



**Εθνικό και Καποδιστριακό
Πανεπιστήμιο Αθηνών
Τμήμα Βιολογίας
Τομέας Φυσιολογίας Ζώων και Ανθρώπου**

**Τριδιάστατη μορφομετρική ανάλυση του σκελετού του
σύγχρονου ελληνικού πληθυσμού με φωτογραμμετρικές
τεχνικές**

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

**Ανδρέας Μπερτσάτος
ΗΛ. ΜΗΧΑΝΙΚΟΣ**

ΑΘΗΝΑ 2021

«Η έγκριση της Διδακτορικής Διατριβής από το Τμήμα Βιολογίας της Σχολής Θετικών
Επιστημών του ΕΚΠΑ δεν υποδηλώνει αποδοχή των απόψεων του συγγραφέα»

Τριμελής Συμβουλευτική Επιτροπή

- Ε. Βαλάκος, Καθηγητής, Πανεπιστήμιο Αθηνών (Επιβλέπων)
- Χ. Παπαγεωργοπούλου, Αν. Καθηγήτρια, Πανεπιστήμιο Θράκης
- Ι. Λυριτζής, Καθηγητής, Πανεπιστήμιο Αιγαίου

Επταμελής Εξεταστική Επιτροπή

- Ε. Βαλάκος, Καθηγητής, Πανεπιστήμιο Αθηνών (Επιβλέπων)
- Χ. Παπαγεωργοπούλου, Αν. Καθηγήτρια, Πανεπιστήμιο Θράκης
- Ι. Λυριτζής, Καθηγητής, Πανεπιστήμιο Αιγαίου
- Ε. Νικήτα, Αν. Καθηγήτρια, Ινστιτούτο Κύπρου
- Α. Κακάμπουρα, Καθηγήτρια, Πανεπιστήμιο Αθηνών
- Π. Παφίλης, Αν. Καθηγητής, Πανεπιστήμιο Αθηνών
- Δ. Χατζηνικολάου, Αν. Καθηγητής, Πανεπιστήμιο Αθηνών

Η παρούσα διατριβή είναι αφιερωμένη στη σύντροφό μου,

Μαριλένα

Περίληψη

Σκοπός της παρούσας διδακτορικής έρευνας ήταν η διεύρυνση της γνώσης πάνω στην ποικιλομορφία του σκελετού του σύγχρονου ελληνικού πληθυσμού και η ανάπτυξη νέων μεθοδολογιών και εργαλείων για την εξαγωγή βιολογικών παραμέτρων, όπως το φύλο και η ηλικία θανάτου, από ανθρώπινα σκελετικά κατάλοιπα. Το σύνολο της έρευνας βασίστηκε στη σύγχρονη σκελετική συλλογή αναφοράς, που φυλάσσεται στο εργαστήριο Φυσικής Ανθρωπολογίας του Τομέα Φυσιολογίας Ζώων και Ανθρώπου του Τμήματος Βιολογίας, η οποία αποτελεί την πληρέστερη σκελετική συλλογή αναφοράς στην Ελλάδα και θεωρείται αντιπροσωπευτικό δείγμα του ελληνικού πληθυσμού του 20ού αιώνα. Βασικός στόχος του διδακτορικού ήταν η ψηφιακή τεκμηρίωση ενός υποσυνόλου του διαθέσιμου σκελετικού δείγματος στο οποίο εστίασαν τα επί μέρους ερευνητικά ερωτήματα. Η ψηφιακή τεκμηρίωση, που βασίστηκε στην τριδιάστατη φωτογραμμετρία, αφορά το κρανίο, την κάτω γνάθο, τα ανώνυμα οστά, τα μηριαία, τις κνήμες και τα βραχιόνια από όλα τα άτομα της σκελετικής συλλογής, όπου ήταν διαθέσιμα. Η ολοκλήρωση της ψηφιακής τεκμηρίωσης έδωσε 2117 τριδιάστατα μοντέλα υψηλής ανάλυσης με φωτορεαλιστική υφή, τα οποία αφενός μεν αξιοποιήθηκαν ως δείγμα σε οστεομετρικές αναλύσεις, πολλές εκ των οποίων δεν είναι δυνατό να εφαρμοστούν στα πραγματικά οστά με κλασσικές οστεομετρικές μεθόδους, αφετέρου δε αποτελούν μία παρακαταθήκη, ως ψηφιακή βιβλιοθήκη, για μελλοντικές έρευνες επιτρέποντας ταυτόχρονα την βέλτιστη φύλαξη και συντήρηση της σύγχρονης σκελετικής συλλογής, στην οποία πλέον η φυσική πρόσβαση δεν είναι απαραίτητη.

Στο πλαίσιο του διδακτορικού έργου, έγιναν μία σειρά από μελέτες με κύριο άξονα τις βιολογικές παραμέτρους του φύλου και της ηλικίας θανάτου. Υπό αυτό το πρίσμα, μελετήθηκε η έκφραση του φυλετικού διμορφισμού του κρανίου και αξιοποιώντας τους βέλτιστους μετρικούς χαρακτήρες υπολογίστηκαν διακριτικές εξισώσεις εκτίμησης φύλου εξειδικευμένες για τον ελληνικό πληθυσμό, οι οποίες έδωσαν ακρίβεια ταυτοποίησης άνω του 95%. Αντίστοιχη μελέτη έγινε και για το οστό της κάτω γνάθου, το οποίο αν και παρουσιάζει μικρότερο φυλετικό διμορφισμό από το κρανίο μπορεί να αξιοποιηθεί για την ταυτοποίηση του φύλου ενός αγνώστου ατόμου. Η εν λόγω μελέτη έδειξε επίσης ότι η έκφραση του φυλετικού διμορφισμού της κάτω γνάθου έχει πολύ έντονη πληθυσμιακή εξειδίκευση.

Επιπλέον, μελετήθηκε η μορφολογία του βοθρίου της κεφαλής του μηριαίου τόσο ως προς το φύλο όσο και ως προς την ηλικία θανάτου και παρόλο που οι παρατηρούμενες μορφολογικές αλλαγές δεν καθιστούν το βοθρίο κατάλληλο σκελετικό δείκτη για την εκτίμηση φύλου ή ηλικίας θανάτου, είναι σαφές πως το μέγεθός του σχετίζεται άμεσα με το φύλο (συστηματικά μεγαλύτερο

στα αρσενικά άτομα) ενώ το σχήμα του με την ηλικία θανάτου (πιο ακανόνιστη περίμετρος με την αύξηση της ηλικίας του ατόμου). Αξιοποιώντας το μέγιστο μήκος των μηριαίων μελετήθηκαν οι διαχρονικές τάσεις στο ανάστημα του ελληνικού πληθυσμού για την περίοδο αναφοράς (1879-1965), οι οποίες, αν και ανοδικές για το συνολικό εύρος της εξεταζόμενης χρονικής περιόδου, έδειξαν ότι τα αρσενικά άτομα είτε ήταν πιο ευάλωτα είτε δέχθηκαν μεγαλύτερη περιβαλλοντική πίεση από τις κοινωνικοοικονομικές συνθήκες που σχετίζονται με τους πολέμους του 1ου μισού του 20ού αιώνα.

Κατά τη δημιουργία της ψηφιακής σκελετικής συλλογής, πραγματοποιήθηκε μία μελέτη ορθότητας κι ακρίβειας στην λήψη οστεομετρικών δεδομένων από εικονικά (ψηφιακά) σκελετικά κατάλοιπα αξιοποιώντας τα μοντέλα των κρανίων, η οποία έδειξε ότι η χρήση τριδιάστατων μοντέλων με φωτορεαλιστική υφή που δημιουργούνται με φωτογραμμετρία όχι μόνο αποτελούν αξιόπιστα υποκατάστατα των πραγματικών οστών, αλλά σε πολλές περιπτώσεις επιτρέπουν καλύτερη ακρίβεια μετρήσεων σε σύγκριση με χειροκίνητα οστεομετρικά όργανα. Αξιοποιώντας περαιτέρω τα μοντέλα των κρανίων, πραγματοποιήθηκε μία αξιολόγηση του διεθνώς αναγνωρισμένου λογισμικού 3D-ID ως προς τη δυνατότητα εκτίμησης καταγωγής και φύλου από κρανιακά κατάλοιπα του ελληνικού πληθυσμού. Τα αποτελέσματα έδειξαν πολύ χαμηλά ποσοστά σωστής ταξινόμησης σε στενά προκαθορισμένες ομάδες φανερώνοντας τις αδυναμίες του λογισμικού αλλά και τους περιορισμούς της ακολουθούμενης μεθόδου (3DGM), ενώ έδειξαν μία καλύτερη απόκριση σε πιο διευρυμένες πληθυσμιακές ομάδες (π.χ. Αφρικανοί, Ασιάτες, Ευρωπαίοι κ.α.).

Με στόχο τη βέλτιστη αξιοποίηση των τριδιάστατων μοντέλων δημιουργήθηκε μία καινοτόμος μέθοδος αυτόματου υπολογισμού των παραμέτρων εγκάρσιων τομών της διάφυσης των μακρών οστών, η οποία αντικαθιστά πλήρως την κοστοβόρα, τόσο σε υλικά όσο και σε απαιτούμενη εργασία, μέθοδο καλουπιών σιλικόνης (Latex Cast Method) εξαλείφοντας ταυτόχρονα το σφάλμα μέτρησης που προκύπτει από την εφαρμογή της τελευταίας. Αυτή η νέα προτεινόμενη μέθοδος υλοποιήθηκε ως ελεύθερο λογισμικό ανοικτού κώδικα, με την ονομασία “long-bone-diaphyseal-CSG-Toolkit”. Η εν λόγω μέθοδος αξιοποιήθηκε για τη μελέτη του φυλετικού διμορφισμού στα μακρά οστά (μηριαίο, κνήμη, βραχιόνιο) του ελληνικού πληθυσμού, από όπου προέκυψε και μία καινούρια μεθοδολογία για την εκτίμηση φύλου από μακρά οστά, η οποία ειδικά για το βραχιόνιο άγγιξε το 97,3% ως προς την ακρίβεια επιτυχούς ταξινόμησης. Τα δεδομένα των παραμέτρων εγκάρσιων τομών των διαφύσεων σε συνδυασμό με κατηγορικά δεδομένα από τις αλλοιώσεις των ενθέσεων του μηριαίου και της κνήμης, ως γενικοί δείκτες δραστηριότητας, αξιοποιήθηκαν περαιτέρω για την μελέτη της επίδρασης της δραστηριότητας του ατόμου σε συγκεκριμένους

σκελετικούς ηλικιακούς δείκτες των ανώνυμων οστών, οι οποίοι αξιοποιούνται ευρέως για την εκτίμηση της ηλικίας θανάτου.

Η τελευταία ερευνητική εργασία του παρόντος διδακτορικού έργου αφορά στη δημιουργία μίας καινοτόμου μεθόδου προσδιορισμού του φύλου από εικονικά κρανία με αυτόματο τρόπο και χωρίς μετρητικό σφάλμα που να σχετίζεται με τον παρατηρητή. Η νέα αυτή μέθοδος, χρησιμοποιώντας στατιστικές μεθόδους μηχανικής εκμάθησης (machine learning) αλλά και νέες μορφομετρικές παραμέτρους εμπνευσμένες από υφιστάμενες μεθόδους υπολογιστικής όρασης (computer vision), συνδυάζει την ακρίβεια και επαναληψιμότητα των μετρικών χαρακτήρων με την διαγνωστική ικανότητα της μακροσκοπικής μορφολογικής ανάλυσης στον εντοπισμό λεπτών διαφορών μεταξύ των δύο φύλλων. Η νέα μέθοδος, η οποία υλοποιήθηκε ως ελεύθερο λογισμικό ανοικτού κώδικα, με την ονομασία “skullanalyzer”, δοκιμάστηκε σε συνδυασμό με ένα επιπλέον δείγμα ευρωπαϊκού πληθυσμού και έδωσε επιτυχή ταξινόμηση άνω του 95% καθώς και μειωμένη πληθυσμιακή εξειδίκευση, κάτι το οποίο αποτελεί σαφές πλεονέκτημα έναντι υφιστάμενων μεθόδων.

Abstract

The purpose of this doctoral study was to expand our knowledge on the diversity of the human skeleton of a modern Greek population and to develop new methodologies and tools for extracting biological parameters, such as sex and age-at-death, from human skeletal remains. The whole research was based on the modern skeletal reference collection, housed in the Physical Anthropology laboratory of the Department of Animal and Human Physiology of the Department of Biology, which is the most complete skeletal reference collection in Greece and is considered a representative sample of the Greek population of the 20th century. A major goal has been the digital documentation of a subset of the available skeletal sample, which an array of individual research questions focused upon. The digital documentation, based on three-dimensional photogrammetry, covers the skull, the mandible, the os coxae, the femurs, the tibiae, and the humeri from all adult individuals of the skeletal collection, where available. The completion of the digital documentation yielded 2,117 high-precision three-dimensional mesh models with high-resolution photorealistic texture. These were used as a sample in numerous osteometric analyses, many of which cannot be applied to dry bones with classical osteometric methods. Furthermore, they are a consignment, as a digital library, for future research while allowing for optimal storage and maintenance of the modern skeletal collection, to which physical access is no longer necessary.

A series of studies were conducted focusing on the biological parameters of sex and age-at-death within the doctorate. In this light, the expression of cranial sexual dimorphism was studied and employing the optimal metric characteristics, population-specific discriminant function equations for estimation of sex have been formulated, which gave an identification accuracy of over 95%. A similar study was conducted for the mandible, which can facilitate sex estimation of an unknown individual, despite being less sexually dimorphism compared to the skull. This study also revealed that the expression of the sexual dimorphism of the mandible is highly population specific.

In addition, the morphological variation of the *fovea capitis femoris* with respect to sex and age-at-death was investigated, and although it may not to be suitable for estimating sex or age-at-death in the forensic context, it was shown that the size is directly related to sex (systematically larger in males) while its shape with age-at-death (more irregular perimeter with increasing age). Furthermore, the secular change in stature of the Greek population for the reference period (1879-1965) was studied utilizing the maximum femoral length. According to the results, the stature increases over the total range of the examined period as expected. However, male individuals were

either more vulnerable or received greater environmental stress from the socio-economic conditions associated with the wars of the first half of the 20th century as opposed to females.

During the development of the digital skeletal collection, a study was performed on the accuracy and precision in obtaining osteometric data from virtual (digital) skeletal remains using skull models. This study revealed that the use of three-dimensional models with photorealistic texture created with photogrammetry not only are reasonable substitutes for dry bones but in many cases allow for better measurement accuracy compared to manual osteometric instruments. Further exploiting the 3D cranial models, an evaluation of the reliability of 3D-ID software in sex and ancestry estimation was performed. The results showed a very low percentage of correct classification in narrowly defined groups, outlining certain weaknesses of the software but also the limitations of the adopted method (3DGM). However, the performance of 3D-ID software increases considerably when it comes to large population groups (eg Africans, Asians, Europeans, etc.).

In order to make the best use out of the three-dimensional models, a novel method for automatically computing the long bones' diaphyseal cross-sectional geometric (CSG) properties was developed. This method, called the "long-bone-diaphyseal-CSG-Toolkit", was implemented as a free open source software aiming to replace the costly and time-consuming latex cast method (LCM) while at the same time it eliminates the measurement error resulting from the implementation of the latter. The CSG-Toolkit was further used to develop an automated method for estimating sex from the lower and upper limbs (femur, tibia, humerus) of the Greek population, which resulted in a classification accuracy reaching 97.3% for the humerus. Additionally, the long bones' diaphyseal cross-sectional geometric properties along with the categorical data from the enthesal changes of the femur and tibia, being evaluated as skeletal activity markers, were used to study the impact of physical activity on specific pelvic age indicators, which are widely used to estimate age-at-death.

The latest research work of this doctorate concerns the development of an innovative method for estimating sex from virtual skulls in an automatic way, which eliminates any observer error. This new method combines the accuracy and repeatability of metric characteristics with the diagnostic capability of macroscopic morphological analysis to identify subtle differences between males and females by using machine learning techniques and new morphometric features inspired by existing computer vision methods. The new method, called "skullanalyzer", is also implemented as free open source software and was evaluated using two European population samples. According to the results, "skullanalyzer" yields accuracy in excess of 95% and is less prone to population specificity, which demonstrates a clear advantage over existing methods.

Ευχαριστίες

Τα προηγούμενα τέσσερα κι πλέον χρόνια αποτέλεσαν μία συναρπαστική διαδρομή (ενίοτε και περιπέτεια) κατά την οποία το σημαντικότερο και καθοριστικό στοιχείο της υπήρξαν οι άνθρωποι με τους οποίους συμπορεύτηκα όλο αυτό το διάστημα τόσο εντός όσο και εκτός εργαστηρίου και ερευνητικού αντικειμένου.

Ευχαριστώ θερμά την Ελισσάβητ (aka κομάντο), την Κατερίνα (aka κρητικιά νύφη), την Κωνσταντίνα (aka γκομενάκι) και τη Νεφέλη (aka γαρουφοχωριό), η συμβολή των οποίων υπήρξε καθοριστική για την περάτωση της ατελείωτης «λάντζας» που απαιτήθηκε για την τριδιάστατη μοντελοποίηση του σκελετικού υλικού πάνω στο οποίο βασίστηκε όλη η ερεύνα που παρουσιάζεται στην παρούσα διατριβή και όχι μόνο. Είμαι ευγνώμων και τυχερός που διασταυρώθηκαν οι δρόμοι μας στο εργαστήριο Φυσικής Ανθρωπολογίας (aka υπόγα) και είμαι πεπεισμένος ότι έμαθα πολύ περισσότερα από αυτές από ότι κατάφερα να τους μάθω επιβλέποντας τις διπλωματικές τους εργασίες. Εύχομαι να είστε εξίσου τυχερές με τους ανθρώπους που θα περάσουν από τις ζωές σας.

Επίσης, θέλω να ευχαριστήσω τη Σοφία (aka μικρή επιστημόνισσα) για τη στήριξή της στα καλά και στα δύσκολα, την Ειρήνη (aka μωρό) για όλα όσα περάσαμε μαζί όλα αυτά τα χρόνια μέσα στο Τμήμα Βιολογίας (aka άσυλο) και όχι μόνο, την κυρία Αναστασία (aka μάνα) για την απρόσκοπτη παροχή του μεσημεριανού μας φαγητού (έχω την εντύπωση ότι ο αριθμός γευμάτων ξεπερνάει τον αριθμό των τριδιάστατων μοντέλων που φτιάξαμε όλο αυτό το διάστημα), την αγαπημένη όλων μας Γωγώ (aka L.A.Woman), της οποίας η παρουσία έκανε τα μεσημεριανά μας διαλείμματα αξέχαστα, τη Šárka (aka τρελή φωτογράφος) για τη φιλοξενία και τις όμορφες αναμνήσεις από την πρακτική μου άσκηση στην Πράγα, καθώς και τον Sam (aka Brian) για το πιο ευχάριστο και Monty-Python_ικό τρίμηνο lockdown (χειμώνας '20) που περάσαμε κλειδωμένοι στην υπόγα!

