7, while the value of ICC (3, k) in particular is always above 0.9, suggesting again an excellent intra-class correlation.

Partial correlation results

A total of 2,400 partial correlations were run in order to capture the association among every type of EC (mean roughness, maximum roughness, minimum roughness, bone resorption) with every CSG property (TA, I_x, I_y, I_{max}, I_{min}) across the five segments in which each diaphysis was divided, separately for the right and left side limbs, while controlling for the effect of age and body size, using stature and body mass independently as body size proxies.

Table 5 presents the number of statistically significant correlations. It is seen that out of the 2,400 tests, merely 142 (5.91%) were statistically significant at $p = 0.05$, whereas 379 (15.8%) were statistically significant at $p = 0.10$. Using body mass as a proxy for body size results in a small increase in the number of significant partial correlations between EC and CSG properties compared to using stature (6.33% at $p = 0.05$ and 17.75% at $p = 0.10$ versus 5.5% at $p = 0.05$ and 13.83% at $p = 0.10$), while left-side elements exhibited more significant partial correlations compared to right-side ones (9% at $p = 0.05$ and 20.41% at $p = 0.10$ versus 2.83% at $p = 0.05$ and 11.17% at $p = 0.10$).

Tables S1 and S2 present all the statistically significant correlations (at $p = 0.05$) along with Spearman's rho and the associated p-values in order to see which specific EC and CSG properties were significantly correlated. It is seen that EC in zone 1 show a much higher number of significant correlations with CSG properties compared to EC in zone 2. In particular, when stature is used as a proxy for body size, EC in zone 1 show a significant correlation with CSG properties in 58 cases (87.88% of the significant comparisons) and in zone 2 in merely 8 cases (12.12%), while when body mass is used as a proxy for body size, the corresponding numbers are 61 (80.26%) and 15 (19.74%). Examining which specific EC showed a more systematic correlation with CSG properties, minimum roughness was the variable with the highest number of significant correlations (42 comparisons -63.63%- when stature is the proxy for size and 50 comparisons -65.78%- when body mass is the body size proxy), followed by mean roughness (11 comparisons -16.67%- for stature and 13 comparisons -17.11%- for body mass), maximum roughness (7 comparisons -10.61%- for stature and 7 comparisons -9.21%- for body mass) and, finally, bone resorption (6 comparisons -9.10%- for stature and 6 comparisons -7.89%- for body mass). All cross-sectional geometric properties showed a comparable number of significant correlations with entheseal changes, ranging from 11 (TA) to 17 (I_x) when stature is used as a body size proxy and from 12 (TA) to 19 (I_{min}) when body mass is the body size

proxy. Finally, the Spearman rho absolute values for the statistically significant results suggested a small to moderate correlation, ranging from 0.32 to 0.554 when stature is used as a body size proxy and from 0.321 to 0.529 when body mass is used as a body size proxy. The data that support the findings of this study are available from the corresponding author upon reasonable request.

Discussion

As pointed out in the Introduction, EC have been used extensively as activity markers in bioarchaeological studies employing skeletal remains with a broad temporal and geographic distribution; however, they have also been criticised on the ground that other factors, most notably age and body size, affect their expression more than activity. In an attempt to enhance the potential of EC as activity markers, different recording schemes have been proposed that capture their expression in a more or less detailed manner (Hawkey & Merbs, 1995; Henderson et al., 2013, 2016; Villotte et al., 2010). With the generalised use of geometric morphometrics, 3D surface data have also been employed in recent EC studies. Noldner and Edgar (2013) compared 3D against ordinal and 2D surface EC data in an assemblage from the Pottery Mound site in New Mexico. Even though the archaeological nature of the assemblage could not allow any conclusive results regarding the potential of 3D recorded entheses as activity markers, the authors highlighted the need to promote the use of 3D scanning technology for quantifying enthesis development. On the other hand, Nolte and Wilczak (2013) showed that upper limb entheseal 3D surface area is primarily dependent upon body size, followed by age and secular change in the Terry Collection. In a series of papers, Karakostis and colleagues focused on the quantification of hand entheses. In particular, Karakostis and Lorenzo (2016) proposed a methodology for the recording of 3D areas of hand entheses and used the index of relative entheseal size as a means of exploring some of the factors affecting hand entheses development. Karakostis et al. (2017) used the same methodology and identified similar patterns in enteheseal surface among individuals with similar levels of activity (intense manual labor vs. less strenuous and/or highly mechanized occupations). More recently, Karakostis et al. (2018) proposed a geometric morphometric approach that allows the quantification of the three-dimensional shape of hand entheses, overcoming the limitation of the previous two papers, which captured exclusively entheseal size. A limitation of existing quantitative studies in human assemblages is that they focus on the surface area of the entheses rather than EC shape. As stated in the Introduction, the only exception to this pattern is the Karakostis et al. (2018) paper, but even in this case the quantification approach adopted did not divide the (fibrocartilaginous) enthesis in zones, neither did it consider changes such as bone resorption or textural change.