Θα ήθελα επίσης να ευχαριστήσω τα μέλη της τριμελούς συμβουλευτικής επιτροπής Χριστίνα Παπαγεωργοπούλου, η οποία μας βόηθησε στα πρώτα μας βήματα, και Ιωάννη Λυριτζή, ο οποίος στάθηκε και η αφορμή (Δελφοί '16) για να ξεκινήσω το διδακτορικό, καθώς και τη φίλη και συνεργάτιδα Ευθυμία Νικήτα για την πολύτιμη βοήθεια και εξαιρετική συνεργασία όλα αυτά τα χρόνια. Τη θεωρώ ένα άτυπο τέταρτο μέλος της συμβουλευτικής επιτροπής σε αυτό διδακτορικό.

Κλείνοντας, τα λόγια περισσεύουν για την αγαπημένη μου Μαριλένα (aka μπικ μπος), η οποία πέραν της επιστημονικής της συνδρομής σε όλη αυτή τη διαδικασία είχε και το κουράγιο να υποστηρίξει ένα δεύτερο διδακτορικό στο ίδιο σπίτι μετά την ολοκλήρωση του δικού της, και για τον επιβλέποντα της καρδιάς μου Στρατή (aka τρελό αγόρι), ο οποίος οραματίστηκε να κάνω διδακτορικό πολύ πριν ξαναμπεί η ιδέα στο μυαλό μου αλλά πρωτίστως είχε το θάρρος να στηρίξει την αίτησή μου για διδακτορικό και τη συνέπεια να αναλάβει την επίβλεψή του εκεί που άλλοι αρνήθηκαν κατά το παρελθόν. Είναι προφανές ότι ο Στρατής, με την απόλυτη εμπιστοσύνη που μου έδειξε αλλά και την κάλυψη που μας προσέφερε απλόχερα όλα αυτά τα χρόνια, δεν είναι απλά ένας εξαιρετικός επιβλέποντας για διδακτορικό αλλά είναι και ο μόνος με τον οποίο θα μπορούσα να το φέρω εις πέρας.

Ακρωνύμια

3DGM	3-Dimensional Geometric Morphometrics
ANOVA	Analysis of Variance
ArPerIndex	Area Perimeter Index
CSG	Cross-Sectional Geometry
CSV	Comma Separated Values
CV	Cross-Validated
DFA	Discriminant Function Analysis
EFDs	Elliptic Fourier Descriptors
FCA	Fovea Capitis Area
FCMD	Fovea Capitis Maximum Distance
FHS	Femoral Head Surface
FMHD	Femoral Maximum Head Diameter
GLM	Generalized Linear Models
GNU GPLv3	GNU General Public License version 3
GPA	Generalized Procrustes Analysis
GUI	Graphical User Interface
HMI	Height Map Image
HOGs	Histogram of Oriented Gradients
ILA	Inter-Landmark angle
ILD	Inter-Landmark distance
LDA	Linear Discriminant Analysis
LOOCV	Leave One Out Cross-Validation
LRMSD	Least Root-Mean-Square Error
OR	ORiginal group
PCA	Principal Component Analysis
RBF	Radial Basis Function
SMA	Second Moments of Area
SVM	Support Vector Machine
TA	Total subperiosteal Area
TEM	Technical Error Measurement

Περιεχόμενα

Περίληψη.....	i
Abstract.....	iv
Ευχαριστίες.....	vi
Ακρωνύμια.....	vii
Εισαγωγή.....	1
1. Ψηφιακή τεκμηρίωση του σκελετικού υλικού με φωτογραμμετρία.....	4
1.1. Φωτογράφιση.....	4
1.2. Φωτογραμμετρία.....	7
1.3. Λάντζα.....	11
2. Σύνοψη ερευνητικού έργου.....	13
3. Μορφομετρική διερεύνηση του φυλετικού διμορφισμού του κρανίου.....	20
3.1. Υλικά και μέθοδοι.....	21
3.2. Αποτελέσματα.....	22
3.3. Συμπεράσματα.....	24
4. Ποικιλομορφία του βοθρίου της κεφαλής του μηριαίου.....	26
4.1. Υλικά και μέθοδοι.....	26
4.2. Αποτελέσματα.....	28
4.3. Συμπεράσματα.....	29
5. Διαχρονικές τάσεις στην εξέλιξη του ύψους στον Ελληνικό πληθυσμό.....	31
5.1. Υλικά και μέθοδοι.....	31
5.2. Αποτελέσματα.....	32
5.3. Συμπεράσματα.....	35
6. «Τι και πώς πρέπει να μοιραζόμαστε μετρικά δεδομένα» Μια σύγκριση του σφάλματος μέτρησης κρανιακών αναφορικών σημείων μεταξύ μεθόδων και παρατηρητών.....	37
6.1. Υλικά και μέθοδοι.....	38
6.2. Αποτελέσματα.....	39
6.3. Συμπεράσματα.....	40
7. Έλεγχος αξιοπιστίας του λογισμικού 3D-ID στην εκτίμηση του φύλου και της καταγωγής με την εφαρμογή σύγχρονου Ελληνικού πληθυσμού.....	42
7.1. Υλικά και μέθοδοι.....	43
7.2. Αποτελέσματα.....	44
7.3. Συμπεράσματα.....	46
8. Προσδιορισμός φύλου του σύγχρονου ελληνικού πληθυσμού από διαστάσεις της κάτω γνάθου.....	48
8.1. Υλικά και μέθοδοι.....	48
8.2. Αποτελέσματα.....	50
8.3. Συμπεράσματα.....	54
9. Μια καινοτόμος μέθοδος για την γεωμετρική ανάλυση της εγκάρσιας τομής της διάφυσης των μακρών οστών.....	55
9.1. Προδιαγραφές λογισμικού και περιγραφή αλγορίθμου.....	56
9.2. Έλεγχος ακρίβειας και αξιοπιστίας του CSG Toolkit.....	58
10. Πλεονεκτήματα στον προσδιορισμό του φύλου αξιοποιώντας γεωμετρικές παραμέτρους της εγκάρσιας τομής της διάφυσης των κάτω και άνω άκρων.....	61
10.1. Υλικά και μέθοδοι.....	62
10.2. Αποτελέσματα.....	64
10.3. Συμπεράσματα.....	66
11. Η επίδραση της δραστηριότητας κατά την εκτίμηση ηλικίας θανάτου από τα οστά της λεκάνης.....	68

11.1. Υλικά και μέθοδοι εκτίμησης ηλικίας θανάτου.....	68
11.2. Εκτίμηση του βαθμού σωματικής δραστηριότητας/καταπόνησης.....	69
11.3. Εκτίμηση σωματικής μάζας.....	70
11.4. Στατιστική ανάλυση.....	70
11.5. Αποτελέσματα και Συμπεράσματα.....	71
12. Προηγμένη μέθοδος εκτίμησης φύλου βάση φυλετικά διμορφικών κρανιακών μορφολογικών χαρακτηριστικών.....	78
12.1. Υλικά και μέθοδοι.....	79
12.1.1. Ανάλυση δεδομένων - διαχωρισμός δειγμάτων.....	80
12.1.2. Ανάλυση δεδομένων - μηχανική μάθηση.....	81
12.1.3. Ανάλυση δεδομένων – υποτύπωση αποτελεσμάτων.....	81
12.1.4. Ανάλυση δεδομένων – συνδυασμός μοντέλων.....	82
12.2. Αποτελέσματα.....	82
12.3. Συμπεράσματα.....	86
Απολογισμός έργου.....	88
Βιβλιογραφία.....	89
Βιογραφικό Σημείωμα.....	106
Ανατύπωση πρωτότυπων δημοσιεύσεων.....	111

Εισαγωγή

Η οστεολογία, ως διακριτό επιστημονικό πεδίο, μελετά τα ιδιαίτερα χαρακτηριστικά της μορφολογίας αλλά και της φυσιολογίας των οστών, τα οποία τους προσδίδουν ισχυρή ανθεκτικότητα και καθιστούν δυνατή την διατήρησή τους στο πέρασ του χρόνου (Lyman, 1994). Σε ακριβώς αυτή τη δυνατότητα διατήρησής τους στο χρόνο εδράζεται και η ουσιώδης συνεισφορά της Οστεολογίας σε διάφορους επιστημονικούς κλάδους, όπως η Φυσική και η Δικαστική Ανθρωπολογία, η Βιοαρχαιολογία, η Παλαιοντολογία, οι οποίοι βασίζονται στην οστεολογία για την άντληση πληροφορίας από τα οστά (ανθρώπινα και μη) με σκοπό την απάντηση των επί μέρους ερωτημάτων τους.

Η οστεολογική μελέτη στοχεύει αφενός μεν στην ανάκτηση πληροφοριών από τα οστά αφετέρου δε στην έρευνα για την ανάπτυξη και εξέλιξη μεθόδων που βελτιώνουν την ποιότητα αυτών των πληροφοριών ή διευρύνουν το πεδίο εφαρμογής τους. Ως εκ τούτου, κάθε οστεολογική μελέτη μπορεί να περιλαμβάνει σειρά από αναλύσεις που αποσκοπούν, μεταξύ άλλων, στον προσδιορισμό του φύλου και την εκτίμηση της ηλικίας θανάτου, στην ερμηνεία παθολογικών καταστάσεων και αίτια θανάτου, καθώς και στην διατροφή, μετακίνηση και καταγωγή. Οι τεχνικές ανάλυσης ποικίλουν σημαντικά ανάλογα με το ερώτημα που στοχεύουν από σύνθετες μοριακές και βιοχημικές τεχνικές, όπως η χρήση ραδιοϊσοτόπων ή σταθερών ισοτόπων και η ανάλυση DNA, έως πιο απλές μορφολογικές αναλύσεις που βασίζονται είτε στην οπτική παρατήρηση είτε στην μορφομετρία – οστεομετρία. Οι μορφολογικές αναλύσεις είναι και οι πλέον διαδεδομένες καθώς αποτελούν μη καταστρεπτικές μεθόδους, οι οποίες όμως μπορούν να μας δώσουν πληθώρα πληροφοριών από τα οστά, και με πολύ χαμηλό κόστος έναντι των σύνθετων μοριακών – βιοχημικών μεθόδων. Ένα χαρακτηριστικό παράδειγμα αποτελεί η εκτίμηση φύλου με οστεομετρικές – μορφολογικές μεθόδους έναντι της ανάλυσης DNA (Brůžek et al., 2017; Bruzek, 2002; Thomas, 2020). Ειδικότερα, η οστεομετρία έχει ευρεία εφαρμογή στην οστεολογική μελέτη καθώς ως ποσοτική μέθοδος ανάλυσης παρέχει την αυξημένη ακρίβεια, επαναληψιμότητα και αναπαραγωγικότητα ως προς το σφάλμα του παρατηρητή σε σύγκριση με αντίστοιχες οπτικές μεθόδους και ως εκ τούτου θεωρείται πιο αξιόπιστη (Bigoni et al., 2010; Spradley, 2016).

Τις προηγούμενες δεκαετίες, η ανάλυση των οστών με τη χρήση μορφομετρικών προσεγγίσεων βασίζονταν αποκλειστικά σε μετρήσεις γραμμικών αποστάσεων και γωνιών με την χρήση κλασικών χειροκίνητων οργάνων μέτρησης όπως ο οστεομετρικός πίνακας, το γωνιόμετρο και το κεφαλόμετρο. Ακολούθως, χρησιμοποιήθηκαν κατά τον ίδιο τρόπο ψηφιακά όργανα τα οποία μειώνουν μεν το συστηματικό σφάλμα των μετρήσεων χωρίς όμως να βελτιώνουν την

αντικειμενικότητα των μετρήσεων που βασίζεται στον χειριστή. Η τεχνολογική έκρηξη των δύο τελευταίων δεκαετιών, βέβαια, έδωσε νέα ώθηση στην ενσωμάτωση νέων τεχνολογιών και εργαλείων στην οστεομετρία. Χαρακτηριστικό παράδειγμα είναι οι ακίδες τριδιάστατης ψηφιοποίησης, όπως το 3D Microscribe, οι οποίες όχι μόνο άλλαξαν τα μετρητικά δεδομένα (data), αλλά άνοιξαν το δρόμο σε νέες μεθόδους μορφολογικής ανάλυσης, όπως η τριδιάστατη γεωμετρική μορφομετρία (3DGM) με πληθώρα ερευνητικού έργου (Bigoni et al., 2010; Choivalopoulou, Bertsatos, et al., 2017; Choivalopoulou et al., 2013, 2016a, 2016b; Choivalopoulou, Papageorgopoulou, et al., 2017; Franklin et al., 2007; Fruciano, 2016; Ross et al., 2010; Ross & Williams, 2008; Slice & Ross, 2018; Spradley & Jantz, 2016; von Cramon-Taubadel et al., 2007; Williams & Richtsmeier, 2003). Φυσικά μέσα στα σύγχρονα εργαλεία μορφολογικής μελέτης των οστών ενσωματώθηκε πλήθος τεχνολογιών από ιατρικά μηχανήματα τριδιάστατης απεικόνισης, όπως MRI και CT scanners, έως εξοπλισμός τριδιάστατης ψηφιακής τεκμηρίωσης δανειζόμενος από την τοπογραφία και τη διατήρηση της πολιτιστικής κληρονομιάς, όπως 3D Laser scanners, structured light scanners και τριδιάστατη φωτογραμμετρία. Όλα αυτά επιτρέπουν την ψηφιακή τεκμηρίωση αλλά και οπτικοποίηση της τριδιάστατης μορφολογίας των οστών σε H/Y, διευκολύνοντας την λήψη των μετρήσεων απευθείας από τα ψηφιακά τριδιάστατα μοντέλα. Η διαδικασία κατασκευής τριδιάστατων μοντέλων είναι ένα πολύτιμο εργαλείο για την αποσαφήνιση των μορφολογικών διαφορών που παρουσιάζονται μεταξύ των σκελετικών δειγμάτων συγκριτικά με αναλύσεις που βασίζονται σε μη μετρικά χαρακτηριστικά τα οποία δεν μπορούν εύκολα να περιγραφούν με τη χρήση απλών γραμμικών μετρήσεων (Waltenberger et al., 2021).

Η οστεομετρική χρήση τριδιάστατων ψηφιακών μοντέλων έναντι των πραγματικών οστών εμφανίζουν σημαντικά πλεονεκτήματα. Κατ' αρχήν, βελτιστοποιείται η διατήρησή τους καθώς όλες οι τεχνικές τριδιάστατης ψηφιακής τεκμηρίωσης είναι ανέπαφες. Εξίσου σημαντικές είναι οι δυνατότητες απομακρυσμένης διαχείρισης και διαμοιρασμού των “ψηφιακών” οστών καθώς πρόκειται για τυπικά αρχεία H/Y αίροντας ουσιαστικά τους όποιους γεωγραφικούς περιορισμούς είτε πρόκειται για απομακρυσμένη εργασία μακριά από τον χώρο του εργαστηρίου, είτε για διεθνείς συνεργασίες, είτε για ομάδα που μπορεί να δουλέψει ταυτόχρονα στο ίδιο υλικό μειώνοντας τον συνολικό χρόνο για την λήψη αποτελεσμάτων. Το πιο σημαντικό, όμως, πλεονέκτημα της ψηφιακής τεκμηρίωσης των οστών είναι η πληρότητα της αποτύπωσης της γεωμετρίας τους χωρίς αυτό να συνεπάγεται ότι δεν υπάρχουν διαβαθμίσεις ως προς την ακρίβεια (Waltenberger et al., 2021) αλλά και ως προς την αρτιότητα της απεικόνισης της συνολικής τους μορφολογίας (Muñoz-Muñoz et al., 2016) σε σχέση με την επιλεγόμενη μέθοδο ψηφιοποίησης. Όλα τα παραπάνω σε συνδυασμό με το προσιτό πλέον κόστος των τεχνολογιών τριδιάστατης ψηφιακής τεκμηρίωσης έχουν κάνει τη χρήση τους πολύ διαδεδομένη στην οστεολογία, τόσο στη

βασική οστεολογική έρευνα όσο και στην εφαρμογή των μεθόδων της σε συγγενικούς κλάδους που προαναφέρθηκαν. Παρ' όλα αυτά, μέχρι και σήμερα, η συντριπτική πλειοψηφία των μελετών, που χρησιμοποιούν μεθόδους ψηφιοποίησης των οστών, εξακολουθεί να βασίζεται στον παρατηρητή – χρήστη κατάλληλου λογισμικού – για τη χειροκίνητη εξαγωγή δεδομένων από τα “ψηφιακά” οστά. Σημαντικότερο ακόμα είναι ότι αυτά τα “ψηφιακά” μετρικά δεδομένα είναι ως επί το πλείστον σε αντιστοιχία με τα μετρικά δεδομένα των κλασσικών χειροκίνητων οργάνων μέτρησης. Αυτό φυσικά δεν σημαίνει ότι εκλείπουν τα πλεονεκτήματα από την ενσωμάτωση τεχνολογιών τριδιάστατης τεκμηρίωσης, αλλά ότι δεν αξιοποιούνται στο πλήρες δυναμικό τους. Οπότε, από το σύνολο της γεωμετρικής και μορφολογικής πληροφορίας ενός τριδιάστατου ψηφιακού μοντέλου αξιοποιείται μόλις ένα πολύ μικρό ποσοστό αυτής, κάνοντας πρόδηλη την αναγκαιότητα για νέες μεθόδους γεωμετρικών και μορφομετρικών αναλύσεων για την εξαγωγή πληρέστερης πληροφορίας από τα σκελετικά κατάλοιπα.

Αυτό ακριβώς το κενό ήταν και ένας από τους βασικούς στόχους της παρούσας διδακτορικής διατριβής, η οποία αποσκοπούσε στη δημιουργία νέων αυτοματοποιημένων μεθόδων εξόρυξης μορφομετρικών δεδομένων για την ανάκτηση βιολογικών παραμέτρων από ψηφιοποιημένα οστά, όπως το φύλλο, η ηλικία θανάτου, το επίπεδο δραστηριότητας και το ανάστημα. Σε αυτό το πλαίσιο υλοποιήθηκε και η ψηφιακή τεκμηρίωση ενός συνόλου οστών από τη διαθέσιμη σκελετική συλλογή αναφοράς του σύγχρονου ελληνικού πληθυσμού που διατηρείται στον Τομέα Φυσιολογίας Ζώων και Ανθρώπου του Τμήματος Βιολογίας του ΕΚΠΑ και η οποία αποτελεί το πρωτογενές υλικό μελέτης του διδακτορικού. Στόχοι της ψηφιακής τεκμηρίωσης ήταν τόσο η δημιουργία των απαραίτητων δεδομένων για το ερευνητικό σκέλος της διατριβής όσο και η ουσιαστική συμβολή στην διατήρηση και περαιτέρω ανάδειξη της σύγχρονης σκελετικής συλλογής αναφοράς, ένα σημαντικό έργο που επιτελεί απρόσκοπτα παραπάνω από 20 χρόνια ο Τομέας Φυσιολογίας Ζώων και Ανθρώπου. Παράλληλα, η παρούσα εργασία στο σύνολό της στόχευσε να αντλήσει πλήθος νέες πληροφορίες που αφορούν στην σκελετική μορφολογία, και όχι μόνο, συμβάλλοντας ουσιαστικά στη διεύρυνση της γνώσης μας για το συγκεκριμένο επιστημονικό πεδίο. Τα κεφάλαια που ακολουθούν παρουσιάζουν αναλυτικά το σύνολο του ερευνητικού έργου που εκπονήθηκε τα τελευταία 4 χρόνια στο πλαίσιο της παρούσας διδακτορικής διατριβής.