The current paper explored the potential of EC as activity markers by testing their partial correlation with different CSG properties while controlling for the effect of age and body size, focused on the proximal humerus (subscapularis, supraspinatus and infraspinatus). It follows the same principles as the Michopoulou et al. (2015, 2017) papers but this time EC are quantified in a continuous manner using a 3D microscopic approach. Our results found a very

small number (5.91%) of statistically significant partial correlations between bone resorption and roughness on the one hand and CSG properties on the other hand while controlling for age and body size (proxied by stature and body mass). Some unexpected patterns emerged, namely a higher number of significant partial correlations in left-side elements compared to right-side ones as well as the fact that minimum roughness was the EC that showed the highest number of significant correlations with CSG properties, followed by mean roughness, maximum roughness and, finally, bone resorption. In case activity was an important factor in the expression of the EC, at least as recorded in the current paper, we would have expected the reverse pattern, that is, a more systematic correlation between EC and CSG properties in the right-side humeri, which is usually the dominant side (Steele, 2000), as well as a higher correlation between maximum roughness and CSG properties, followed by mean roughness. In light of our overall findings, these unexpected results lend further support to the fact that entheseal changes do not effectively express activity patterns, at least based on the variables used to capture these changes and activity patterns in our sample.

Before dismissing the use of EC as skeletal activity markers, we need to acknowledge a number of potential limitations to our study. First of all, we have only focused on three entheseal sites, all of which are located in the proximal humerus. We selected these sites because it is easy to identify Zones 1 and 2 (as defined by the New Coimbra method), they have been adopted in many relevant studies (e.g. Schrader, 2015; Thomas, 2014), and it was easy to capture them together during the casting procedure. Nonetheless, it is important to test more EC before drawing any firm conclusions. In addition, our analyses should be repeated with even larger and more diverse samples before the results can be generalised. At this point it must be stressed that, as a general rule in statistical analysis, the lack of a statistically significant effect does not necessarily mean that the effect is absent, but that no such effect was traced by the model (Stockburger, 2007), and with small sample sizes it becomes harder to find a significant effect. We should, however, stress that even if the correlation between EC and CSG properties becomes significant by adopting larger samples, the small Spearman's rho values traced in the current study are unlikely to change considerably. We should also highlight once again that despite the fact that we examined 29 individuals, these exhibited considerable variation in their documented professions as well as in the estimated cross-sectional geometric properties; therefore, we cannot attribute the poor performance of EC as activity markets to a lack of variability in labour history in our assemblage. As stated in the Methods section, even though roughness and bone resorption as captured in our study express many of the variables used in existing ordinal and binary EC recording schemes, they do not express a one-to-one correspondence with them. Finally, the CSG properties adopted in this paper capture the rigidity of skeletal elements against bending and compressive forces. Compared to previous studies (Michopoulou et al., 2015, 2017), we have now calculated the CSG properties not only at 35% distance from the distal epiphysis but at five sections along the bone shaft in order to capture more effectively bone response to mechanical stress. Nonetheless, as highlighted in Michopoulou et al. (2017), different types of mechanical stress, such as high peak strain or shear strain, may have been more important in producing EC. Additionally, the ontogeny of CSG and EC responses may be different and this issue becomes even more prominent given the likelihood that profession/activity would change during an individual's lifetime.

Despite the above potential limitations, our results are in accordance to those of previous studies that have found a small association between EC and CSG properties (Michopoulou et al., 2015, 2017), as well as EC and activity patterns more generally when other contributing factors such as age and body size are taken into account (e.g., Alves Cardoso & Henderson, 2010; Benjamin et al., 2009; Lieverse et al., 2009; Milella et al., 2012; Niinimaki, 2011; Weiss, 2007). This small association cannot be attributed to the mode of recording EC, that is, binary or ordinal variables, as the current paper adopted a continuous approach. Thus, bearing in mind the possible restrictions highlighted above, the findings of the current paper suggest an inherent limitation of EC in expressing activity.