1. Ψηφιακή τεκμηρίωση του σκελετικού υλικού με φωτογραμμετρία

Η τεχνική της φωτογραμμετρίας είναι σχεδόν τόσο παλιά όσο και η φωτογραφική τεχνολογία περί τα μέσα του 19ου αιώνα. Πρωτοχρησιμοποιήθηκε γύρω στο 1840 από τον Γάλλο τοπογράφο Dominique François Jean Arago για τον υπολογισμό τοπογραφικών χαρτών ενώ ως όρος, η φωτογραμμετρία, πρωτοπροτάθηκε το 1867 από τον Πρώσο αρχιτέκτονα Albrecht Meydenbauer (1867) και αναφέρεται στην τεχνολογία προσδιορισμού των διαστάσεων ενός αντικειμένου από τη φωτογραφική αποτύπωση αυτού. Η τριδιάστατη φωτογραμμετρία είναι μία υποκατηγορία της φωτογραμμετρίας και αναφέρεται στη δημιουργία τριδιάστατων ψηφιακών μοντέλων ως ακριβή αποτύπωση πραγματικών αντικείμενων. Η τριδιάστατη φωτογραμμετρία βασίζεται στη χρήση αλγόριθμων ταυτοποίησης κοινών σημείων σε επικαλυπτόμενες φωτογραφίες ενός αντικειμένου από πολλές οπτικές γωνίες και μεθόδων υπολογιστικής γεωμετρίας για την δημιουργία τριδιάστατων μοντέλων από τα αντικείμενα, που αποτυπώνονται στις εν λόγω φωτογραφίες. Η μεθοδολογική ροή της τριδιάστατης φωτογραμμετρίας χωρίζεται σε δύο τμήματα. Το πρώτο αφορά στη λήψη πολλαπλών φωτογραφιών του αντικειμένου προς μοντελοποίηση και το δεύτερο στη φωτογραμμετρική επεξεργασία τους με κατάλληλο λογισμικό σε Η/Υ για την παραγωγή του τριδιάστατου μοντέλου του πραγματικού αντικειμένου που φωτογραφήθηκε.

1.1. Φωτογράφιση

Η λήψη των φωτογραφιών που απαιτούνται για την τριδιάστατη φωτογραμμετρία είναι το πιο ουσιαστικό/σημαντικό στάδιο της ψηφιακής τεκμηρίωσης που καθορίζει εν πολλοίς και την ποιότητα του τελικού τριδιάστατου μοντέλου. Παρόλα αυτά δεν υπάρχει ένα αυστηρό πρωτόκολλο που να διέπει την τεχνική της φωτογράφισης παρά μία σειρά από οδηγίες οι οποίες επιτρέπουν τη βελτιστοποίηση του παραγόμενου μοντέλου τόσο σε ακρίβεια και πιστότητα καθώς και την αποδοτικότητα της όλης διαδικασίας. Ως εκ τούτου, θα λέγαμε ότι η διαδικασία της φωτογράφισης για τριδιάστατη φωτογραμμετρία είναι ένα μείγμα τέχνης και τεχνικής. Πέραν της προαναφερθείσας απαραίτητης αλληλεπικάλυψης των φωτογραφιών ως προς το αντικείμενο ενδιαφέροντος, οι βασικοί κανόνες (rules of thumb) για μία καλή φωτογραφική λήψη για τη συγκεκριμένη δουλειά είναι οι εξής:

1. Το αντικείμενο θα πρέπει να είναι όσο το δυνατόν εξ' ολοκλήρου εστιασμένο (in focus) ενώ ο περιβάλλον χώρος (background) είναι προτιμότερο να είναι εκτός εστίασης (out of focus). Αν και το βάθος πεδίου μπορεί να ελεγχθεί με το διάφραγμα του φακού, ανάλογα με το μέγεθος το σχήμα του αντικειμένου, η βέλτιστη ρύθμιση συνήθως καταλήγει σε ένα παζάρι μεταξύ ευκρίνειας (sharpness) και βάθος πεδίου (depth of field).

2. Το αντικείμενο θα πρέπει να καλύπτει όσο το δυνατόν μεγαλύτερο μέρος του κάδρου (frame) ώστε να είναι αξιοποιήσιμα περισσότερα pixels από τη διαθέσιμη ανάλυση της εκάστοτε κάμερας. Η υψηλότερη ανάλυση αυξάνει γραμμικά τον χρόνο φωτογραμμετρικής επεξεργασίας αλλά και τις απαιτήσεις (κυρίως σε ελάχιστη μνήμη RAM) του υπολογιστικού συστήματος που απαιτείται για την φωτογραμμετρική επεξεργασία των εικόνων.
3. Κάθε σημείο του αντικειμένου θα πρέπει να αποτυπώνεται σε τουλάχιστον τρεις λήψεις υπό διαφορετική γωνία και κατά προτίμηση έξι για τον βέλτιστο χωροταξικό προσδιορισμό κάθε σημείου. Και πάλι ο βέλτιστος αριθμός καθορίζεται από την επί μέρους μορφολογία του αντικειμένου καθώς και την υφή του.
4. Για κάθε είδος αντικειμένου (υφή, μέγεθος, σχήμα) υπάρχει ένα συγκεκριμένο εύρος βέλτιστου αριθμού απαιτούμενων λήψεων για την ορθή φωτογραμμετρική του αναπαράσταση. Λιγότερες λήψεις μειώνουν δραστικά την ποιότητας του τελικού αποτελέσματος, περισσότερες λήψεις δεν αυξάνουν περαιτέρω την ποιότητά του αλλά εκθετικά τον χρόνο φωτογραμμετρικής επεξεργασίας που απαιτείται.
5. Ο φωτισμός του αντικειμένου πρέπει να είναι κατά το δυνατόν ομοιογενής ώστε να επιτυγχάνεται σωστή χρωματική αποτύπωση και να αποφεύγονται έντονες σκιές, ιδιαίτερος πάνω στο αντικείμενο.
6. Το αντικείμενο πρέπει να είναι σταθερό και να αλλάζει η θέση της κάμερας.

Με βάση τους παραπάνω κανόνες, αλλά και με το γεγονός ότι η παρούσα διδακτορική έρευνα απαιτούσε την ψηφιακή τεκμηρίωση σε τριδιάστατα μοντέλα σχεδόν 2500 οστών, η βελτιστοποίηση της φωτογράφισης βασίστηκε σε μία μεγάλη σειρά δοκιμών (trial 'n' retrial), η οποία ευτυχώς προηγήθηκε της έναρξης του διδακτορικού.

Στο πλαίσιο του διδακτορικού έργου χρησιμοποιήθηκε μία ψηφιακή μηχανή DSLR Canon EOS 100D με ανάλυση 18 Mpixels και φακός σταθερής εστιακής απόστασης Canon EF 40mm f/2.8 STM. Δεδομένου ότι η ψηφιακή τεκμηρίωση έπρεπε να συμπεριλαμβάνει ολόκληρο το οστό, αυτό φωτογραφήθηκε σε δύο διαφορετικές θέσεις με αντιδιαμετρικό προσανατολισμό μεταξύ τους και για κάθε θέση ελήφθησαν 65 με 85 φωτογραφίες από διαφορετικές γωνίες λήψης ανάλογα το οστό. Τα οστά που ψηφιοποιήθηκαν είναι το κρανίο, η κάτω γνάθος, τα βραχιόνια, τα μηριαία, οι κνήμες καθώς και τα δύο ανώνυμα οστά από κάθε σκελετό της σύγχρονης σκελετικής συλλογής αναφοράς του ελληνικού πληθυσμού που φυλάσσεται στο εργαστήριο Φυσικής Ανθρωπολογίας του Τομέα Φυσιολογίας Ζώων και Ανθρώπου του Τμήματος Βιολογίας του ΕΚΠΑ. Και καθώς οι

βελτιστοποίηση μιας διαδικασίας δεν σταματάει με τις βέλτιστες ρυθμίσεις αλλά ολοκληρώνεται μόνο με πατέντα, σχεδιάστηκε και υλοποιήθηκε η ιδιοκατασκευή της παρακάτω φωτογραφίας



Εικόνα 1.1: Περιστρεφόμενη βάση φωτογράφισης

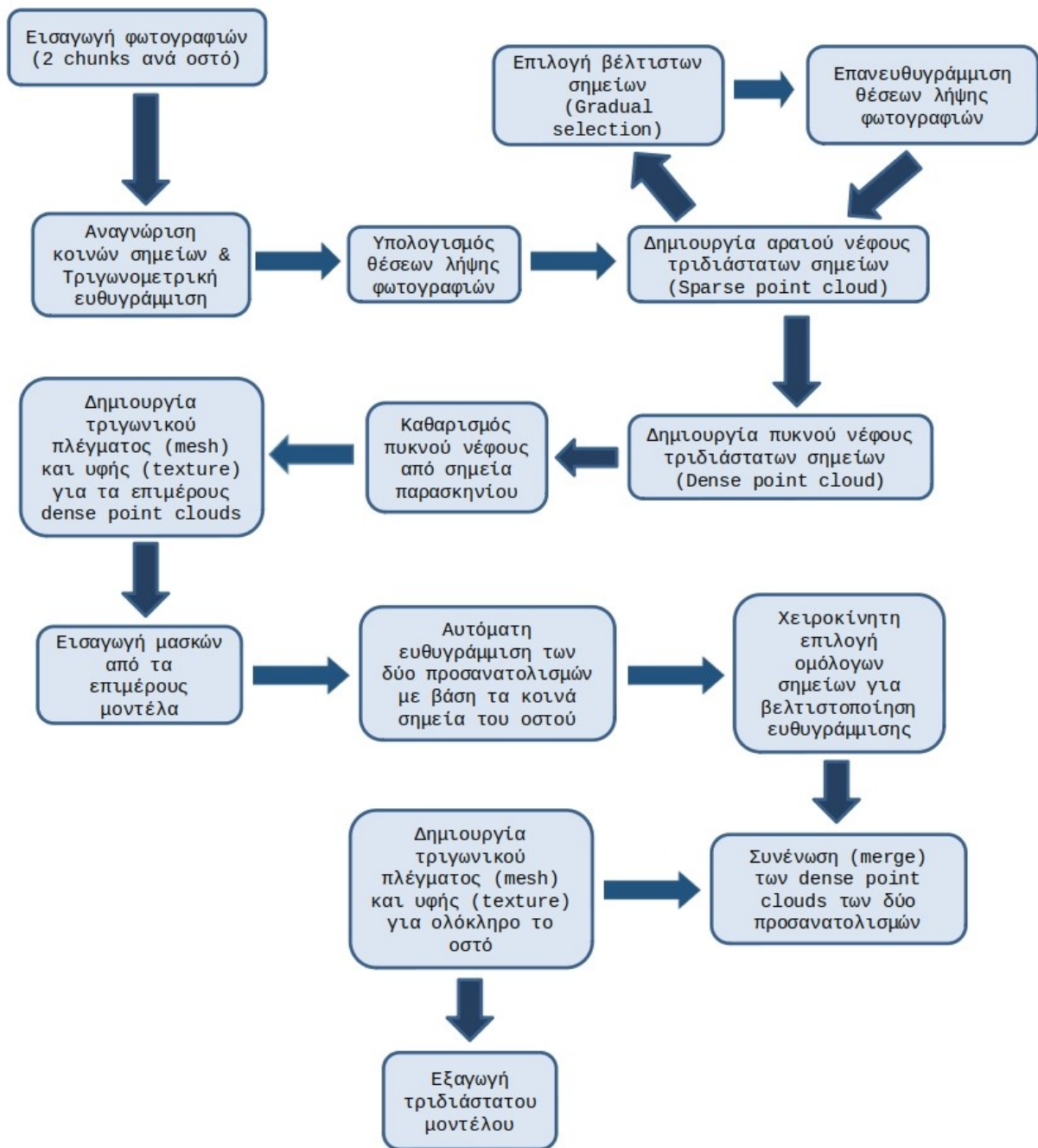
Με βάση τις βέλτιστες ρυθμίσεις, όλες οι φωτογραφίες των κρανίων, των κάτω γνάθων και των ανώνυμων οστών ελήφθησαν σε ημιαυτόματη λειτουργία σταθερού διαφράγματος ρυθμισμένη στα 14 F/stops, ενώ όλα τα μακρά οστά με σταθερό διάφραγμα στα 18 F/stops. Λόγω του μικρού διαφράγματος για τη μεγιστοποίηση του βάθους πεδίου σε κάθε λήψη, ο απαιτούμενος χρόνος έκθεσης καθιστά απαραίτητη τη χρήση σταθερής βάσης (πχ τρίποδο) για τη φωτογραφική μηχανή ώστε να μην επηρεάζεται η ευκρίνεια της φωτογραφίας λόγω κουνήματος. Η εν λόγω ιδιοκατασκευή επιτρέπει την πλήρη φωτογράφιση ενός οστού (2 προσανατολισμοί) σε λιγότερο από 8 λεπτά μειώνοντας δραστικά (x15) τον απαιτούμενο χρόνο για την πλήρη φωτογράφιση ενός αντικειμένου σε σύγκριση με τη χρήση φωτογραφικού τρίποδου για την σταθεροποίηση της κάμερας. Κατά τα αυτόν τον τρόπο, κάθε οστό τοποθετήθηκε στην επίπεδη κυλινδρική βάση ώστε να εξασφαλιστεί η σταθερότητά του κατά την περιμετρική λήψη φωτογραφιών, ενώ η κάμερα, με την βοήθεια μηχανικού βραχίονα, σταθεροποιούταν διαδοχικά σε διαφορετικές γωνίες ανύψωσης (0°, 20°, 45°, 85°) και το βαγονέτο-βάση του βραχίονα κινούταν κυκλικά περιστρεφόμενο γύρω από το οστό ακολουθώντας την πατούρα της ζάντας αυτοκινήτου που είναι και το βασικό στοιχείο της συγκεκριμένης ιδιοκατασκευής.

1.2. Φωτογραμμετρία

Το δεύτερο τμήμα της φωτογραμμετρικής μοντελοποίησης περιλαμβάνει την επεξεργασία των φωτογραφιών σε Η/Υ. Η τυπική ροή φωτογραμμετρικής επεξεργασίας περιλαμβάνει πέντε βασικά στάδια:

- α) την αναγνώριση κοινών χαρακτηριστικών σημείων (feature detection) στις φωτογραφίες,
- β) τον υπολογισμό των θέσεων λήψης της κάμερας μέσω της τριγωνομετρικής ευθυγράμμισης των κοινών σημείων,
- γ) τη δημιουργία αραιού νέφους σημείων (sparse point cloud) με τον τριδιάστατο χωροταξικό υπολογισμό των ευθυγραμμισμένων κοινών χαρακτηριστικών σημείων (Carrivick et al., 2016),
- δ) τη δημιουργία πυκνού νέφους σημείων (dense point cloud) με τη χρήση στερεοσκοπικής ζεύξης μεταξύ των φωτογραφιών (stereo pair selection) και εικόνων βάθους (depth maps) για κάθε φωτογραφία (Shen, 2013), και
- ε) υπολογισμός τριγωνικού πλέγματος (mesh) και υφής επιφανείας (texture) για τη δημιουργία του τελικού τριδιάστατου ψηφιακού μοντέλου.

Αν και υπάρχουν αρκετά διαθέσιμα λογισμικά τα οποία καλύπτουν επί μέρους ή και εξ ολοκλήρου τα παραπάνω στάδια φωτογραμμετρικής επεξεργασίας, για τις απαιτήσεις της ψηφιακής τεκμηρίωσης του διαθέσιμου σκελετικού υλικού επιλέχθηκε το Photoscan (έκδοση Pro v1.4, Agisoft LLC, Russia) της Agisoft, το οποίο καλύπτει ολόκληρη τη ροή επεξεργασίας και επιπλέον δίνει συγκεκριμένες δυνατότητες και ρυθμίσεις που επιτρέπουν τη βελτιστοποίηση της ποιότητας του τελικού παραγόμενου τριδιάστατου μοντέλου αλλά και υψηλή αποδοτικότητα τόσο στην ευκολία χειρισμού όσο και στην αξιοποίηση των πόρων των διαθέσιμων Η/Υ. Το διάγραμμα ροής στην παρακάτω εικόνα δίνει μία πιο λεπτομερή περιγραφή της ροής που ακολουθήθηκε από τις αρχικές φωτογραφίες κάθε οστού έως το τριδιάστατο μοντέλο.



Εικόνα 1.2: Διάγραμμα ροής φωτογραμμετρικής επεξεργασίας

Ένα από τα σημαντικά πλεονεκτήματα του Photoscan είναι η δυνατότητα επεξεργασίας παρτίδας (batch processing), η οποία καθιστά την χρήση του πολύ αποδοτική καθώς επιτρέπει να αξιοποιούνται οι Η/Υ αδιάκοπα. Να σημειωθεί ότι ο συνολικός χρόνος (όλα τα στάδια) που απαιτείται χειρισμός από το χρήστη δεν υπερβαίνει τα δέκα λεπτά για κάθε οστό, ενώ ο συνολικός χρόνος επεξεργασίας ξεπερνά τις δύο ώρες. Κατ' αυτόν τον τρόπο, οι φωτογραφίες από κάθε προσανατολισμό εισήχθησαν σε ξεχωριστά σύνολα (chunks), δύο για κάθε οστό. Για κάθε σύνολο έγινε η αρχική φωτογραμμετρική επεξεργασία για τη δημιουργία αραιού νέφους σημείων (sparse

point cloud) και κατόπιν αξιοποιήθηκε η δυνατότητα του Photoscan για βαθμιαία επιλογή βέλτιστων σημείων (gradual selection) με βάση τις παραμέτρους υπολογισμού κάθε σημείου. Οι παράμετροι αυτοί αφορούν στο σφάλμα υπολογισμού (projection accuracy) και στην αβεβαιότητα υπολογισμού (reconstruction uncertainty) της θέσης κάθε σημείου στο νέφος. Αφαιρώντας τα σημεία των οποίων η ακρίβεια υπολογισμού ξεπερνούσε τις 8 μονάδες (εσωτερικές παράμετροι του Agisoft για τις οποίες δεν αναφέρονται συγκεκριμένες μονάδες μέτρησης, απλά η επιλογή του συγκεκριμένου ορίου έγινε κατόπιν δοκιμών και εμπειρικών συμβουλών άλλων χρηστών σε διαδικτυακές συζητήσεις για το συγκεκριμένο ζήτημα), γινόταν επανευθυγράμμιση των θέσεων λήψης των φωτογραφιών (camera realignment) και επανυπολογισμός των αραιού νέφους με τα εναπομείναντα σημεία.

Μετάπειτα, και για κάθε σύνολο ξεχωριστά, υπολογίστηκε το πυκνό νέφος τριδιάστατων σημείων (dense point cloud) σε υψηλή ποιότητα ανάλυσης, το οποίο είναι και το πιο δαπανηρό επεξεργαστικά κομμάτι όλης της διαδικασίας. Το πυκνό νέφος σημείων αποτελεί μία λεπτομερή τριδιάστατη απεικόνιση του οστού καθώς και του περιβάλλοντα χώρου και στην υψηλή ποιότητα ανάλυσης (ρύθμιση στο Agisoft) κάθε πυκνό νέφος αποτελείται από 6-7εκ. σημεία για κάθε προσανατολισμό. Στη συνέχεια απομακρύνθηκαν τα σημεία που δεν αντιστοιχούν στο οστό, μία διαδικασία μερικών λεπτών κατά την οποία απαιτείται χειροκίνητη επεξεργασία και αποτελεί ένα ακόμα πλεονέκτημα του Photoscan, το οποίο επιτρέπει εύκολους και ακριβείς χειρισμούς μέσα από το γραφικό περιβάλλον ενιαίας φωτογραμμετρικής ροής. Μετά τον καθαρισμό, δημιουργήθηκε τριγωνικό πλέγμα (mesh) με επιλογή μεσαίου αριθμού τριγώνων (περί τα 50000) και υφή (texture) με ανάλυση 4096x4096 pixels για τα επιμέρους πυκνά νέφη σημείων κάθε προσανατολισμού των οστών. Αυτό έγινε για τον ενδιάμεσο υπολογισμό μάσκας για κάθε φωτογραφία προκειμένου να ευθυγραμμιστούν αυτόματα τα ζεύγη των διαφορετικών προσανατολισμών ανά οστό αποκλειστικά από τα κοινά σημεία πάνω στα οστά και όχι από το παρασκήνιο (περιβάλλοντα χώρο). Σημειωτέον, χωρίς αυτό το βήμα, η αυτόματη ευθυγράμμιση θα κατέληγε στη ευθυγράμμιση της επίπεδης βάσης πάνω στην οποία στηρίζονται τα οστά με αποτέλεσμα τα δύο πυκνά νέφη (από κάθε προσανατολισμό) να είναι ανάποδα. Κατ' αυτόν τον τρόπο, χρησιμοποιήθηκε η δυνατότητα αυτόματης ευθυγράμμισης (point based alignment) με τη χρήση "μάσκας" (limit tie points by mask) στις φωτογραφίες, οι οποίες μάσκες υπολογίστηκαν αυτόματα από τα επιμέρους μοντέλα τριδιάστατου πλέγματος με την αντίστοιχη εντολή (import masks from models).

Μετά την αυτόματη ευθυγράμμιση, πραγματοποιήθηκε και χειροκίνητη ευθυγράμμιση (marker based alignment) ώστε να περιοριστεί το μέγιστο σφάλμα σε λιγότερο από 1/10 του χιλιοστού. Σε πολλές περιπτώσεις (ανάλογα με την μορφολογία του αντικειμένου αλλά και την υφή επιφανείας

του) η αυτόματη ευθυγράμμιση είναι αρκετή για να δώσει αποτέλεσμα που να πληρεί τις παραπάνω προδιαγραφές μέγιστου σφάλματος αλλά όχι πάντα. Ακόμα πιο σημαντικός παράγοντας που καθιστά αναγκαίο το στάδιο της χειροκίνητης ευθυγράμμισης είναι η επιβεβαίωση του σφάλματος. Κατά την αυτόματη ευθυγράμμιση, το Agisoft επιστρέφει την τιμή μεσοσταθμικού σφάλματος από όλα τα σημεία που χρησιμοποιήθηκαν και παρόλο που γνωρίζουμε τον αριθμό τους, δεν έχουμε περαιτέρω πληροφορίες για την εκτίμηση του μέγιστου σφάλματος. Αυτό συνεπάγεται ότι σε μία ευθυγράμμιση με μικρό μεσοσταθμικό σφάλμα μπορεί να υπάρχουν συγκεκριμένες περιοχές του οστού που το σφάλμα να είναι κατά πολύ μεγαλύτερο, ενώ με την χειροκίνητη επιλογή κατά την οποία συνήθως επιλέγονται τα κατά δύναμη πιο απομακρυσμένα σημεία πάνω στο οστό, το μεσοσταθμικό σφάλμα που επιστρέφει το Agisoft αφορά τις μέγιστες διαστάσεις του οστού και κατ' επέκταση τη μέγιστη απόκλιση που μπορεί να υπάρχει ως σφάλμα ευθυγράμμισης. Εν συνεχεία έγινε συνένωση (merge chunks) των πυκνών νεφών των ζευγών διαφορετικών προσανατολισμών σε ένα καινούριο σύνολο (chunk) ανά οστό. Ακολούθως, δημιουργήθηκαν, για κάθε καινούριο συνενωμένο πυκνό νέφος σημείων (merged chunk), το τελικό τριγωνικό πλέγμα (mesh) με επιλογή προκαθορισμένου αριθμού κορυφών και τριγώνων για κάθε οστό και η αντίστοιχη υφή επιφανείας (texture) με ανάλυση 4096x4096 ή 8192x8192 pixels. Τέλος, έγινε η εξαγωγή των τελικών μοντέλων σε πρότυπο Alias/Wavefront OBJ file format.

Τα τριδιάστατα μοντέλα που παράγονται με την προαναφερθείσα μεθοδολογία φωτογραμμετρικής επεξεργασίας, να μεν είναι αντιπροσωπευτικά των πραγματικών οστών ως προς την γεωμετρία τους και τη φωτορεαλιστική απεικόνιση της επιφανειακής υφής του, ωστόσο οι διαστάσεις τους δεν ανταποκρίνονται στις πραγματικές διαστάσεις των οστών. Ως εκ τούτου, απαιτείται ένα επιπλέον βήμα για τον προσδιορισμό της κλίμακας του μοντέλου και τη μετατροπή του (scaling) σε κατάλληλες μονάδες αντιπροσωπευτικές των πραγματικών διαστάσεων των αντίστοιχων οστών. Όλα τα τριδιάστατα μοντέλα μετατράπηκαν με κατάλληλο συντελεστή κλίμακας στις πραγματικές τους διαστάσεις με βασική μονάδα μήκους το χιλιοστό (mm) με επαναληπτικές μετρήσεις και ανάλογα με τον τύπο οστού. Έτσι, λοιπόν, τα κρανία και οι κάτω γνάθοι μετατράπηκαν σε πραγματικές διαστάσεις βάσει τριών ζευγών ομόλογων ψηφιακών (digital) και χειροκίνητων (manual) μετρήσεων μεταξύ δυνητικά απομακρυσμένων σημείων, τα οποία μπορούσαν να διακριθούν με ακρίβεια τόσο στο ψηφιακό μοντέλο όσο και στο πραγματικό οστό. Αυτές οι ψηφιακές μετρήσεις έγιναν με τη χρήση του ελεύθερου λογισμικού ανοικτού κώδικα MeshLab και οι χειροκίνητες χρησιμοποιώντας παχύμετρο με ρολόι μέγιστου μήκους 150mm, μάρκας Mitutoyo, με ορθότητα $\pm 0,03\text{mm}$ και βαθμονόμηση 0,02mm. Η αλλαγή της κλίμακας του κάθε μοντέλου στις πραγματικές διαστάσεις έγινε επίσης με τη χρήση του λογισμικού MeshLab

(Cignoni et al., 2008) ενώ ο συντελεστής κλίμακας (scaling factor) υπολογίστηκε ως η μέση τιμή των λόγων των μετρήσεων ανά ζεύγη, όπως περιγράφεται από τον παρακάτω τύπο.

$$Sf = \frac{\left(\frac{m_1}{d_1} + \frac{m_2}{d_2} + \frac{m_3}{d_3}\right)}{3}, \text{ όπου } m: \text{ η αναλογική μέτρηση και } d: \text{ η αντίστοιχη ψηφιακή.}$$

Από την άλλη, τα τριδιάστατα ψηφιακά μοντέλα των βραχιόνιων, μηριαίων, κνημών, και ανώνυμων οστών μετατράπηκαν στις πραγματικές τους διαστάσεις με βάση την επαναληπτική μέτρηση του μέγιστου μήκους τους. Το μέγιστο μήκος στα ψηφιακά μοντέλα υπολογίστηκε αυτόματα με εξειδικευμένο πρόγραμμα γραμμένο σε GNU Octave (Eaton et al., 2019) και C++, το οποίο πραγματοποιούσε και την κλίμακα (scaling) σε πραγματικές διαστάσεις με βάση τις ομόλογες μετρήσεις στα πραγματικά οστά, οι οποίες έγιναν με ψηφιακό παχύμετρο μέγιστου μήκους 600mm, μάρκας DIGITRONIC - Moore & Wright 150 DDL Series, με ορθότητα $\pm 0,02\text{mm}$ και βαθμονόμηση $0,01\text{mm}$. Για λόγους επιβεβαίωσης της ορθής ομολογίας των σημείων μέτρησης μεταξύ της χειροκίνητης μέτρησης (που αντιστοιχούν στο μέγιστο μήκος του εκάστοτε οστού) και του αυτόματου υπολογισμού τους από το τριδιάστατο μοντέλο, τα εκάστοτε σημεία απεικονίζονταν πάνω στα ψηφιακά μοντέλα με τη βοήθεια του λογισμικού MeshLab.

1.3. Λάντζα

Η τριδιάστατη φωτογραμμετρία ως μέθοδος ψηφιοποίησης επιλέχθηκε καθώς είναι από τις πλέον οικονομικά προσιτές μεθόδους και συνάμα η πλέον αξιόπιστη ως προς το συνολικό τελικό αποτέλεσμα του παραγόμενου τριδιάστατου μοντέλου, τόσο ως προς την ακρίβεια της γεωμετρικής αναπαράστασης αλλά και την πιστότητα απεικόνισης της υφής επιφάνειας. Μολονότι η ψηφιακή τεκμηρίωση ενός οστού με τη μέθοδο της τριδιάστατης φωτογραμμετρίας, όπως περιγράφηκε παραπάνω, είναι μία σχετικά εύκολη και γρήγορη διαδικασία, καθώς ο συνολικός χρόνος για τους απαιτούμενους χειρισμούς δεν ξεπερνάει τα 25-30 λεπτά, η ψηφιοποίηση δέκα διαφορετικών οστών (ανά άτομο) από μία ολόκληρη συλλογή 250 ατόμων αποτελεί μία τεράστια πρόκληση, στην οποία υπεισέρχονται σημαντικά ζητήματα διαχείρισης πόρων. Από το σύνολο της σύγχρονης σκελετικής συλλογής αναφοράς, ψηφιοποιήθηκαν όλα τα διαθέσιμα οστά μόνο από τα ενήλικα άτομα. Αυτό είχε ως αποτέλεσμα τη δημιουργία 2117 ψηφιακών τριδιάστατων μοντέλων, των οποίων κάποια πληροφοριακά στοιχεία ανά οστό συνοψίζονται στον παρακάτω πίνακα.

Πίνακας 1.1: Σύνοψη τριδιάστατων ψηφιακών μοντέλων.

Οστό	Αριθμός τριγώνων	Ανάλυση υφής επιφανείας	Τυπικό μέγεθος μοντέλου*	Αριθμός 3Δ μοντέλων	Συνολικό μέγεθος αρχείων
Κρανίο	500.000	8192x8192	58 MB	223	12,8 GB
Κάτω γνάθος	200.000	4096x4096	20 MB	205	4,1 GB
Βραχιόνιο	200.000	8192x8192	24 MB	427	9,9 GB
Ανώνυμο	400.000	8192x8192	47 MB	431	20,0 GB
Μηριαίο	300.000	8192x8192	35 MB	405	13,8 GB
Κνήμη	300.000	8192x8192	34 MB	426	14,2 GB

* οι τιμές αφορούν μέγεθος στο δίσκο και περιλαμβάνουν και τα τρία αρχεία (obj: αρχείο γεωμετρίας σε ASCII format, mtl: αρχείο βιβλιοθήκης πρότυπων υλικών, jpg: αρχείο εικόνας υφής επιφανείας) τα οποία απαιτούνται για τριδιάστατα μοντέλα πλέγματος με υφή επιφανείας κατά το πρότυπο Alias/Wavefront OBJ

Είναι σημαντικό να σημειωθεί ότι η ψηφιακή τεκμηρίωση στο σύνολό της είναι αποτέλεσμα ομαδικής εργασίας πέντε ανθρώπων, η οποία αφενός μεν έγινε υπό την καθοδήγησή μου αλλά και την συνεχή επίβλεψη καθ' όλα τα στάδια της φωτογραμμετρικής ροής, όπως προαναφέρθηκαν, αφετέρου δε η συμβολή των Ελισσάβετ Γκανιάτσου, Κατερίνας Χρηστάκη, Κωνσταντίνας Αθανασοπούλου και Νεφέλης Γαρούφη τόσο κατά την φωτογραμμετρική επεξεργασία στο Photoscan αλλά ιδιαιτέρως κατά το στάδιο της φωτογράφισης ήταν καθοριστική για την ολοκλήρωση της ψηφιακής τεκμηρίωσης σε εύλογο χρονικό διάστημα. Ενδεικτικό του όγκου δουλειάς είναι ότι κατά τη διαδικασία της φωτογράφισης ελήφθησαν σχεδόν 350.000 φωτογραφίες, από τις οποίες τελικά αξιοποιήθηκαν οι 288.685 για την φωτογραμμετρική επεξεργασία των παραπάνω οστών. Ως αποτέλεσμα αυτής της ομαδικής προσπάθειας, η ψηφιακή τεκμηρίωση όλου του διαθέσιμου σκελετικού υλικού ολοκληρώθηκε σε κάτι λιγότερο από 18 μήνες, ενώ κατ' εκτίμηση χρειάστηκαν περίπου 1500 ανθρωποώρες εργασίας και γύρω στις 6500 ώρες αυτοματοποιημένης φωτογραμμετρικής επεξεργασίας, η οποία μοιράστηκε σε τρία μηχανήματα (2 H/Y: 6-core FX6300, Radeon RX550-2GB DDR5, 16GB RAM, 1 H/Y: 8-core FX8379, Radeon R9 280X-3GB DDR5, 32GB RAM). Κλείνοντας αυτό το κεφάλαιο, αξίζει να αναφερθεί ότι όλος ο εξοπλισμός που χρησιμοποιήθηκε (ιδιοκατασκευή περιστροφικής βάσης, ψηφιακή κάμερα και λοιπός εξοπλισμός, καθώς και οι τρεις H/Y) αποκτήθηκε με ίδια μέσα (κόστος αγοράς, χρόνος κατασκευής κ.α.) με μοναδικές εξαιρέσεις τις δύο κάρτες γραφικών (Radeon RX550-2GB DDR5), οι οποίες ήταν δωρεά από την Δρ. Ευθυμία Νικήτα, και το ψηφιακό παχύμετρο μεγάλου μήκους (DIGITRONIC - Moore & Wright 150 DDL Series), το οποίο αγοράστηκε από το Μουσείο Ζωολογίας του ΕΚΠΑ και παραχωρήθηκε για τις ανάγκες υπολογισμού της πραγματικής κλίμακας των τριδιάστατων μοντέλων από το διευθυντή του μουσείου, κ. Καθ, Ευστράτιο Βαλάκο.

2. Σύνοψη ερευνητικού έργου

Η ψηφιακή τεκμηρίωση της σύγχρονης σκελετικής συλλογής αναφοράς έδωσε τη δυνατότητα διερεύνησης μίας σειράς σημαντικών ερωτημάτων για το σύγχρονο ελληνικό πληθυσμό καθώς και την ανάπτυξη καινοτόμων μεθόδων και εργαλείων για την ανάκτηση βιολογικής πληροφορίας από τα οστά, οι οποίες υπερβαίνουν το παρόν σκελετικό δείγμα και τον ελληνικό πληθυσμό. Το παρόν κεφάλαιο παρουσιάζει μία σύντομη σύνοψη των ερευνητικών εργασιών που υλοποιήθηκαν στο πλαίσιο του διδακτορικού, οι οποίες παρουσιάζονται αναλυτικότερα στα επόμενα κεφάλαια. Η σειρά παρουσίασης ακολουθεί το χρονοδιάγραμμα υλοποίησής τους, η οποία οποία ξεκίνησε σχεδόν ταυτόχρονα και έγινε παράλληλα με τη ψηφιακή τεκμηρίωση του σκελετικού υλικού. Για αυτό το λόγο οι πρώτες εργασίες έγιναν με υφιστάμενα δεδομένα (συντεταγμένες αναφορικών σημείων του κρανίου) ή με ενδιάμεσα δεδομένα, που προέκυψαν κατά τη διαδικασία της ψηφιοποίησης (φωτογραφίες μηριαίου), και αξιοποιήθηκαν κατάλληλα για να απαντηθούν σύγχρονα ερωτήματα που αφορούν στη μορφολογία του σκελετού. Επίσης, κατά την ψηφιακή τεκμηρίωση δόθηκε σχετική προτεραιότητα ανά οστό (σε σειρά: κρανίο, κάτω γνάθος, μηριαίο, βραχιόνιο, κνήμη, ανώνυμο) με το στόχο με την αποδοτικότερη (πρακτικά) χρήση του παραγόμενου ψηφιοποιημένου υλικού για τους επί μέρους ερευνητικούς σκοπούς χωρίς όμως να επιβαρύνεται (χρονικά) η διαδικασία της ψηφιοποίησης. Κατ' αυτόν τον τρόπο κατέστη δυνατή η αποτελεσματικότερη αξιοποίηση του υλικού καθ' όλη τη διάρκεια του διδακτορικού και αυτός ο σχεδιασμός αποτέλεσε και τη βάση για τη σειρά με την οποία πραγματοποιήθηκαν οι ερευνητικές εργασίες που παρουσιάζονται παρακάτω.

1. Η ανάλυση 80 αναφορικών σημείων του κρανίου με στόχο την εύρεση των πιο φυλετικά διμορφικών μετρικών χαρακτήρων (αποστάσεις και γωνίες) και τη δημιουργία διακριτικών εξισώσεων για την ταυτοποίηση του φύλου αγνώστου ατόμου από τον ελληνικό πληθυσμό. Στόχος ήταν η διεξοδική στατιστική ανάλυση όλων των δυνατών συνδυασμών αποστάσεων (3.160) και γωνιών (246.480) που προκύπτουν από τα προαναφερθέντα ανατομικά σημεία, η οποία και πραγματοποιήθηκε προγραμματιστικά σε κώδικα GNU Octave. Η επικυρωμένη επιτυχής ταξινόμηση ξεπέρασε το 95% (LOOCV 95,2%) για πολυμεταβλητή εξίσωση με τη χρήση μόλις τριών μετρικών χαρακτήρων (1 γωνία, 2 αποστάσεις) (Bertsatos, Parageorgoroulou, et al., 2018). Πέραν της εξαιρετικής ακρίβειας ορθής ταξινόμησης φύλου, η παρούσα εργασία έδειξε ότι ένας σημαντικός αριθμός μετρικών χαρακτήρων μεταξύ των οποίων και η πιο φυλετικά διμορφική μέτρηση στον ελληνικό πληθυσμό (απόσταση μεταξύ δεξιού *onale mediale* και δεξιού *mastoidale*) δίνουν καλύτερα αποτελέσματα ταξινόμησης φύλου σε σχέση με τις κλασσικές κρανιομετρικές μετρήσεις.