Acknowledgements

Literature cited

al-Oumaoui, I., Jiménez-Brobeil, S., & du Souich, P. (2004). Markers of activity patterns in some populations of the Iberian Peninsula. *International Journal of Osteoarchaeology*, **14**, 343–359. DOI: 10.1002/oa.719

Alves Cardoso, F.A., & Henderson, C.Y. (2010). Enthesopathy formation in the humerus: data from known age-at-death and known occupation skeletal collections. *American Journal of Physical Anthropology*, **141**, 550–560. DOI: 10.1002/ajpa.21171

Benjamin, M., Kumai, T., Milz, S., Boszczyk, B.M., Boszczyk, A.A., & Ralphs, J.R. (2002). The skeletal attachment of tendons – tendon "entheses". *Comparative Biochemistry and Physiology A*, **133**, 931–945. DOI: 10.1016/S1095-6433(02)00138-1

Benjamin, M., Toumi, H., Suzuki, D., Hayashi, K., & McGonagle, D. (2009). Evidence for a distinctive pattern of bone formation in enthesophytes. *Annals of the Rheumatic Diseases*, **68**, 1003–1010. DOI: 10.1136/ard.2008.091074

Bertsatos, A. (2018). Long-bone-diaphyseal-CSG-Toolkit: A GNU Octave CSG Toolkit (Version v1.0.1). *Zenodo*. DOI:10.5281/zenodo.1467045

Bertsatos, A., & Chovalopoulou, M.-E. (2018). Validation study of correct operation for the "long-bone-diaphyseal-CSG-Toolkit". *Zenodo*. DOI: 10.5281/zenodo.1466135

Campbell, M.J. (2006). *Statistics at Square Two* (2nd ed.). London: Blackwell BMJ Books.

Cardoso, F.A., Henderson, C. (2013). The categorisation of occupation in identified skeletal collections: A source of bias? *International Journal of Osteoarchaeology*, **23**, 186-196. DOI: 10.1002/oa.2285

Chapman, M. (1997). Evidence for Spanish influence on activity induced musculoskeletal stress markers at Pecos Pueblo. *International Journal of Osteoarchaeology*, **7**, 497–506. DOI: 10.1002/(SICI)1099-1212(199709/10)7:5<497::AID-OA394>3.0.CO;2-H

Colao, A., Ferone, D., Marzullo, P., & Lombardi, G. (2004). Systemic complications of acromegaly: epidemiology, pathogenesis, and management. *Endocrine Reviews*, **25**,102–152. DOI: 10.1210/er.2002-0022

Eliopoulos, C., Lagia, A., & Manolis, S.K. (2007). A modern, documented human skeletal collection from Greece. *Homo*, **58**, 221–228. DOI: 10.1016/j.jchb.2006.10.003

Eshed, V., Gopher, A., Galili, E., & Hershkovitz, I. (2004). Musculoskeletal stress markers in Natufian hunter-gatherers and Neolithic farmers in the Levant: the upper limb. *American Journal of Physical Anthropology*, **123**, 303–315. DOI: 10.1002/ajpa.10312

Goodall, R.H., Darras, L.P., & Purnell, M.A. (2015). Accuracy and precision of silicon based impression media for quantitative areal texture analysis. *Scientific Reports*, **5**, 10800. DOI: 10.1038/srep10800

Hawkey, D.E., & Merbs, C.F. (1995). Activity-induced musculoskeletal stress markers (MSM) and subsistence strategy changes among ancient Hudson Bay Eskimos. *International Journal of Osteoarchaeology*, **5**, 324–338. DOI: 10.1002/oa.1390050403

Henderson, C.Y. (2013). Technical note: Quantifying size and shape of entheses. *Anthropological Science*, **121**, 63-73. DOI: 10.1537/ase.121017

Henderson, C.Y. (2008). When hard work is disease: the interpretation of enthesopathies. In: M. Brickley & M. Smith (Eds.), *Proceedings of the 8th Annual Conference of the British Association of Biological Anthropology and Osteoarchaeology* (pp. 17-25). Oxford: Archaeopress.

Henderson, C.Y., Mariotti, V., Pany-Kucera, D., Villotte, S., & Wilczak, C. (2013). Recording specific entheseal changes of fibrocartilaginous entheses: Initial tests using the Coimbra Method. *International Journal of Osteoarchaeology*, **23**, 152–162. DOI: 10.1002/oa.2287

Henderson, C.Y., Mariotti, V., Pany-Kucera, D., Villotte, S., & Wilczak, C. (2016). The new 'Coimbra Method': A biologically appropriate method for recording specific features of fibrocartilaginous entheseal changes. *International Journal of Osteoarchaeology*, **26**, 925-932. DOI: 10.1002/oa.2477

Jurmain, R. (1999). *Stories from the Skeleton: Behavioral Reconstruction in Human Osteology*. London: Routledge.