2. Η φωτογραμμετρική ανάλυση της μορφολογίας του βοθρίου της κεφαλής του μηριαίου (*fovea capitis*) σε σχέση με το φύλο και την ηλικία με τη χρήση μεταβλητών μεγέθους και σχήματος μεταξύ των οποίων και η καινοτόμος μεταβλητή *ArPerIndex*. Συνοπτικά, ο φυλετικός διμορφισμός στο βοθρίο της κεφαλής του μηριαίου αφορά στις μεταβλητές μεγέθους, ενώ η επιρροή της ηλικίας του ατόμου εμφανίζεται στο σχήμα του. Ειδικότερα, το εμβαδόν και η μέγιστη περίμετρος του βοθρίου της κεφαλής του μηριαίου είναι σαφώς μεγαλύτερα στα αρσενικά από ότι στα θηλυκά άτομα, ενώ ταυτόχρονα το περίγραμμά του γίνεται πιο ακανόνιστο με την αύξηση της ηλικίας (Bertsatos, Chovaloroulou, et al., 2018). Πέραν της μορφολογικής ανάλυσης του βοθρίου της κεφαλής του μηριαίου, η παρούσα εργασία αποτέλεσε και την αφορμή για την δημιουργία της μεταβλητής *ArPerIndex*, της οποίας η εξαιρετική δυνατότητα ως μορφολογικός δείκτης περιγράμματος επιβεβαιώθηκε και από μετέπειτα οστεολογικές έρευνες σε διαφορετικά ερωτήματα, όπως η εκτίμηση φύλου από μακρά οστά (Bertsatos, Garoufi, et al., 2020) καθώς η μονατομική ταξινόμηση σκελετικού υλικού από μαζικές ταφές (Bertsatos & Chovaloroulou, 2020).
3. Αξιοποιώντας το εύρος στη χρονολογία γέννησης των ατόμων της σύγχρονης σκελετικής συλλογής αναφοράς αλλά και τα μετρικά δεδομένα (μέγιστο μήκος) που προέκυψαν κατά την μοντελοποίηση των μηριαίων οστών, έγινε μία ανάλυση των διαχρονικών τάσεων στο ανάστημα του ελληνικού πληθυσμού και συσχέτιση με τις πολιτικοοικονομικές συνθήκες που επικράτησαν στη χώρα ανά χρονικές περιόδους. Όπως αναμενόταν, το ανάστημα κατά τη διάρκεια της περιόδου μεταξύ 1879 και 1965 αυξήθηκε τόσο στα αρσενικά όσο και στα θηλυκά άτομα. Παρόλα αυτά, η ανάλυση διαχρονικών τάσεων μεταξύ διαφορετικών περιόδων, έδειξε ότι τα αρσενικά άτομα υπέστησαν μία μικρή μείωση στο ανάστημα κατά τη διάρκεια των πολέμων και της μεσοπολεμικής περιόδου μεταξύ 1912 και 1950. Αυτή η μείωση, μολονότι μη στατιστικά σημαντική, υποδεικνύει ότι οι οικονομικές και διατροφικές στερήσεις ως απόρροια των πολέμων και της μεγάλης οικονομικής ύφεσης του μεσοπολέμου (κραχ 1929) επηρέασαν εμφανώς το ανάστημα του ελληνικού πληθυσμού, τουλάχιστον των αντρών που γεννήθηκαν κατά τη διάρκεια αυτής της περιόδου (Bertsatos & Chovaloroulou, 2018a).
4. Πρωταρχικό ζήτημα της τριδιάστατης μοντελοποίησης των δέκα (10) οστών που ψηφιοποιήθηκαν από τη σύγχρονη σκελετική συλλογή αναφοράς είναι η ακρίβεια και η ορθότητα των παραχθέντων τριδιάστατων μοντέλων. Αυτό αφορά τόσο τη γεωμετρία τους όσο και την εικονική δυνατότητα που προσφέρουν στο να ταυτοποιηθούν συγκεκριμένα ανατομικά σημεία. Σε αυτό το πλαίσιο χρησιμοποιήθηκαν 50 τριδιάστατα μοντέλα κρανίων

από τα οποία συλλέχθηκαν 56 αναφορικά σημεία και τα οποία συγκρίθηκαν με αντίστοιχα δεδομένα συντεταγμένων ψηφιοποιημένα με Microscribe-3DX από προηγούμενη διδακτορική διατριβή (Χοβαλοπούλου, 2014) καθώς και με μετρήσεις αποστάσεων (σε αντιστοιχία με τα συλλεχθέντα αναφορικά σημεία) που έγιναν με τη χρήση παχύμετρου στα συγκεκριμένα κρανία. Έγινε ανάλυση του τεχνικού σφάλματος μέτρησης (TEM) μεταξύ των διαφορετικών μεθόδων μέτρησης/ψηφιοποίησης (inter-method), μεταξύ διαφορετικών τύπων αναφορικών σημείων (types I, II, III), μεταξύ επαναλαμβανόμενων μετρήσεων του ίδιου ερευνητή (intra-observer) καθώς και διαφορετικών ερευνητών (inter-observer) ως προς τη τριδιάστατη διάταξη αναφορικών σημείων (3DGM) καθώς και ως προς τις αποστάσεις μεταξύ αναφορικών σημείων (ILDs). Σύμφωνα με τα αποτελέσματα, η ακρίβεια της φωτογραμμετρικής μοντελοποίησης των οστών κυμαίνεται στα 0,2 χιλιοστόμετρα (mm) και είναι εφάμιλλη του ψηφιοποιητή ακίδας, ενώ παράλληλα τα αναφορικά σημεία τύπου I έχουν υψηλότερη ακρίβεια και με τις δύο μεθόδους σε σχέση με τα τύπου II και III. Παρόλα αυτά, τα φωτογραμμετρικά τριδιάστατα μοντέλα προσφέρουν μικρότερο σφάλμα μεταξύ επαναλαμβανόμενων μετρήσεων καθιστώντας τα πιο αξιόπιστα. Επιπλέον, οι μετρήσεις αποστάσεων μεταξύ αναφορικών σημείων είναι λιγότερο ευάλωτες σε σφάλμα μέτρησης σε σύγκριση με τις τριδιάστατες διατάξεις αναφορικών σημείων. Επίσης, οι μετρήσεις αποστάσεων έχουν παρόμοιο σχετικό σφάλμα μέτρησης (περίπου 1,5%) μεταξύ όλων των μεθόδων που εξετάστηκαν, κάτι το οποίο επιτρέπει την συνάθροιση και συνεπεξεργασία μετρικών δεδομένων τόσο από διαφορετικούς παρατηρητές όσο και με διαφορετικές μεθόδους μέτρησης. Το ίδιο δεν ισχύει απαραίτητα και για τις τριδιάστατες διατάξεις αναφορικών σημείων όταν αυτές συλλέγονται με διαφορετικές μεθόδους (Bertsatos, Gkaniatsou, et al., 2019).

5. Τα αναφορικά σημεία που συλλέχθηκαν από τα τριδιάστατα μοντέλα των κρανίων της σύγχρονης σκελετικής συλλογής αναφοράς χρησιμοποιήθηκαν για την αξιολόγηση του λογισμικού 3D-ID για την ορθή ταξινόμηση της καταγωγής και ταυτοποίησης φύλου αγνώστων ατόμων με βάση διαφορετικούς σύγχρονους πληθυσμούς αναφοράς από όλο τον κόσμο που περιέχει στη βάση δεδομένων του και στηρίζεται σε ανάλυση τριδιάστατης γεωμετρικής μορφομετρίας. Από τα 158 κρανία που χρησιμοποιήθηκαν από την ψηφιοποιημένη σκελετική συλλογή το ποσοστό επιτυχούς ταυτοποίησης φύλου άγνωστης καταγωγής κυμάνθηκε από 74,1% έως 86,7% ανάλογα με τις επιλογές διάταξης αναφορικών σημείων (χρησιμοποιώντας είτε 19 είτε 25 αναφορικά σημεία), ενώ με γνωστή την καταγωγή (κατατάσσοντας τον ελληνικό πληθυσμό ως νοτιοανατολικό Ευρωπαϊκό) το ποσοστό ανέβηκε σε 89,9%. Η ταξινόμηση καταγωγής του ελληνικού πληθυσμού ως

νοτιοανατολικός Ευρωπαϊκός είχε πενιχρά αποτελέσματα, ενώ η ταξινόμηση των υπό εξέταση δειγμάτων ως Ευρωπαϊκά άγγιξε το 70,7% υποδεικνύοντας τις αδυναμίες του 3D-ID που σχετίζονται τόσο με τη περιορισμένη βάση δεδομένων του όσο και με τους περιορισμούς της τριδιάστατης γεωμετρικής μορφομετρίας για την εκτίμηση πληθυσμιακής καταγωγής (Bertsatos, Christaki, et al., 2019).

6. Η μορφομετρική ανάλυση της κάτω γνάθου από τα διαθέσιμα τριδιάστατα ψηφιακά μοντέλα της σύγχρονης σκελετικής συλλογής αναφοράς με στόχο τη δημιουργία πληθυσμιακά εξειδικευμένων στατιστικών μοντέλων εκτίμησης του φύλου για τον ελληνικό πληθυσμό. Αξιοποιώντας 194 δείγματα ενήλικων ατόμων υπολογίστηκαν είκοσι (20) αποστάσεις και τρεις (3) γωνίες από τις τριδιάστατες συντεταγμένες συγκεκριμένων ανατομικών σημείων και αναλύθηκαν στατιστικά μέσω διαχωριστικής ανάλυσης (DFA) και ανάλυσης της διακύμανσης (ANOVA) για την διερεύνηση της έκφρασης του φυλετικού διμορφισμού. Τα ύψη του κλάδου και της κορωνοειδούς απόφυσης καθώς και το μέγιστο πλάτος της κάτω γνάθου αποτελούν τους πλέον φυλετικά διμορφικούς μετρικούς χαρακτήρες για τον ελληνικό πληθυσμό, ενώ η βέλτιστη διαχωριστική εξίσωση προσδιορισμού του φύλου έδωσε 85,7% κατόπιν διασταυρωμένης επικύρωσης. Ενώ συγκεκριμένα πρότυπα έκφρασης φυλετικού διμορφισμού στην κάτω γνάθο είναι παρόμοια μεταξύ των πληθυσμών, η συγκριτική εφαρμογή διαχωριστικών εξισώσεων από ξένους πληθυσμούς στο ελληνικό δείγμα επαλήθευσε την έντονη πληθυσμιακή εξειδίκευση της έκφρασης του φυλετικού διμορφισμού επισημαίνοντας την αναγκαιότητα κατάλληλων διαχωριστικών εξισώσεων εκτίμησης φύλου για τον ελληνικό πληθυσμό (Bertsatos, Athanasopoulou, et al., 2019).
7. Η αντικατάσταση της ευρέως χρησιμοποιούμενης μεθόδου λήψης δακτυλίων με οδοντιατρική σιλικόνη (Latex Cast Method) για την γεωμετρική ανάλυση της εγκάρσιας τομής της διάφυσης των μακρών οστών με μία αυτόματη υπολογιστική μέθοδο των αντίστοιχων γεωμετρικών παραμέτρων απευθείας από τριδιάστατα ψηφιακά μοντέλα και χωρίς σφάλμα μέτρησης σχετιζόμενο με τον παρατηρητή ή τη μέθοδο. Η καινοτόμος μέθοδος που δημιουργήθηκε είναι απόλυτα ακριβής και με πλήρη επαναληψιμότητα καθώς βασίζεται σε αλγόριθμους υπολογιστικής γεωμετρίας για την εξαγωγή των απαραίτητων γεωμετρικών παραμέτρων. Η καινούρια μέθοδος ανάλυσης της γεωμετρίας των εγκάρσιων τομών της διάφυσης των μακρών οστών έχει υλοποιηθεί σε λογισμικό ανοικτού κώδικα με την ονομασία long-bone-diaphyseal-CSG-Toolkit (Bertsatos, 2018), το οποίο είναι ελεύθερα διαθέσιμο. Το λογισμικό δίνει τη δυνατότητα αυτόματης επεξεργασίας τριδιάστατων

μοντέλων μηριαίων, κνημών καθώς και βραχιόνιων. Η πλήρης εξάλειψη μετρητικού σφάλματος σχετιζόμενο με το χρήστη έχει βασιστεί σε εξειδικευμένους υπολογιστικούς αλγόριθμους βελτιστοποίησης αναφορικών σημείων, τα οποία είναι προαπαιτούμενα για τον υπολογισμό του ανατομικού προσανατολισμού του εκάστοτε ψηφιακού οστού. Το ελεύθερο λογισμικό έχει γραφτεί σε γλώσσες προγραμματισμού GNU Octave και C++, ενώ η τελευταία έκδοση αναβάθμισης (v1.1.0) (Bertsatos, 2020) επιτρέπει τον υπολογισμό και εξαγωγή όλων των παραμέτρων σε περίπου 2 δευτερόλεπτα ανά οστό. Ο συνολικός χρόνος μαζί με τις προπαρασκευαστικές ενέργειες από το χρήστη δεν υπερβαίνει το ένα λεπτό, ενώ η εξαγωγή των αντίστοιχων παραμέτρων με τη μέθοδο λήψης δακτυλίων με οδοντιατρική σιλικόνη απαιτεί τουλάχιστον μία ώρα εντατικής δουλειάς (Bertsatos & Chovalopoulou, 2019). Η καινούρια υπολογιστική μέθοδος έχει ήδη δείξει την αξία της σε μία σειρά από ερευνητικές εργασίες, που είτε μελετούν τις εγκάρσιες τομές της διάφυσης μακρών οστών ως δείκτη δραστηριότητας (Nikita et al., 2019) είτε αξιοποιούν αυτές τις γεωμετρικές παραμέτρους για τον προσδιορισμό φύλου (Bertsatos, Garoufi, et al., 2020) και την μονατομική ταξινόμηση σκελετικού υλικού (Bertsatos & Chovalopoulou, 2020).

8. Στόχος ήταν η δημιουργία ενός αυτοματοποιημένου εργαλείου για τον προσδιορισμό του φύλου από μακρά οστά με τη δυνατότητα εκτίμησης της πιθανότητας σωστής ταυτοποίησης. Χρησιμοποιώντας το λογισμικό long-bone-diaphyseal-CSG-Toolkit για την εξαγωγή γεωμετρικών παραμέτρων εγκάρσιων τομών από τις διαφύσεις του συνόλου των διαθέσιμων μακρών οστών από τη σύγχρονη σκελετική συλλογή αναφοράς, πραγματοποιήθηκε επεξεργασία των δεδομένων με στατιστική διαχωριστική ανάλυση (DFA) καθώς και με τεχνικές μηχανικής εκμάθησης (SVM). Τα αποτελέσματα έδειξαν την χρησιμότητα των διάφορων μεταβλητών της γεωμετρίας της εγκάρσιας τομής από το μηριαίο, την κνήμη και το βραχιόνιο δίνοντας βελτιωμένα ποσοστά επιτυχούς ταξινόμησης (κατόπιν διασταυρωμένης επικύρωσης) σε σχέση με υφιστάμενες μεθόδους. Πιο συγκεκριμένα η επιτυχής εκτίμηση φύλου από το μηριαίο ανήλθε στο 94,8%, από την κνήμη στο 94,7%, ενώ στο βραχιόνιο άγγιξε το 97,3%. Επίσης, η περαιτέρω στατιστική ανάλυση έδειξε ότι η καινούρια προτεινόμενη μέθοδος φυλετικού προσδιορισμού είναι σχετικά ανθεκτική σε μορφολογικές διακυμάνσεις οφειλόμενες σε διαχρονικές τάσεις. Για την προτεινόμενη μέθοδο δημιουργήθηκε ένα ελεύθερο λογισμικό (ανοικτού κώδικα) με ένα απλοποιημένο γραφικό περιβάλλον εργασίας, το οποίο διευκολύνει τον χρήστη μέσα από χρήσιμες επιλογές στον επιτυχή προσδιορισμό του φύλου με γνώμονα και την αυτόματα υπολογιζόμενη πιθανότητα σωστής ταυτοποίησης (Bertsatos, Garoufi, et al., 2020).

9. Διερεύνηση του κατά πόσο τα αυξημένα επίπεδα δραστηριότητας, όπως αυτά υποδεικνύονται από διάφορους σκελετικούς δείκτες δραστηριότητας που σχετίζονται με τις αλλοιώσεις των ενθέσεων και τη γεωμετρία της εγκάρσιας τομής του μηριαίου, έχουν ως αποτέλεσμα αυξημένο ρυθμό εκφυλιστικών αλλοιώσεων στην ηβική σύμφυση και την ωτοειδή επιφάνεια, όπως αυτές εκφράζονται από τους σκελετικούς δείκτες ηλικίας για τις εν λόγω περιοχές του ανώνυμου οστού. Η υπό διερεύνηση συσχέτιση δοκιμάστηκε τόσο με στατιστική συσχέτιση Spearman όσο και με την εφαρμογή γενικευμένων γραμμικών μοντέλων (GLM). Επιπλέον, ελέγχθηκε κατά πόσο διαφοροποιούνται οι δείκτες δραστηριότητας μεταξύ υπερεκτιμημένων και υποεκτιμημένων ως προς την πραγματική ηλικία θανάτου ατόμων με βάση τους αντίστοιχους σκελετικούς δείκτες. Βρέθηκε αρνητική συσχέτιση των σταδίων της σκελετικής ηλικίας με συγκεκριμένες γεωμετρικές παραμέτρους εγκάρσιων τομών, που συνεπάγεται μείωση της σκελετικής ακαμψίας με την αύξηση της ηλικίας, και θετική συσχέτιση με το βαθμό αλλοιώσεων των ενθέσεων, γεγονός που υποδηλώνει ότι τα ηλικιωμένα άτομα παρουσιάζουν πιο έντονες αλλοιώσεις. Η διαφορά στους δείκτες δραστηριότητας μεταξύ υπερεκτιμημένων και υποεκτιμημένων ηλικιακά ατόμων ήταν στατιστικά σημαντική μόνο για τα άτομα άνω των 50 ετών. Η παρούσα μελέτη υπογραμμίζει ότι η επίδραση της δραστηριότητας στους δείκτες ηλικίας της πυέλου δεν είναι έντονη, καθώς βρέθηκε περιορισμένη συσχέτιση μεταξύ της δραστηριότητας και του σκελετικού εκφυλισμού στην ηβική σύμφυση και στην λαγόνια κοιλιακή επιφάνεια. Επιπλέον, τα αποτελέσματά μας υποδηλώνουν ότι η συνεκτίμηση των βαθμού αλλοίωσης των ενθέσεων κατά την αξιολόγηση της σκελετικής ηλικίας μπορεί να παρέχει περαιτέρω πληροφορίες σχετικά με την εκτίμηση ηλικίας-θανάτου σε ηλικιωμένα άτομα (Bertsatos, Choaloroulou, Boskovits, et al., 2020).
10. Η δημιουργία μιας αυτοματοποιημένης υπολογιστικής μεθόδου για την εκτίμηση του φύλου από τριδιάστατα ψηφιακά μοντέλα κρανίων με βάση τα μορφολογικά χαρακτηριστικά που εκφράζουν έντονο φυλετικό διμορφισμό. Η προτεινόμενη μέθοδος αναπτύχθηκε και αξιολογήθηκε χρησιμοποιώντας δύο δείγματα ευρωπαϊκών πληθυσμών, ένα τσεχικό δείγμα αποτελούμενο από 170 κρανία που ανακατασκευάστηκε από ανώνυμες αξονικές τομογραφίες και ένα ελληνικό δείγμα 156 κρανίων από τη σύγχρονη σκελετική συλλογή αναφοράς. Η ακρίβεια επιτυχούς προσδιορισμού του φύλου αξιολογήθηκε υπό το πρίσμα πληθυσμιακής εξειδίκευσης αξιολογώντας τα δύο πληθυσμιακά δείγματα είτε χωριστά είτε ομαδοποιημένα. Η πληθυσμιακά εξειδικευμένη ορθή ταξινόμηση για μεμονωμένα μορφομετρικά χαρακτηριστικά κυμάνθηκε από 78,5 έως 96,7%, ενώ η πληθυσμιακά μη-εξειδικευμένη κυμάνθηκε από 71,7 έως 90,8%. Ο συνδυασμός όλων των διαγνωστικών

χαρακτηριστικών του φύλου στην εκτίμηση φύλου πολλαπλών χαρακτηριστικών απέδωσε σωστή απόδοση ταξινόμησης άνω του 91% για ολόκληρο το δείγμα, ενώ το φύλο των περίπου τριών τετάρτων του δείγματος θα μπορούσε να προσδιοριστεί με ακρίβεια 100% σύμφωνα με την εκτίμηση οπίσθιας πιθανότητας. Η προτεινόμενη μέθοδος, ελεύθερα διαθέσιμη με το πρόγραμμα Η/Υ ανοικτού κώδικα skullanalyzer και της βοηθητικής συνάρτησης estimate_sex.m για περιβάλλον GNU Octave, παρέχει έναν αποτελεσματικό και αξιόπιστο τρόπο για την εκτίμηση του φύλου από κρανιακά κατάλοιπα και προσφέρει σημαντικά πλεονεκτήματα σε σχέση με τις υπάρχουσες μεθόδους (Bertsatos, Choalopoulou, Brůžek, et al., 2020).