Jurmain, R., Alves Cardoso, F., Henderson, C., & Villotte, S. (2011). Bioarchaeology's Holy Grail: the reconstruction of activity. In: A.L. Grauer (Ed.), *A Companion to Paleopathology* (pp. 531-552). New York: Wiley-Blackwell.

Karakostis, F.A., & Lorenzo, C. (2016). Morphometric patterns among the 3D surface areas of human hand entheses. *American Journal of Physical Anthropology*, **160**, 694–707. DOI: 10.1002/ajpa.22999

Karakostis, F.A., Hotz, G., Scherf, H., Wahl, J., & Harvati, K. (2017). Occupational manual activity is reflected on the patterns among hand entheses. *American Journal of Physical Anthropology*, **164**, 30–40. DOI: 10.1002/ajpa.23253

Karakostis, F.A., Hotz, G., Scherf, H., Wahl, J., & Harvati, K. (2018). A repeatable geometric morphometric approach to the analysis of hand entheseal three-dimensional form. *American Journal of Physical Anthropology*, **166**, 246-260. DOI: 10.1002/ajpa.23421

Lieberman, D.E., Pearson, O.M., Polk, J.D., Demes, B., & Crompton, A.W. (2003). Optimization of bone growth and remodeling in response to loading in tapered mammalian limbs. *The Journal of Experimental Biology*, **206**, 3125–3138. DOI: 10.1242/jeb.00514

Lieverse, A.R., Bazaliiskii, V.I., Goriunova, O.I., & Weber, A.W. (2009). Upper limb musculoskeletal stress markers among middle Holocene foragers of Siberia's Cis-Baikal region. *American Journal of Physical Anthropology*, **138**, 458–472. DOI: 10.1002/ajpa.20964

Lo, W.L.A., Zhao, J.L., Chen, L., Lei, D., Huang, D.F., Tong, K.F. (2017). Between-days intra-rater reliability with a hand held myotonometer to quantify muscle tone in the acute stroke population. *Scientific Reports*, **7**, 14173. DOI:10.1038/s41598-017-14107-3

Mariotti, V., Facchini, F., & Belcastro, M.G. (2007). The study of entheses: proposal of a standardised scoring method for twenty-three entheses of the postcranial skeleton. *Collegium Antropologicum*, 31, 191–313.

Mays, S. (2000). Age-dependent cortical bone loss in women from 18th and early 19th century London. *American Journal of Physical Anthropology*, **112**, 349–361. DOI: 10.1002/1096-8644(200007)112:3<349::AID-AJPA6>3.0.CO;2-0

Michopoulou, E., Nikita, E., & Valakos, E.D. (2015). Evaluating the efficiency of different recording protocols for entheseal changes in regards to expressing activity patterns using archival data and cross-sectional geometric properties. *American Journal of Physical Anthropology*, **158**, 557-568. DOI: 10.1002/ajpa.22822

Michopoulou, E., Nikita, E., & Henderson, C.Y. (2017). A test of the effectiveness of the Coimbra method in capturing activity-induced entheseal changes. *International Journal of Osteoarchaeology*, **27**, 409-417. DOI: 10.1002/oa.2564

Milella, M., Belcastro, M.G., Zollikofer, C.P.E., & Mariotti, V. (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. DOI: 10.1002/ajpa.22060

Molnar, P. (2008). Dental wear and oral pathology: possible evidence and consequences of habitual use of teeth in a Swedish Neolithic sample. *American Journal of Physical Anthropology*, **136**, 423–431. DOI: 10.1002/ajpa.20824

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**,12–23. DOI: 10.1002/ajpa.20234

Molnar, P., Ahlstrom, T.P., & Leden, I. (2011). Osteoarthritis and activity—an analysis of the relationship between eburnation, musculoskeletal stress markers (MSM) and age in two Neolithic hunter–gatherer populations from Gotland, Sweden. *International Journal of Osteoarchaeology*, **21**, 283-291. DOI: 10.1002/oa.1131

Niinimäki, S. (2011). What do muscle marker ruggedness scores actually tell us? *International Journal of Osteoarchaeology*, **21**, 292–299. DOI: 10.1002/oa.1134

Nikita, E., Chovalopoulou, M.E. (2017). Regression equations for the estimation of stature and body mass using a Greek documented skeletal collection. *Homo*, **68**, 422-432. DOI: 10.1016/j.jchb.2017.11.002

Noldner, L.K., & Edgar, H.J. (2013). 3D representation and analysis of enthesis morphology. *American Journal of Physical Anthropology*, **152**, 417-424. DOI: 10.1002/ajpa.22367

Nolte, M., & Wilczak, C. (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. DOI: 10.1002/oa.2292

O'Neill, M.C., & Ruff, C.B. (2004). Estimating human long bone cross-sectional geometric properties: a comparison of noninvasive methods. *Journal of Human Evolution*, **47**, 221–235. DOI: 10.1016/j.jhevol.2004.07.002

Palmer, J.L., Hoogland, M.H., & Waters-Rist, A.L. (2016). Activity reconstruction of Post-Medieval Dutch rural villagers from upper limb osteoarthritis and entheseal changes. *International Journal of Osteoarchaeology*, **26**, 78-92. DOI: 10.1002/oa.2397

Parfitt, A.M. (2004). The attainment of peak bone mass: what is the relationship between muscle growth and bone growth? *Bone*, **34**, 767–770. DOI: 10.1016/j.bone.2004.01.023

Petit, M.A., Beck, T.J., Lin, H.-M., Bentley, C., Legro, R.S., & Lloyd, T. (2004). Femoral bone structural geometry adapts to mechanical loading and is influenced by sex steroids: the Penn State young women's health study. *Bone*, **35**, 750–759. DOI: 10.1016/j.bone.2004.05.008

Robb, J.E. (1998). The interpretation of skeletal muscle sites: a statistical approach. *International Journal of Osteoarchaeology*, **8**, 363–377. DOI: 10.1002/(SICI)1099-1212(1998090)8:5<363::AID-OA438>3.0.CO;2-K

Ruff, C.B. (2008). Biomechanical analyses of archaeological human skeletons. In M.A. Katzenberg & S.R. Saunders (Eds.), *Biological Anthropology of the Human Skeleton* (2nd ed.) (pp. 183-206). New York: Wiley-Blackwell.

Schrader, S.A. (2012). Activity patterns in New Kingdom Nubia: an examination of entheseal remodeling and osteoarthritis at Tombos. *American Journal of Physical Anthropology*, **149**, 60–70. DOI: 10.1002/ajpa.22094

Schrader, S.A. (2015). Elucidating inequality in Nubia: An examination of entheseal changes at Kerma (Sudan). *American Journal of Physical Anthropology*, **156**, 192-202. DOI: 10.1002/ajpa.22637

Steele, J. (2000). Handedness in past human populations: skeletal markers. *Laterality*, **5**, 193–220. DOI: 10.1080/713754380

Stefanović, S., & Porčić, M. (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. DOI: 10.1002/oa.1243

Stockburger, D.W. (2007). Hypothesis and hypothesis testing. In N.J. Salkind (Ed.), *Encyclopedia of Measurement and Statistics* (pp. 446-450). Thousand Oaks, CA: SAGE Publications, Inc.

Thomas, A. (2014). Bioarchaeology of the middle Neolithic: Evidence for archery among early european farmers. *American Journal of Physical Anthropology*, **154**, 279-290. DOI: 10.1002/ajpa.22504

Villotte, S., Castex, D., Couallier, V., Dutour, O., Knüsel, C.J., & Henry-Gambier, D. (2010). Enthesopathies as occupational stress markers: evidence from the upper limb. *American Journal of Physical Anthropology*, **142**, 224–234. DOI: 10.1002/ajpa.21217

Wallace, I.J., Winchester, J.M., Su, A., Boyer, D.M., Konow, N. (2017). Physical activity alters limb bone structure but not entheseal morphology. *Journal of Human Evolution*, **107**, 14-18. DOI: 10.1016/j.jhevol.2017.02.001

Weiss, E., & Jurmain, R. (2007). Osteoarthritis revisited: a contemporary review of aetiology. *International Journal of Osteoarchaeology*, **17**, 437-450. DOI: 10.1002/oa.889

Weiss, E. (2006). Osteoarthritis and body mass. *Journal of Archaeological Science*, **33**, 690-695. DOI: 10.1016/j.jas.2005.10.003

Weiss, E. (2007). Muscle markers revisited: activity pattern reconstruction with controls in a central California Amerind population. *American Journal of Physical Anthropology*, **133**, 931–940. DOI: 10.1002/ajpa.20607

Weiss, E., Corona, L., & Schultz, B. (2012). Sex differences in musculoskeletal stress markers: Problems with activity pattern reconstructions. *International Journal of Osteoarchaeology*, **22**, 70–80. DOI: 10.1002/oa.1183

Whitehouse, D.J. (1994) *Handbook of Surface Metrology* (1st Ed.). London: Taylor & Francis.