3. Μορφομετρική διερεύνηση του φυλετικού διμορφισμού του κρανίου

Η ταυτοποίηση του φύλου αγνώστου ατόμου είναι από τα κύρια καθήκοντα κατά την εξέταση ανθρώπινων σκελετικών υπολειμμάτων. Με την πάροδο των ετών, έχουν προταθεί διάφορες ποιοτικές και ποσοτικές μέθοδοι βασιζόμενες σε κρανιακά και μετακρανιακά χαρακτηριστικά (Krishan et al., 2016; Phenice, 1969; Spradley & Jantz, 2011; Walker, 2008), οι οποίες εμφανίζουν ποικίλα ποσοστά επιτυχούς ταξινόμησης. Ωστόσο, τα υψηλότερα ποσοστά επιτυχούς ταξινόμησης φύλου τα εμφανίζουν μέθοδοι που σχετίζονται με τη λεκάνη και το κρανίο (Krishan et al., 2016). Σε γενικές γραμμές, οι ποιοτικές μέθοδοι εφαρμόζονται ευκολότερα, αλλά βασίζονται σε μεγάλο βαθμό στην εμπειρία του παρατηρητή. Κατά συνέπεια τείνουν να είναι πιο υποκειμενικές και επιρρεπείς σε σφάλματα αναπαραγωγιμότητας (Wilczak et al., 2017). Αντιθέτως, οι ποσοτικές μέθοδοι στηρίζονται σε μετρήσεις, οι οποίες μπορεί να είναι πιο χρονοβόρες και συχνά απαιτούν εξειδικευμένο εξοπλισμό (Oikonomopoulou et al., 2017), ωστόσο, είναι λιγότερο επιρρεπείς σε παρατηρησιακή μεροληψία και συχνά παράγουν πιο ακριβή αποτελέσματα (Gómez-Valdés et al., 2012; Wilczak et al., 2017). Τα τελευταία χρόνια οι ποσοτικές μέθοδοι προχώρησαν σε πιο περίπλοκες τεχνικές χρησιμοποιώντας γεωμετρική μορφομετρία, οι οποίες επιτρέπουν πιο προηγμένη ανάλυση της διακύμανσης του σχήματος από τις παραδοσιακές μεθόδους μέτρησης (Bigoni et al., 2010; Gonzalez et al., 2009; Krishan et al., 2016).

Η χρήση αναφορικών σημείων στη γεωμετρική μορφομετρία για την ανάλυση της διακύμανσης του σχήματος του ανθρώπινου κρανίου σε σχέση με το φύλο έχει αξιολογηθεί από προηγούμενες έρευνες (Bigoni et al., 2010; Chovalopoulou et al., 2013, 2016a, 2016b; Chovalopoulou, Papageorgopoulou, et al., 2017; Manthey et al., 2017). Ωστόσο, η Γενικευμένη Ανάλυση Προκρούστη (GPA) και η Ανάλυση Κύριων Συνιστωσών (PCA), οι οποίες αποτελούν την συνηθισμένη διαδικασία ανάλυσης τέτοιων δεδομένων, παράγουν αποτελέσματα που είναι δύσκολο να εφαρμοστούν σε δικανικές περιπτώσεις. Εξαιρεση αποτελούν οι περιπτώσεις όπου δεν υπάρχουν ελλείποντα δεδομένα και τα δεδομένα αναφοράς είναι επίσης διαθέσιμα στους δικαστικούς ανθρωπολόγους ή βιοαρχαιολόγους, που καλούνται να μελετήσουν ένα απομονωμένο δείγμα. Αυτό έχει οδηγήσει σε κάποια κριτική ενάντια στην ευχρηστία και αποτελεσματικότητα της γεωμετρικής μορφομετρίας στην εκτίμηση του φύλου (Langley et al., 2018; Oikonomopoulou et al., 2017). Αντιθέτως, οι παραδοσιακές ποσοτικές μέθοδοι, οι οποίες είναι απλές στην εφαρμογή, συχνά στηρίζονται σε κρανιακά χαρακτηριστικά, τα οποία είναι γνωστό ότι εμφανίζουν φυλετικό διμορφισμό βάσει οπτικής παρατήρησης (Franklin et al., 2013; Mahakkanukrauh et al., 2015). Ως αποτέλεσμα, ένα σημαντικό ποσοστό φυλετικής ποικιλομορφίας, το οποίο καταγράφεται μέσω των

GPA και PCA, δεν διερευνάται ούτε χρησιμοποιείται στις διακριτικές εξισώσεις για την ταυτοποίηση του φύλου αγνώστου ατόμου (Krishan et al., 2016). Στόχος της συγκεκριμένης εργασίας ήταν να καλύψει αυτό το χάσμα διενεργώντας διεξοδική στατιστική ανάλυση όλων των δυνατών συνδυασμών αποστάσεων και γωνιών που προκύπτουν από επιλεγμένα κρανιακά αναφορικά σημεία και τη δημιουργία διακριτικών εξισώσεων για την ταυτοποίηση του φύλου αγνώστου ατόμου από τον ελληνικό πληθυσμό.

3.1. Υλικά και μέθοδοι

Για τις ανάγκες της παρούσας εργασίας μελετήθηκαν συνολικά 176 κρανία ενήλικων ατόμων (94 άνδρες και 82 γυναίκες). Οι τρισδιάστατες συντεταγμένες 80 αναφορικών σημείων (12 στο οβελιαίο επίπεδο και 34 αμφίπλευρα) στην εξωτερική επιφάνεια του κρανίου ψηφιοποιήθηκαν με Microscribe 3DX (Immersion Cor, San Jose, California), του οποίου η ονομαστική ακρίβεια είναι $\pm 0,23$ mm. Τα αναφορικά σημεία επιλέχθηκαν με βάση την επαρκή απεικόνιση της γεωμετρίας του κρανίου (Πίνακας 1).

Πίνακας 3.1: Λίστα κρανιακών αναφορικών σημείων

1	Nasion	17	Occipitocondylion anterior *	32	Supra Auricular *
2	Glabella	18	Caroticum mediale *	33	Zygotemporale superior *
3	Bregma	19	Spinale *	34	Jugale *
4	Lambda	20	Ovale mediale *	35	Frontomalare temporale *
5	Opisthocranium	21	Postalverion *	36	Porion *
6	Opisthion	22	Infratemporale *	37	Zygotemporale inferior *
7	Basion	23	Mastoidale *	38	Zygomaxillare *
8	Nasospinale	24	Asterion *	39	Frontomalare orbitale *
9	Hormion	25	Entomion *	40	Supraconchion *
10	Staphylion	26	Supramastoid crest – squamous suture intersection *	41	Maxillofrontale *
11	Staurion	27	Crotaphion *	42	Subconchion *
12	Foramen incisivum	28	Coronale *	43	Ectoconchion *t
13	Foraminolaterale *	29	Sphenion *	44	Infraorbitale *
14	Occipitocondylion mediale *	30	Frontotemporale *	45	Apertion *
15	Occipitocondylion posterior *	31	Landmark x *	46	Maxillonasofrontale *
16	Occipitocondylion laterale *				

* Αμφίπλευρα αναφορικά σημεία, τα οποία ψηφιοποιήθηκαν και από τις δύο πλευρές

Όλες οι διαθέσιμες συντεταγμένες των ογδόντα αναφορικών σημείων αναλύθηκαν σε περιβάλλον GNU Octave. Στοχευμένα λογισμικά προγραμματισμού (scripts), τα οποία είναι διαθέσιμα κατόπιν αιτήματος, υπολόγισαν όλους τους πιθανούς συνδυασμούς αποστάσεων και γωνιών από το σύνολο των αναφορικών σημείων που περιλαμβάνονται στην παρούσα εργασία. Κατόπιν, για κάθε μοναδικό σύνολο 3.160 αποστάσεων και 246.480 γωνιών πραγματοποιήθηκε γραμμική διακριτική ανάλυση για τον υπολογισμό του ποσοστού επιτυχούς ταξινόμησης κάθε μεμονωμένης μεταβλητής (απόσταση ή γωνία). Δεκατρείς αποστάσεις και επτά γωνίες με το υψηλότερο ποσοστό επιτυχούς ταξινόμησης αναλύθηκαν περαιτέρω στο SPSS (IBM SPSS έκδοση 23.0, Armonk, NY). Για τις προαναφερθείσες 20 μεταβλητές υπολογίστηκε και η διασταυρούμενη επικύρωση. Επιπλέον, στο σύνολο των αποστάσεων και των γωνιών εφαρμόστηκε μία σειρά από πολυπαραγοντικές διακριτικές αναλύσεις. Στην πολυπαραγοντική διακριτική ανάλυση χρησιμοποιήθηκε τόσο η κλιμακωτή μέθοδος επιλογής ανεξάρτητων μεταβλητών όσο και η μέθοδος χρήσης όλων των ανεξάρτητων μεταβλητών.

3.2. Αποτελέσματα

Η προσέγγιση που υιοθετήθηκε στην παρούσα εργασία ανέδειξε πολλές φυλετικά διμορφικές απόστασης και γωνίας με αρκετά υψηλή επιτυχή ταξινόμηση. Πιο συγκεκριμένα, σύμφωνα με τη διακριτική ανάλυση, 60 αποστάσεις και 61 γωνίες εμφάνισαν επιτυχή ταξινόμηση υψηλότερη από 82% και 75% αντίστοιχα. Ωστόσο, στην πολυπαραγοντική διακριτική ανάλυση χρησιμοποιήθηκαν οι 13 αποστάσεις με ποσοστό επιτυχούς ταξινόμησης υψηλότερο από 85% και οι 7 γωνίες, οι οποίες εμφάνισαν ποσοστό επιτυχούς ταξινόμησης υψηλότερο από 78%.

Τα προαναφερθέντα ποσοστά επιτυχούς ταξινόμησης υπολογίστηκαν στο GNU Octave. Ωστόσο, τα αποτελέσματα επαληθεύθηκαν από την διακριτική ανάλυση στο SPSS, ενώ ταυτόχρονα υπολογίστηκαν και τα ποσοστά της διασταυρούμενης επικύρωσης. Τα αποτελέσματα της περιγραφικής στατιστικής καθώς και τα ποσοστά επιτυχούς ταξινόμησης από το SPSS για τις συγκεκριμένες 20 μεταβλητές, αναγράφονται στον Πίνακα 2.

Ενώ η υψηλότερη επικυρωμένη επιτυχής ταξινόμηση στο σύνολο του δείγματος για την απόσταση D1 και τη γωνία A1 ήταν 87,4% και 80,6% αντίστοιχα, η πολυπαραγοντική διακριτική ανάλυση απέδωσε ακόμη υψηλότερα ποσοστά. Τα αποτελέσματα της πολυπαραγοντικής διακριτικής ανάλυσης των 20 μεταβλητών τόσο με την κλιμακωτή μέθοδο επιλογής ανεξάρτητων μεταβλητών όσο και τη μέθοδο χρήσης όλων των ανεξάρτητων μεταβλητών συνοψίζονται στον Πίνακα 3.

Πίνακας 3.2: Αποτελέσματα περιγραφικής στατιστικής και ποσοστά επιτυχούς ταξινόμησης

Από. ^α	Αναφορικό σημείο		Μέσος όρος ^β	95% διάστημα εμπιστοσύνης		Ποσοστό επιτυχούς ταξινόμησης %			
	Αναφορικό σημείο 1	Αναφορικό σημείο 2		Κατώτερο ^β	Ανώτερο ^β	Συνολικά ^γ	Αρσενικά ^γ	Θηλυκά ^γ	
D1	Ovale mediale Δεξί	Mastoidale Δεξί	53,3 / 47,5	52,7 / 46,9	53,8 / 48,1	87,4	87,1	87,7	
D2	Mastoidale Αριστερό	Frontomalare temporale Δεξί	142,5 / 134,3	141,5 / 133,3	143,3 / 135,3	86,7 / 86,1	88,0	85,2 / 84,0	
D3	Sphenion Αριστερό	Mastoidale Δεξί	141,3 / 131,5	140,2 / 130,6	142,4 / 132,5	86,6	85,7	87,7	
D4	Mastoidale Αριστερό	Coronale Δεξί	159,2 / 149,0	158,1 / 147,9	160,4 / 150,1	86,2	83,7	89,0	
D5	Mastoidale Αριστερό	Jugale Δεξί	132,5 / 125,0	131,6 / 124,0	133,4 / 125,9	86,1 / 85,5	89,1 / 88,0	82,7	
D6	Mastoidale Αριστερό	Sphenion Δεξί	141,6 / 132,5	140,6 / 131,4	142,7 / 133,5	86,0	87,8	84,1	
D7	Supraconchion Αριστερό	Mastoidale Δεξί	139,0 / 129,5	137,9 / 128,5	140,0 / 130,5	86,0	85,9	86,3	
D8	Frontotemporale Αριστερό	Mastoidale Δεξί	148,1 / 138,5	147,0 / 137,5	149,2 / 139,5	85,6	86,0	85,2	
D9	Ectoconchion Αριστερό	Mastoidale Δεξί	135,5 / 127,2	134,5 / 126,3	136,4 / 128,2	85,5 / 84,3	85,9	85,0 / 82,5	
D10	Zygotemporale superior Αριστερό	Mastoidale Δεξί	132,4 / 123,6	131,4 / 122,7	133,3 / 124,5	85,4	84,6	86,3	
D11	Jugale Αριστερό	Mastoidale Δεξί	133,1 / 124,6	132,1 / 123,7	134,1 / 125,6	85,4	85,7	85,0	
D12	Ovale mediale Αριστερό	Mastoidale Αριστερό	53,6 / 48,5	53,0 / 47,9	54,1 / 49,0	85,1	83,9	86,6	
D13	Glabella	Mastoidale Δεξί	134,3 / 124,8	133,3 / 123,7	135,3 / 125,8	85,1	84,9	85,2	
Γων. ^α	Αναφ. Σημ. 1	Αναφ. Σημ. 2	Αναφ. Σημ. 3						
A1	Glabella	Supra-Auricular Αριστερό	Maxillonasofrontale Αριστερό	5,3 / 6,6	5,1 / 6,4	5,5 / 6,9	80,6	87,2	72,8
A2	Mastoidale Δεξί	Opisthocranion	Zygotemporale superior Δεξί	25,0 / 22,6	24,7 / 22,2	25,4 / 23,0	79,5 / 78,9	83,5 / 82,4	75,0
A3	Glabella	Asterion Αριστερό	Maxillonasofrontale Αριστερό	4,5 / 5,5	4,3 / 5,3	4,6 / 5,7	79,3	84,9	72,8
A4	Glabella	Maxillonasofrontale Δεξί	Bregma	67,9 / 59,2	66,5 / 58,0	69,3 / 60,3	78,9	78,7	79,0
A5	Glabella	Maxillofrontale Δεξί	Bregma	71,2 / 64,1	70,1 / 63,1	72,4 / 65,2	78,6	77,2	80,2
A6	Ovale mediale Αριστερό	Mastoidale Δεξί	Coronale Δεξί	66,4 / 69,8	65,9 / 69,2	67,0 / 70,3	78,6	82,6	74,1
A7	Glabella	Supramastoid crest – squamous suture intersection Αριστερό	Maxillonasofrontale Αριστερό	5,2 / 6,3	5,0 / 6,1	5,4 / 6,5	78,2	86,0	69,1

^αΟι αποστάσεις είναι σε mm και οι γωνίες σε μοίρες· ^β Αρσενικά / Θηλυκά· ^γΜεμονωμένη τιμή δίνεται όταν το ποσοστά επιτυχούς ταξινόμησης είναι ίδιο με αυτό της διασταυρούμενης επικύρωσης, στην αντίθετη περίπτωση δίνονται και οι δύο τιμές με τη μορφή: αρχικά / διασταυρωμένη επικύρωση

Πίνακας 3.3: Αποτελέσματα πολυπαραγοντικής διακριτικής ανάλυσης. Ποσοστά επιτυχούς ταξινόμησης και διακριτικές εξισώσεις

Μεταβλητές που συμπεριλήφθηκαν ^α	Μέγεθος Δείγματος	Σταθερά	Συντελεστές	Ποσοστό επιτυχούς ταξινόμησης %		
				Συνολικά ^β	Αρσενικά ^β	Θηλυκά ^β
Όλες οι γωνίες A1-A7	169	-0,087	-0,260 +0,303 -2,065 +0,029 +0,039 -0,136 +1,649	86,4 / 85,2	88,8 / 86,5	83,8
A2+A3+A4+A6+A7	169	0,563	0,310 -2,159 +0,056 -0,130 +1,402	86,5 / 85,3	88,9 / 86,7	83,8
Όλες οι αποστάσεις D1-D13	167	-30,233	0,107 -0,017 +0,052 +0,069 -0,055 +0,056 +0,008 -0,148 -0,029 +0,193 -0,058 +0,118 0,073	92,2 / 88,6	90,8 / 86,2	93,8 / 91,3
D1+D4+D6	167	-30,745	0,214 +0,061 +0,077	90,6	87,8	93,8
Όλες οι γωνίες και οι αποστάσεις A1-A7 & D1-D13	166	-23,655	0,020 +0,093 -2,217 +0,004 +0,027 -0,020 +1,401 +0,011 -0,020 +0,046 +0,013 -0,072 +0,060 +0,028 -0,194 -0,054 +0,270 -0,064 +0,199 +0,102	96,4 / 91,6	96,5 / 90,7	96,3 / 92,5
A3+A7+D1+D6	166	-23,874	-1,951 +1,200 +0,241 +0,106	94,1 / 93,5	94,4 / 93,3	93,8
A5+D1+D6 ^γ	166	-31,370	0,084 +0,214 +0,109	95,8 / 95,2	95,5 / 94,3	96,3

^α Όλες οι μεταβλητές συμπεριλήφθηκαν τόσο στις αναλύσεις με την κλιμακωτή μέθοδο επιλογής ανεξάρτητων μεταβλητών όσο και στις αναλύσεις με τη μέθοδο χρήσης όλων των ανεξάρτητων μεταβλητών

^β Μεμονωμένη τιμή δίνεται όταν το ποσοστά επιτυχούς ταξινόμησης είναι ίδιο με αυτό της διασταυρούμενης επικύρωσης, στην αντίθετη περίπτωση δίνονται και οι δύο τιμές με τη μορφή: αρχικά / διασταυρωμένη επικύρωση

^γ Κλιμακωτή μέθοδος επιλογής ανεξάρτητων μεταβλητών χρησιμοποιώντας όλες τις αποστάσεις και τις γωνίες με εξαίρεση τις μεταβλητές A1, A3 και A7

Παρόλο που η μέθοδος χρήσης όλων των ανεξάρτητων μεταβλητών έδωσε υψηλότερα ποσοστά επιτυχούς ταξινόμησης και για τις γωνίες και για τις αποστάσεις (μέθοδος enter), η κλιμακωτή μέθοδος επιλογής ανεξάρτητων μεταβλητών αποκάλυψε ότι ο συνδυασμός τεσσάρων μόνο μεταβλητών, πιο συγκεκριμένα οι γωνίες A3 και A7 και οι αποστάσεις D1 και D6, εμφανίζει τη μεγαλύτερη ακρίβεια στην ταξινόμηση φύλου (επικυρωμένη επιτυχής ταξινόμηση 93,5%). Επιπλέον, ο ίδιος συνδυασμός παρέχει σχεδόν την ίδια ακρίβεια για άνδρες και γυναίκες (93,3% και 93,8% αντίστοιχα).

3.3. Συμπεράσματα

Σχετικά με τον εντοπισμό γνωρισμάτων, τα οποία να παρέχουν ακρίβεια στην ταυτοποίηση φύλου αγνώστου ατόμου, τα αποτελέσματά της παρούσας εργασίας ανέδειξαν φυλετικά διμορφικά χαρακτηριστικά που δεν είχαν προηγουμένως αναφερθεί στη βιβλιογραφία, όπως για παράδειγμα η απόσταση μεταξύ των αναφορικών σημείων *ovale mediale* και *mastoidale*. Επιπλέον, όλες οι υπόλοιπες αποστάσεις, οι οποίες εμφανίζουν ποσοστά επιτυχούς ταξινόμησης υψηλότερα από 85%,

αφορούν μήκη ευθύγραμμων τμημάτων μεταξύ διαμετρικά αντίθετων αναφορικών σημείων, τα οποία ωστόσο δεν είναι συμμετρικά (π.χ. αμφίπλευρα αναφορικά σημεία). Με άλλα λόγια, σύμφωνα με την παρούσα εργασία, οι πιο φυλετικά διμορφικές αποστάσεις δεν αντιστοιχούν σε αμφίπλευρα αναφορικά σημεία και συνήθως δεν αξιολογούνται από άλλους ερευνητές (Franklin et al., 2013; Mahakkanukrauh et al., 2015; Marinescu et al., 2014). Αντίστοιχο μοτίβο παρατηρείται και για τις μεταβλητές που αφορούν τις γωνίες, ανεξαρτήτως του γεγονότος ότι εμφανίζουν χαμηλότερα ποσοστά επιτυχούς ταξινόμησης σε σύγκριση με τις αποστάσεις. Τα πλεονεκτήματα των μη τυπικών διαστάσεων μεταξύ αναφορικών σημείων έναντι των τυπικών καθώς και των γεωμετρικών μορφομετρικών μεθόδων έχουν επίσης αναφερθεί από τους Spradley και Jantz (2016).

Αν και η χρήση εξοπλισμού όπως το Microscribe 3DX και οι γεωμετρικές μορφομετρικές μέθοδοι, οι οποίες βασίζονται σε αναφορικά σημεία, γίνονται ολοένα και πιο δημοφιλείς στην ιατροδικαστική ανθρωπολογία, η χρήση διακριτικών εξισώσεων που βασίζονται σε απλές μετρήσεις παραμένει πολύ σημαντική, καθώς χαλαρώνει τους χρονικούς περιορισμούς και την ανάγκη εξοπλισμού.

4. Ποικιλομορφία του βοθρίου της κεφαλής του μηριαίου

Το βοθρίο εντοπίζεται ελαφρώς οπίσθια και κατώτερα από το ακριβές κέντρο της επιφάνειας της κεφαλής του μηριαίου (Michaels & Matles, 1970; Wenger et al., 2007). Χρησιμοποιείται ως δείκτης σε διάφορες μελέτες και η μορφολογία του έχει αξιολογηθεί για διάφορες ανθρωπολογικές έρευνες. Επιπλέον, αποτελεί τυπικό αναφορικό σημείο κατά τον υπολογισμό της διακύμανσης του προσανατολισμού της κοτύλης, όταν αυτή πραγματοποιείται μέσω ανάλυσης εικόνας (Beltran et al., 2012). Το βοθρίο χρησιμεύει επίσης ως δείκτης για την αξιολόγηση δυσπλαστικών ισχίων (Siebenrock et al., 2013).

Από ανθρωπολογική σκοπιά, η μορφολογία του βοθρίου εκτός από το ότι χρησιμοποιείται για τον διαχωρισμό θραυσμάτων μακρών οστών ακόμη και ανήλικων ατόμων, πρόσφατα αξιολογήθηκε και η χρήση του ως δείκτης για την ταυτοποίηση του φύλου αγνώστου ατόμου (Murton et al., 2015). Ωστόσο, η βιβλιογραφία σχετικά με την ανατομία του βοθρίου είναι περιορισμένη (Acar et al., 2017; Perumal et al., 2017).

Σκοπός της παρούσας εργασίας ήταν να συμβάλει στην υπάρχουσα βιβλιογραφία αξιολογώντας κατά πόσον το βοθρίο μπορεί να χρησιμοποιηθεί στην εκτίμηση ηλικίας θανάτου ή στην εκτίμηση φύλου ενός ανθρώπινου σκελετού. Πιο συγκεκριμένα, η παρούσα εργασία αξιολόγησε τη διακύμανση του μεγέθους και του σχήματος του βοθρίου σε σχέση με την ηλικία και το φύλο χρησιμοποιώντας ξηρά οστά.

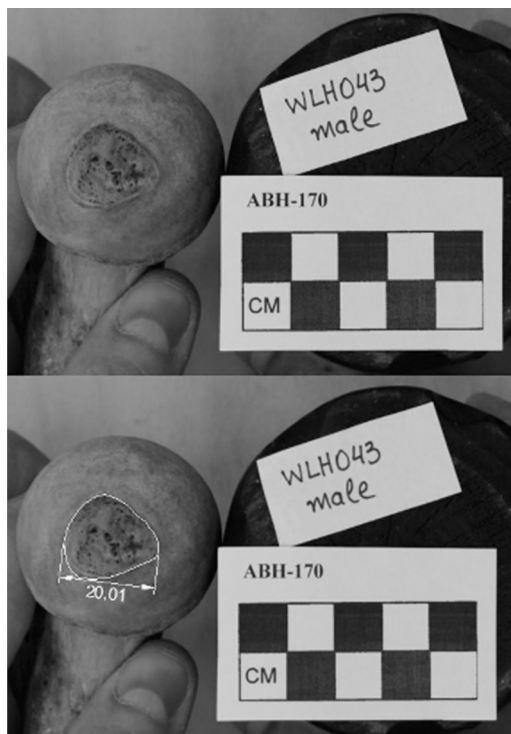
4.1. Υλικά και μέθοδοι

Για τις ανάγκες της παρούσας έρευνας χρησιμοποιήθηκαν 212 ενήλικα άτομα (120 αρσενικά και 92 θηλυκά), τα οποία δεν εμφάνιζαν παθολογική ανωμαλία στα οστά του μηριαίου, από τη σύγχρονη συλλογή αναφοράς. Δεδομένα συλλέχθηκαν από αμφότερα το δεξί και το αριστερό μηριαίο του κάθε ατόμου.

Για την καταγραφή της μέγιστης διαμέτρου της κεφαλής του μηριαίου (FMHD) χρησιμοποιήθηκε ηλεκτρονικό παχύμετρο Mitutoyo 150mm, ενώ η λήψη εικόνων του βοθρίου πραγματοποιήθηκε με ψηφιακή κάμερα Canon EOS 100D και φακό Canon EF 40mm f / 2.8 STM σταθερού εστιακού μήκους. Όλες οι εικόνες τραβήχτηκαν από ακίνητη θέση με κατακόρυφη προβολή κάθετη προς την κλίμακα αναφοράς σε σταθερή απόσταση 35 cm. Τα μηριαία τοποθετήθηκαν χειροκίνητα με την επιφάνεια του βοθρίου στραμμένη προς τα πάνω και στο ίδιο επίπεδο με την κλίμακα αναφοράς όπως φαίνεται στην Εικόνα 4.1. Οι διαστάσεις συντεταγμένες

ενός πολύγωνου, το οποίο περιγράφει τα όρια του βοθρίου εξήχθησαν με τη βοήθεια του λογισμικού AutoCAD.

Συνολικά, με το λογισμικό GNU Octave, υπολογίστηκαν 5 μεταβλητές (2 μεταβλητές σχήματος και 3 μεταβλητές μεγέθους) προκειμένου να διερευνηθεί η ποικιλομορφία του βοθρίου της κεφαλής του μηριαίου. Πιο συγκεκριμένα, οι 3 μεταβλητές μεγέθους είναι το εμβαδόν της επιφάνειας (FCA) και η μέγιστη διάμετρος (FCMD) του βοθρίου καθώς και ο λόγος της επιφάνειας του βοθρίου προς την επιφάνεια της κεφαλής του μηριαίου (FCA/FHS). Οι μεταβλητές σχήματος περιλαμβάνουν τους εξείς δύο λόγους: 1) το εμβαδόν της επιφάνειας του βοθρίου της κεφαλής του μηριαίου προς το τετράγωνο της περιμέτρου του ($ArPerIndex$), και 2) ο λόγος των I_{MIN} προς I_{MAX} (SMA ratio).



Εικόνα 4.1: Στο πάνω τμήμα απεικονίζεται η φωτογραφία πριν από την επεξεργασία με το λογισμικό, όπου φαίνεται η αναφορική κλίμακα. Στο κάτω τμήμα απεικονίζεται η φωτογραφία μετά την επεξεργασία όπου φαίνεται το πολύγωνο περιγραφής των ορίων του βοθρίου και έχει σημειωθεί η μέγιστη διάστασή του σε cm, που αντιστοιχεί στην μεταβλητή FCMD.

Αρχικά πραγματοποιήθηκε έλεγχος κατευθυνόμενης ασυμμετρίας σε όλες τις αρχικές μεταβλητές, μέσω paired sample t-test ή Wilcoxon test, προκειμένου να ελεγχθεί εάν η αριστερή και η δεξιά πλευρά θα έπρεπε να αναλυθούν ξεχωριστά. Στην συνέχεια με τη βοήθεια των Independent sample t-test ή Mann-Whitney test ελέγχθηκε η ύπαρξης φυλετικού διμορφισμού. Επιπλέον, το δείγμα χωρίστηκε σε 3 κατηγορίες ηλικιακών ομάδων: 1) 20-39 ετών, 2) 40-59 ετών και 3) ≥ 60 ετών προκειμένου να ελεγχθεί εάν εντοπίζονται μορφολογικές αλλαγές, οι οποίες να σχετίζονται με την ηλικία (ANOVA / Kruskal-Wallis H test). Τέλος, η καταλληλότητα του βοθρίου της κεφαλής του μηριαίου ως δείκτης εκτίμησης φύλου και ηλικίας αγνώστου ατόμου αξιολογήθηκε μέσω της διακριτικής ανάλυσης (DEA). Όλες οι στατιστικές αναλύσεις πραγματοποιήθηκαν με το λογισμικό SPSS (PASW Statistics 23.0).

4.2. Αποτελέσματα

Οι μεταβλητές FCA/FHS και SMA ratio παρουσίασαν κατευθυνόμενη ασυμμετρία στα θηλυκά άτομα. Κατά συνέπεια, για τις συγκεκριμένες μεταβλητές, η αριστερή και η δεξιά πλευρά αναλύθηκαν ξεχωριστά. Αντιθέτως, για τις μετέπειτα αναλύσεις των μεταβλητών FCA, FCMD και ArperIndex χρησιμοποιήθηκαν οι μέσες τιμές της αριστεράς και της δεξιάς πλευράς.

Στον Πίνακα 4.1 παρατίθενται τα αποτελέσματα των t-test ή Mann-Whitney test σχετικά με την ύπαρξη φυλετικού διμορφισμού. Σύμφωνα με τα αποτελέσματα, μόνο οι μεταβλητές μεγέθους FCA και FCMD είναι φυλετικά διμορφικές και η εφαρμογή διακριτικής ανάλυσης στις εν λόγω μεταβλητές οδήγησε στην δημιουργία στατιστικά σημαντικών μοντέλων (Wild's lambda: 0,888 και 0,884 αντίστοιχα, $p < 0,001$ και για τις δύο). Η επικυρωμένη επιτυχής ταξινόμηση κυμαίνεται από 62,7% έως 65,6% για τις μεταβλητές FCA και FCMD αντίστοιχα (Πίνακας 4.2).

Πίνακας 4.1: Αποτελέσματα Independent sample t-test (A) και Mann-Whitney test (B) σχετικά με την ύπαρξη φυλετικού διμορφισμού στις μεταβλητές του βοθρίου της κεφαλής του μηριαίου.

A)			
	t	df	Sig.
FCA	5.335	208	<.001
FCMD	5.151	209	<.001
SMA ratio (Δεξιά πλευρά)	-0.788	196	.431
FCA/FHS (Αριστερή πλευρά)	-0.969	198	.334
FCA/FHS (Δεξιά πλευρά)	-0.773	196	.440
Significant difference ($p \leq 0.05$)			
B)			
	Mann-Whitney U	Sig.	
ArperIndex	4778	.114	
SMA ratio (Αριστερή πλευρά)	4696	.589	
Στατιστικά σημαντική διαφορά ($p \leq 0.05$)			

Πίνακας 4.2: Αποτελέσματα διακριτικής ανάλυσης. Ποσοστά επιτυχούς ταξινόμησης και διακριτικές εξισώσεις για τις μεταβλητές FCA και FCMD.

	N	Σταθερά	Συντελεστές	Ποσοστό επιτυχούς ταξινόμησης %		
				Συνολικά ^α	Αρσενικά ^α	Θηλυκά ^α
FCA	212	-3.271	0.016	62.7	69.2	54.3
FCMD	212	-6.375	0.333	65.6 / 65.1	75.8	52.2 / 51.1

^α Μεμονωμένη τιμή δίνεται όταν το ποσοστό επιτυχούς ταξινόμησης είναι ίδιο με αυτό της διασταυρούμενης επικύρωσης, στην αντίθετη περίπτωση δίνονται και οι δύο τιμές με τη μορφή: αρχικά / διασταυρωμένη επικύρωση

Όσον αφορά τις εκφυλιστικές μορφολογικές αλλαγές του βοθρίου της κεφαλής του μηριαίου, μόνο το ArPerIndex παρουσίασε στατιστικά σημαντική διαφορά ($p = 0,028$) μεταξύ των ηλικιακών ομάδων και πιο συγκεκριμένα μόνο μεταξύ των ηλικιακών ομάδων «20-39 ετών» και « ≥ 60 ετών» στη μεταβλητή ArPerIndex (Πίνακας 4.3).

Πίνακας 4.3: Αποτελέσματα ANOVA (A) and Kruskal Wallis test (B) σχετικά με τις εκφυλιστικές μορφολογικές αλλαγές των μεταβλητών του βοθρίου της κεφαλής του μηριαίου

A)			
	F	df	Sig.
FCMD (Αρσενικά άτομα)	0.601	2 / 117	.550
FCMD (Θηλυκά άτομα)	0.662	2 / 89	.518
SMA ratio (Δεξιά πλευρά)	0.041	2 / 196	.960
FCA/FHS (Αριστερή πλευρά)	0.801	2 / 198	.450

Στατιστικά σημαντική διαφορά ($p \leq 0.05$)
 Οι βαθμοί ελευθερίας αναγράφονται υπό τη μορφή: Μεταξύ ομάδων / Εντός ομάδων

B)			
	Chi-Square	df	Sig.^a
FCA (Αρσενικά άτομα)	0.019	2	.991
FCA (Θηλυκά άτομα)	0.009	2	.996
SMA ratio (Αριστερή πλευρά)	5.664	2	.059
FCA/FHS (Δεξιά πλευρά)	0.300	2	.861
ArPerIndex	7.137	2	.028

	Dunn's pairwise tests	Median	Adj. Sig.^b
	20-39 / 40-59	0.895	.511
	40-59 / ≥ 60	0.885	.593
	≥ 60 / 20-39	0.867	.025

^a Στατιστικά σημαντική διαφορά ($p \leq 0.05$)

^b Η στατιστικά σημαντική διαφορά είναι προσαρμοσμένη βάσει των διορθώσεων Bonferroni ($p \leq 0.05$)

4.3. Συμπεράσματα

Σύμφωνα με τα αποτελέσματα, οι μεταβλητές μεγέθους FCA και FCMD είναι φυλετικά διμορφικές, ενώ μόνο η μεταβλητή σχήματος ArPerIndex εμφανίζει εκφυλιστικές αλλαγές. Η μεταβλητή ArPerIndex, αποτελεί καινοτομία δεδομένου ότι δεν υπήρχε στη βιβλιογραφία πριν τη δημοσίευση της συγκεκριμένης μελέτης. Επιπλέον, η εξαιρετική δυνατότητα της συγκεκριμένης μεταβλητής ως μορφολογικός δείκτης περιγράμματος επιβεβαιώνεται από την δυνατότητα περιγραφής των εκφυλιστικών αυτών αλλαγών του βοθρίου της κεφαλής του μηριαίου. Πιο συγκεκριμένα, η περίμετρος του βοθρίου της κεφαλής του μηριαίου τείνει να έχει πιο ακανόνιστο σχήμα σε ηλικιωμένα άτομα (το ArPerIndex μειώνεται). Το γεγονός αυτό θα μπορούσε να είναι

αποτέλεσμα οστικής αναπαραγωγής, η οποία σχετίζεται με τη διαδικασία γήρανσης (Boskey & Coleman, 2010) ή πιθανώς ένας συνδυασμός γήρανσης και προτύπων δραστηριότητας που μπορεί να οδηγήσουν στην παρατηρούμενη μορφολογική διακύμανση.

Ωστόσο, όπως φαίνεται από τα αποτελέσματα της διακριτικής ανάλυσης καθώς και τα αποτελέσματα των Dunn's pairwise tests, το μέγεθος των αλλαγών και στις δύο περιπτώσεις είναι μικρό και κατά συνέπεια το βοθρίο της κεφαλής του μηριαίου δεν μπορεί να χρησιμοποιηθεί για την εκτίμηση της ηλικίας θανάτου ή την εκτίμηση του φύλου αγνώστου ατόμου.

5. Διαχρονικές τάσεις στην εξέλιξη του ύψους στον Ελληνικό πληθυσμό.

Η τεκμηρίωση της ανθρώπινης ποικιλομορφίας όσον αφορά το βιοτικό επίπεδο σε βάθος χρόνου, αποτελεί κύριο στόχο των κοινωνικών ανθρωπολόγων, των ιστορικών και των οικονομολόγων, προκειμένου να εξηγήσουν τις χωρικές και χρονικές διακυμάνσεις στην ανθρώπινη υγεία και την ευημερία του πληθυσμού.