Wieland, M., Hanggi, P., Hotz, W., Textor, M., Keller, B.A., & Spencer, N.D. (2000). Wavelength-dependent measurement and evaluation of surface topographies: application of a new concept of window roughness and surface transfer function. *Wear*, **237**, 231-252. DOI: 10.1016/S0043-1648(99)00347-6

Wilczak, C.A. (1998). Consideration of sexual dimorphism, age, and asymmetry in quantitative measurements of muscle insertion sites. *International Journal of Osteoarchaeology*, **8**, 311–325. DOI: 10.1002/(SICI)1099-1212(1998090)8:5<311::AID-OA443>3.0.CO;2-E

Woo, S.L., Kuei, S.C., Amiel, D., Gomez, M.A., Hayes, W.C., White, F.C., & Akeson, W.H. (1981). The effect of prolonged physical training on the properties of long bone: a study of Wolff 's Law. *The Journal of Bone and Joint Surgery, American volume*, **63**, 780–786.

Zumwalt, A.C. (2005). A new method for quantifying the complexity of muscle attachment sites. *The Anatomical Record (Part B: New Anat.)*, **286B**, 21-28. DOI: 10.1002/ar.b.20075

Zumwalt, A.C. (2006). The effect of endurance exercise on the morphology of muscle attachment sites. *Journal of Experimental Biology*, **209**, 444–454. DOI: 10.1242/jeb.02028

TABLES

Table 1. Sample composition

Age group	Age range	No of individuals	Mean age	Standard deviation
Young adult	19-39	10	29	5.8
Middle adult	40-59	9	50.4	6.5
Old adult	60-89	10	75.5	11.5

Table 2. Cross-sectional geometric properties estimated in the current study
(from Ruff, 2008)

Property	Definition
Total subperiosteal area (TA)	Area within subperiosteal surface (resistance to tension, compression, or shear)
Second moment of area about M-L (x) axis (I_x)	Anterior-posterior bending rigidity
Second moment of area about A-P (y) axis (I_y)	Medial-lateral bending rigidity
Maximum second moment of area (I_{max})	Maximum bending rigidity
Minimum second moment of area (I_{min})	Minimum bending rigidity

Table 3. Descriptive statistics for the humeral cross-sectional geometric properties obtained at the midshaft.

	TA	I_x	I_y	I_{max}	I_{min}
	R side				
Mean	326.4	9101.5	8806.2	10649.9	7257.7
Standard deviation	52.8	2978.0	2958.9	3514.4	2392.3
Minimum	230.7	4745.9	3955.7	5284.6	3416.9
Maximum	435.6	15588.6	15974.6	18875.7	12757.6
	L side				
Mean	316.2	8537.9	8155.6	9950.1	6743.4
Standard deviation	43.8	2437.9	2385.6	2829.3	1934.4
Minimum	244.2	4978.9	4005.1	5256.0	3867.4
Maximum	396.8	13400.0	12969.4	16028.3	10582.5

Table 4. Intra-class correlation and associated p-values

Element/ Comparison	Enthesis	Variable	ICC (3,1)	p-value	ICC (3,k)	p-value
Bone 1	subscapularis	roughness	0.763	<0.001	0.942	<0.001
	supraspinatus		0.814	<0.001	0.956	<0.001
	infraspinatus		0.777	<0.001	0.946	<0.001
	subscapularis	porosity	0.994	<0.001	0.999	<0.001
	supraspinatus		0.995	<0.001	0.999	<0.001
	infraspinatus		0.940	<0.001	0.987	<0.001
Bone 2	subscapularis	roughness	0.722	<0.001	0.929	<0.001
	supraspinatus		0.728	<0.001	0.931	<0.001
	infraspinatus		0.728	<0.001	0.930	<0.001
	subscapularis	porosity	0.998	<0.001	0.999	<0.001
	supraspinatus		0.992	<0.001	0.998	<0.001
	infraspinatus		0.948	<0.001	0.989	<0.001
Resin 1	subscapularis	roughness	0.731	<0.001	0.931	<0.001
	supraspinatus		0.879	<0.001	0.973	<0.001
	infraspinatus		0.761	<0.001	0.941	<0.001
	subscapularis	porosity	0.991	<0.001	0.998	<0.001
	supraspinatus		0.995	<0.001	0.999	<0.001
	infraspinatus		0.934	0.001	0.986	0.001
Resin 2	subscapularis	roughness	0.815	<0.001	0.957	<0.001
	supraspinatus		0.798	<0.001	0.952	<0.001
	infraspinatus		0.708	<0.001	0.924	<0.001
	subscapularis	porosity	0.999	<0.001	0.999	<0.001
	supraspinatus		0.979	<0.001	0.996	<0.001
	infraspinatus		0.974	<0.001	0.995	<0.001
Bone 1 vs. cast 1	subscapularis	roughness	0.732	<0.001	0.845	<0.001
	supraspinatus		0.720	<0.001	0.837	<0.001
	infraspinatus		0.689	<0.001	0.816	<0.001
	subscapularis	porosity	0.995	<0.001	0.998	<0.001
	supraspinatus		0.991	<0.001	0.995	<0.001
	infraspinatus		0.922	<0.001	0.959	<0.001
Bone 2 vs. cast 2	subscapularis	roughness	0.737	<0.001	0.848	<0.001
	supraspinatus		0.662	<0.001	0.797	<0.001
	infraspinatus		0.799	<0.001	0.888	<0.001
	subscapularis	porosity	0.999	<0.001	0.999	<0.001
	supraspinatus		0.986	<0.001	0.993	<0.001
	infraspinatus		0.938	<0.001	0.968	<0.001

Table 5. Number of statistically significant partial correlations
(No of significant correlations at p=0.05 / No of significant correlations at p=0.10)

| Side	Body size proxy	Type of enth_eseal change				Total
		Mean roughness	Maximum roughness	Minimum roughness	Bone resorption	
Right	Stature	0/0 out of 150	2/22 out of 150	0/9 out of 150	3/7 out of 150	5/38 out of 600
	Body mass	6/21 out of 150	4/26 out of 150	15/36	4/13 out of 150	29/96 out of 600
Left	Stature	11/39 out of 150	5/23 out of 150	42/60 out of 150	3/6 out of 150	61/128 out of 600
	Body mass	7/34 out of 150	3/19 out of 150	35/61 out of 150	2/3 out of 150	47/117 out of 600

Figure legends
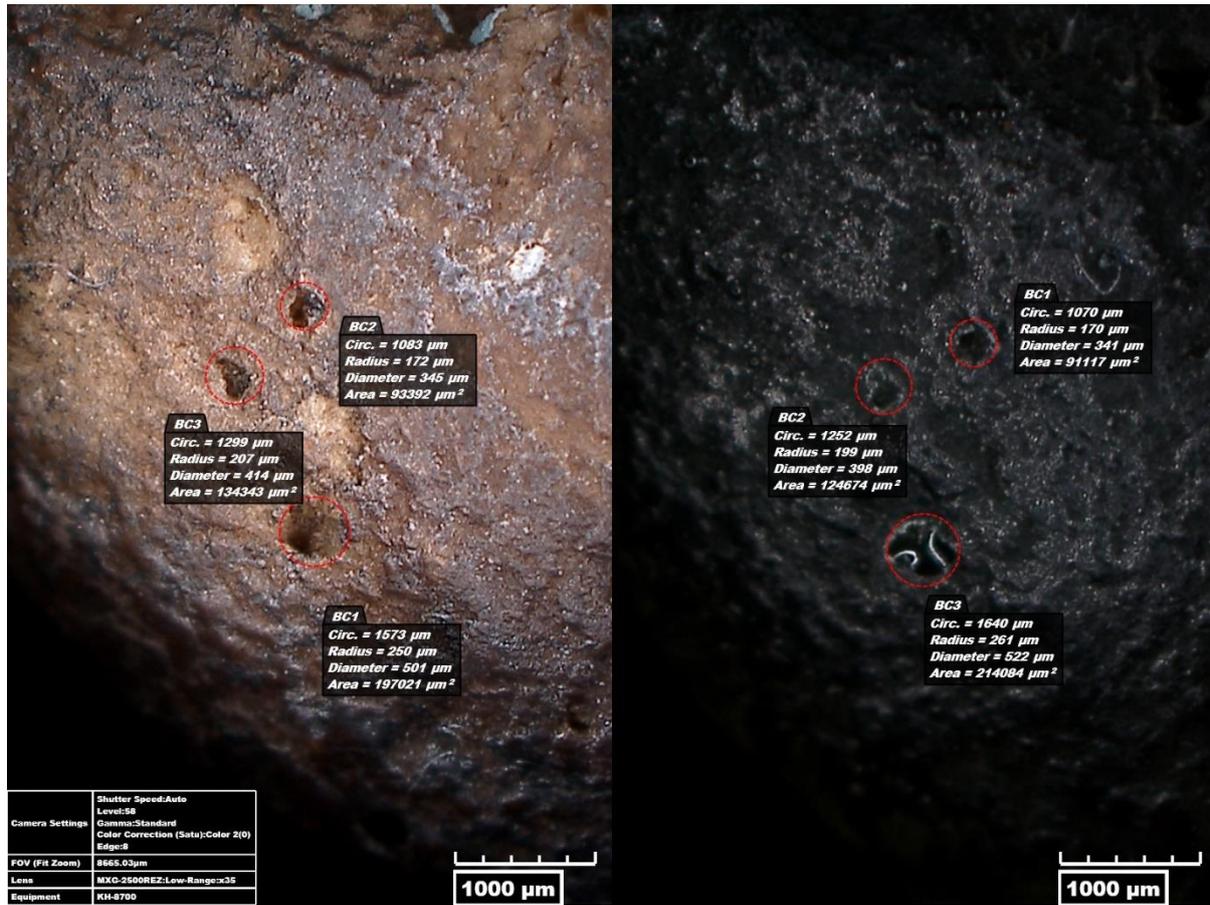
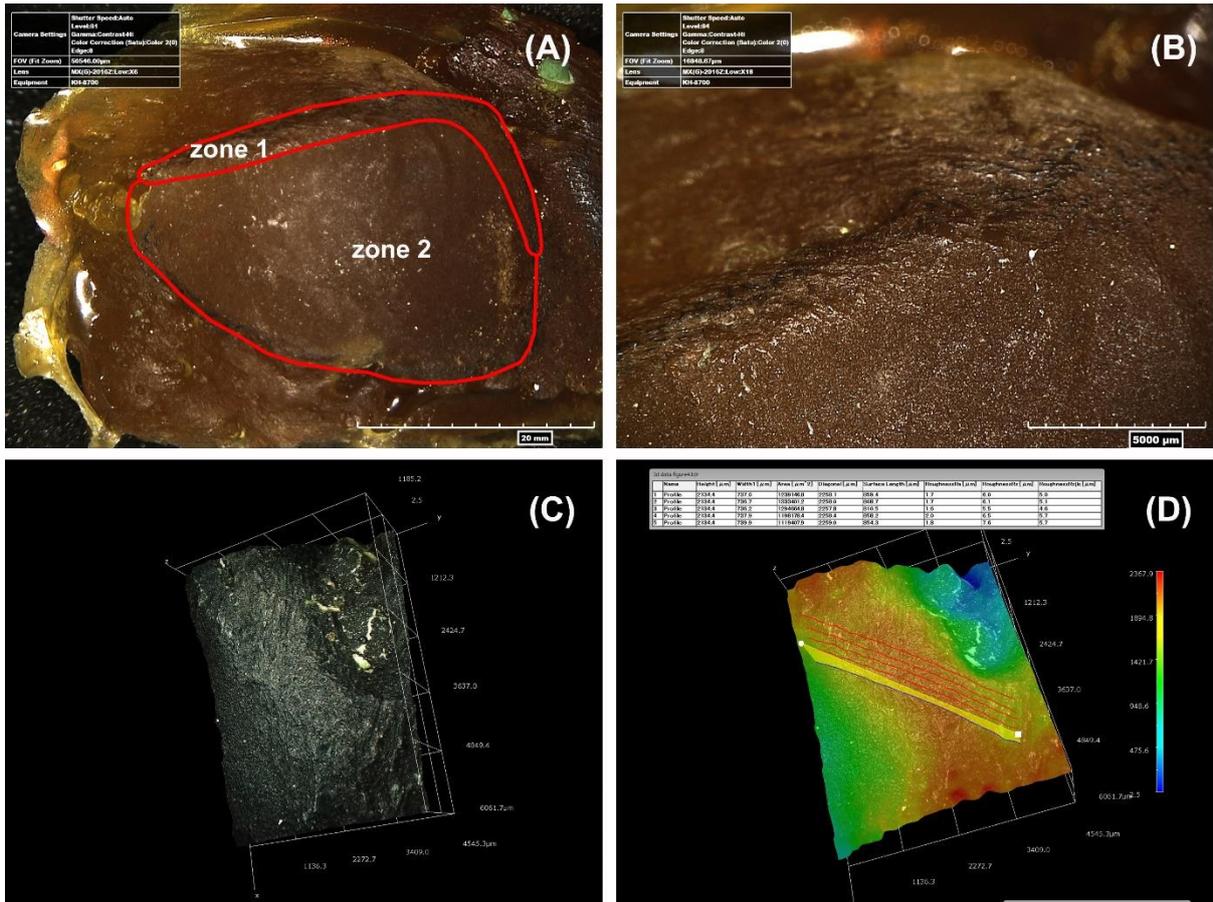Figure 1. Porosity measurement on bone (left) and corresponding cast (right) of test sample 1 at magnification (x35).

Figure 2. Steps in the measurement of roughness. M. subscapularis in male individual ABH133 (A); middle section of zone 1 at magnification (x18) (B); 3D image of the area where roughness was measured at even higher magnification (x50) (C); same image as before but with different colors showing surface elevation; the table shows the roughness measurement per chord (D)