Το ανάστημα θεωρείται καλός δείκτης της κοινωνικοοικονομικής κατάστασης ή του βιοπολιτισμικού περιβάλλοντος των ανθρώπινων πληθυσμών (Komlos, 2009; Steckel, 2008). Παρόλο που το ανάστημα επηρεάζεται κυρίως από γενετικούς παράγοντες (Lango Allen et al., 2010), οι συνθήκες υγιεινής, η διατροφική ποιότητα καθώς και η συνολική υγεία μπορούν να επιδράσουν σε αυτό (Keep & Bogin, 1999; Susanne, 1985).

Καθ' όλη τη διάρκεια του 19ου αιώνα, η Ελλάδα ήταν ως επί το πλείστον γεωργική κοινωνία. Ωστόσο, στα τέλη του 19ου και στις αρχές του 20ου αιώνα ο αγροτικός ελληνικός πληθυσμός ξεκίνησε να προσαρμόζεται σε έναν πιο αστικό τρόπο ζωής μετά την αλλαγή πολιτικής του ελληνικού κράτους προς τις δημόσιες επενδύσεις σε αστικές υποδομές. Εκτός από τη μετάβαση του αγροτικού ελληνικού πληθυσμού σε αστικό, το πρώτο μισό του 20ού αιώνα βρίσκει την Ελλάδα σε τέσσερις μεγάλους πολέμους. Πιο συγκεκριμένα, τους Βαλκανικούς πολέμους που διήρκεσαν από το 1912 έως το 1913, τον Α' Παγκόσμιο Πόλεμο, ο οποίος διήρκεσε 3 χρόνια (1916 –1918), τον Ελληνοτουρκικό πόλεμο που διήρκεσε από το 1919 έως το 1922, και τον Β' Παγκόσμιο Πόλεμο, που διήρκεσε από το 1940 μέχρι το 1945. Επιπλέον, μετά τον Β' Παγκόσμιο Πόλεμο ακολούθησε ο Εμφύλιος Πόλεμος, ο οποίος διήρκεσε άλλα 5 χρόνια (1946-1950). Η μακρά μεσοπολεμική περίοδος (1912–1950) εμπόδισε σημαντικά τη γεωργική ανάπτυξη του 19ου αιώνα, ενώ η συνακόλουθη οικονομική κρίση μαζί με τη Μεγάλη Ύφεση της δεκαετίας του 1930 οδήγησε στη διατλαντική μετανάστευση κατά τις αρχές του 20ου αιώνα.

Μελετώντας την χρονική περίοδο από τα τέλη του 19ου έως τα μέσα του 20ού αιώνα, στόχος ήταν να καταγράψουμε ολόκληρη την περίοδο μετασχηματισμού της ελληνικής κοινωνίας και να διερευνήσουμε πώς οι κοινωνικοοικονομικές αλλαγές στην Ελλάδα επηρέασαν το ανάστημα του πληθυσμού της.

5.1. Υλικά και μέθοδοι

Για την παρούσα εργασία μελετήθηκαν 189 ενήλικες (104 άνδρες και 85 γυναίκες) από τη Σύγχρονη Συλλογή Αναφοράς. Η ερευνητική προσέγγιση που επιλέχθηκε περιλάμβανε τον

υπολογισμό του αναστήματος βάσει των εξισώσεων παλινδρόμησης του Ruff και των συναδέλφων του (2012), και οι οποίες στηρίζονται στο μέγιστο μήκος του μηριαίου, ενώ η διαχρονική τάση αξιολογήθηκε με γραμμική παλινδρόμηση του αναστήματος σε σχέση με το έτος γέννησης.

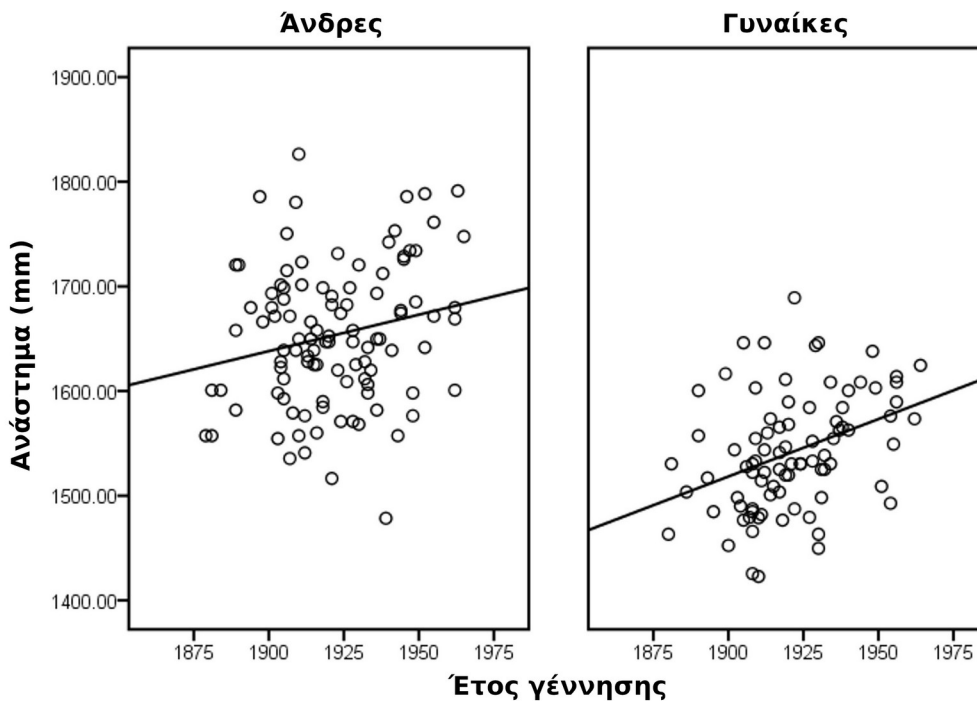
Επιπλέον, με βάση το έτος γέννησης κάθε ατόμου, το δείγμα χωρίστηκε σε τρεις χρονικές περιόδους. Πιο συγκεκριμένα, τη χρονική περίοδο 1879 έως 1911, η οποία αντιστοιχεί στη γεωργική στασιμότητα στις αρχές του αιώνα, τη χρονική περίοδο 1912 έως 1950, που αντιστοιχεί στο διάστημα όπου η Ελλάδα συμμετείχε σε μια σειρά πολέμων, και τη χρονική περίοδο 1951 έως 1965, που σηματοδοτεί τη μεταπολεμική ταχεία οικονομική ανάπτυξη και τον εκσυγχρονισμό της ελληνικής κοινωνίας. Στις τρεις χρονικές περιόδους πραγματοποιήθηκε ανάλυση διακύμανσης (ANOVA) προκειμένου να εκτιμηθεί εάν οι δραματικές μεταβολές στην οικονομία επηρέασαν το ανάστημα του ελληνικού πληθυσμού.

Θα πρέπει να σημειωθεί ότι η συγκεκριμένη μελέτη στηρίχθηκε στην υπόθεση ότι δεν πραγματοποιήθηκε ουσιώδης γενετική αλλαγή που να επηρέασε το ανάστημα κατά τη διάρκεια της χρονικής περιόδου ενδιαφέροντος. Η υπόθεση αυτή βασίζεται σε μεγάλο βαθμό στα ιστορικά στοιχεία σχετικά με τα μεταναστευτικά πρότυπα στην ελληνική χερσόνησο.

Από τον σχηματισμό του ελληνικού κράτους, πραγματοποιήθηκαν δύο μεγάλα κύματα μετανάστευσης. Η πρώτη μαζική μετανάστευση προκλήθηκε από την ταχεία πτώση της τιμής των σταφίδων και την επακόλουθη οικονομική κρίση του 1893, ενώ το δεύτερο κύμα πραγματοποιήθηκε μετά τον Β΄ Παγκόσμιο Πόλεμο. Το μόνο κύμα μετανάστευσης που πραγματοποιήθηκε κατά το πρώτο μισό του 20ού αιώνα αφορά τον επαναπατρισμό των Ελλήνων της Μικράς Ασίας. Ωστόσο, δεν θεωρείται αλλοδαπός πληθυσμός, καθώς είναι ελληνικής καταγωγής και πολιτισμικά και βιολογικά μπορούν να θεωρηθούν Έλληνες.

5.2. Αποτελέσματα

Σύμφωνα με τα αποτελέσματά, κατά την περίοδο 1879 και 1965, το ανάστημα αυξήθηκε σημαντικά τόσο για τους άνδρες όσο και για τις γυναίκες και παρόλο που κλίση της ευθείας παλινδρόμησης για τους άνδρες ήταν 0,696 και η αντίστοιχη για τις γυναίκες ήταν 1,098 (Εικόνα 5.1), ο συντελεστής παλινδρόμησης δε διέφερε σημαντικά μεταξύ των δύο φύλων.



Εικόνα 5.1: Ευθείες παλινδρόμησης αρσενικών (αριστερά) και θηλυκών (δεξιά) ατόμων ελληνικού πληθυσμού.

Τα αποτελέσματα της ANOVA έδειξαν ότι οι διαφορές στο μέσο ανάστημα μεταξύ των χρονικών περιόδων είναι στατιστικά σημαντικές τόσο για τους άνδρες όσο και για τις γυναίκες. Ωστόσο, τα αποτελέσματα των πολλαπλών συγκρίσεων αποκάλυψαν ότι τα αρσενικά άτομα εμφάνισαν στατιστικά σημαντική διαφορά μόνο μεταξύ της δεύτερης και της τρίτης περιόδου, ενώ τα θηλυκά είχαν στατιστικά σημαντική αύξηση ύψους μεταξύ της πρώτης και της δεύτερης περιόδου καθώς και μεταξύ της πρώτης και της τρίτης περιόδου (Πίνακας 5.1).

Πιο συγκεκριμένα, όπως φαίνεται στην Εικόνα 5.2, κατά την περίοδο 1879 - 1911 το ανάστημα των γυναικών ήταν 151,4 cm, αυξήθηκε στα 155,4 cm κατά την περίοδο 1912 - 1950, και αυξήθηκε περαιτέρω στα 157,1 cm κατά την περίοδο 1951 - 1965. Αντιθέτως, τα αρσενικά εμφάνισαν μικρή μείωση στο μέσο ανάστημα μεταξύ των δύο πρώτων χρονικών περιόδων από 165,7 cm σε 164,4 cm, αν και αυξήθηκε σημαντικά την τρίτη περίοδο φτάνοντας τα 170,6 cm.

Πίνακας 5.1: Αποτελέσματα ανάλυσης διακύμανσης: Σύγκριση αναστήματος αρσενικών και θηλυκών ατόμων μεταξύ χρονικών περιόδων.

ANOVA						
	Άθροισμα τετραγώνων αποκλίσεων	df	Μέσα τετράγωνα αποκλίσεων	F	Sig.	Levene's Test Sig.
Αρσενικά	30688,58	2	15344,29	3,506	0,034	0,426
Θηλυκά	37220,76	2	18610,38	6,872	0,002	0,881

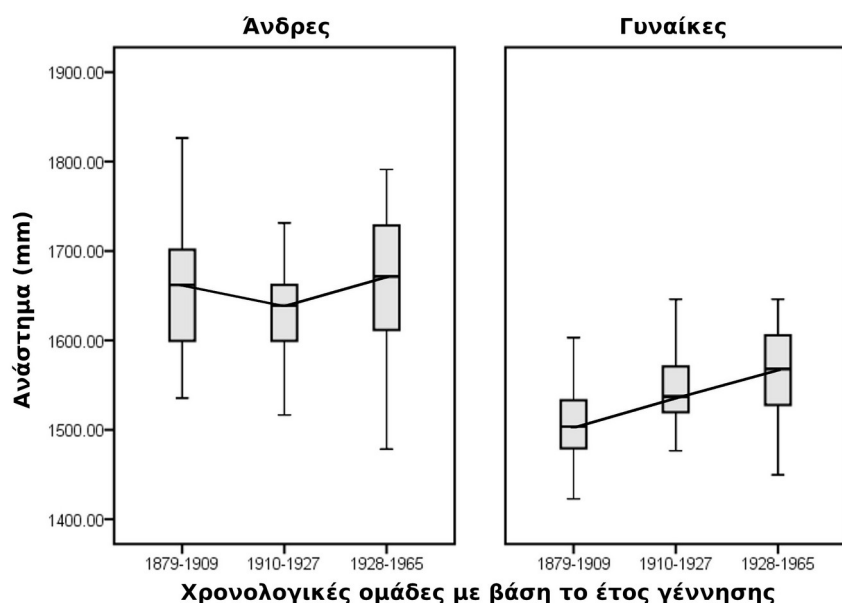
Πολλαπλές συγκρίσεις μέσου αναστήματος αρσενικών ατόμων μεταξύ των χρονικών περιόδων

		Διαφορά Μέσων Όρων (mm)	Std. Error	Sig.	95% Διάστημα εμπιστοσύνης	
					Κατώτερο όριο	Ανώτερο όριο
1879-1911	1912-1950	13,02	13,99	0,729	-20,94	46,97
	1951-1965	-49,12	24,66	0,139	-108,95	10,71
1912-1950	1879-1911	-13,02	13,99	0,729	-46,97	20,93
	1951-1965	-62,13*	23,68	0,030	-119,58	-4,69
1951-1965	1879-1911	49,12	24,66	0,139	-10,71	108,95
	1912-1950	62,13*	23,68	0,030	4,69	119,58

Πολλαπλές συγκρίσεις μέσου αναστήματος θηλυκών ατόμων μεταξύ των χρονικών περιόδων

		Διαφορά Μέσων Όρων (mm)	Std. Error	Sig.	95% Διάστημα εμπιστοσύνης	
					Κατώτερο όριο	Ανώτερο όριο
1879-1911	1912-1950	-40,02*	12,29	0,005	-69,95	-10,09
	1951-1965	-57,25*	19,86	0,015	-105,61	-8,89
1912-1950	1879-1911	40,02*	12,29	0,005	10,09	69,95
	1951-1965	-17,22	18,94	0,742	-63,35	28,89
1951-1965	1879-1911	57,25*	19,86	0,015	8,89	105,61
	1912-1950	17,22	18,94	0,742	-28,89	63,35

* Στατιστικά σημαντικές διαφορές με $\alpha = 0,05$.



Εικόνα 5.2: Σύγκριση αναστήματος αρσενικών (αριστερά) και θηλυκών (δεξιά) ατόμων μεταξύ χρονικών περιόδων.

5.3. Συμπεράσματα

Αρκετοί συγγραφείς κατά τη διάρκεια του 20ου αιώνα έχουν δημοσιεύσει βιομετρικά αρχεία αναστήματος από ζώντες ανθρώπους από διάφορα μέρη της Ελλάδας, τα οποία καλύπτουν την ίδια χρονική περίοδο στην οποία επικεντρώθηκε η παρούσα εργασία. Ωστόσο, αυτές οι μελέτες εστίασαν σε περιορισμένα χρονικά διαστήματα. Για παράδειγμα, ο Βαλαώρας (Valaoras, 1970) αξιολόγησε τις διαχρονικές τάσεις στην εξέλιξη του ύψους των ανδρών που γεννήθηκαν στην Ελλάδα κατά την περίοδο 1927 έως 1945 με βάση τα βιομετρικά χαρακτηριστικά στρατιωτών.

Όσον αφορά την πρώτη χρονική περίοδο (1879–1911), παρατηρούνται διαφορές μεταξύ των ευρημάτων μας και τις παρατηρήσεις άλλων ερευνητών τόσο στον ανδρικό όσο και στο γυναικείο πληθυσμό. Σύμφωνα με τις έρευνες του Εξαρχόπουλου (1931) και του Λαμπαδάρου (1921), ο πληθυσμός ήταν ψηλότερος. Ωστόσο, και τα δύο δείγματα, τα οποία μελέτησαν ο Εξαρχόπουλος και ο Λαμπαδάριος, προέρχονταν από μαθητές γυμνασίου της Αθήνας, υποδηλώνοντας μια κοινωνικοοικονομική κατάσταση ανώτερης τάξης, καθώς εκείνη την εποχή οι απόφοιτοι λυκείου βρίσκονταν κυρίως μεταξύ των προνομιούχων του ελληνικού πληθυσμού. Επιπλέον, ο Roberts (1954), καθώς και οι Hasluck και Morant (1929) κατέγραψαν το ύψος του γενικού πληθυσμού των Κρητικών και των Ελλήνων Μακεδόνων, αντίστοιχα. Παρόλο που και οι δύο πληθυσμοί φαίνονται να είναι ψηλότεροι από το δείγμα μας, η διαφορά αυτή είναι αναμενόμενη δεδομένου ότι πρόκειται για τους δύο ψηλότερους υποπληθυσμούς της ελληνικής εθνότητας όπως αναφέρει και ο Βαλαώρας στα ευρήματά του το 1970.

Το σημαντικό εύρημα της παρούσας εργασίας σχετικά με τη δεύτερη χρονική περίοδο (1912–1950) ήταν ότι οι άνδρες παρουσίασαν μείωση στο ανάστημα, ενώ οι γυναίκες ακολούθησαν μια συνεχή αύξηση. Αυτό είναι σε συμφωνία με προηγούμενα δημοσιευμένα έργα με εξαίρεση την έρευνα του Βαλαώρα και του Παπαϊωάννου το 1948, η οποία στηρίχτηκε σε μαθητές γυμνασίου της Αθήνας. Ως εκ τούτου, θα μπορούσε κανείς να υποστηρίξει ότι την ίδια χρονική περίοδο όπου το ύψος του γενικού αντρικού πληθυσμού φαίνεται να μειώνεται ως αποτέλεσμα του συνεχούς πολέμου και της οικονομικής στέρησης, η κοινωνικοοικονομική ελίτ στις αστικές περιοχές επωφελείται από την τεχνολογική πρόοδο στην υγειονομική περίθαλψη και τη διατροφή.

Η τρίτη περίοδος είναι καλά τεκμηριωμένη λόγω των έργων του Φθενάκη (Fthenakis, 1970), καθώς και του Μανώλη και των συναδέλφων του (Manolis et al., 1995). Το μόνο μειονέκτημα είναι ότι όλα τα δείγματα περιελάμβαναν φοιτητές τριτοβάθμιας εκπαίδευσης, γεγονός που μπορεί να καθιστά το δείγμα μη αντιπροσωπευτικό ως προς την κοινωνικοοικονομική κατάσταση του γενικού πληθυσμού. Ωστόσο, και τα δύο πανεπιστημιακά δείγματα όπως και το δείγμα της συγκεκριμένης

έρευνας έδειξαν ότι το ανάστημα του ελληνικού πληθυσμού αυξήθηκε στη μεταπολεμική περίοδο τόσο στους άνδρες όσο και στις γυναίκες.

Συμπερασματικά, η διατροφική στέρηση κατά την παιδική ηλικία ως αποτέλεσμα του πολέμου μπορεί να είχε αντίκτυπο στο ύψος εξηγώντας έτσι τη μείωση που παρατηρείται στην διαχρονική τάση της εξέλιξης του ύψους των ανδρών μεταξύ του 1912 και του 1950. Ωστόσο, αυτό δεν είναι εμφανές στα δεδομένα όσον αφορά τον γυναικείο πληθυσμό, του οποίου το ύψος συνεχίζει να αυξάνεται κατά τη διάρκεια των πολέμων. Μια πιθανή εξήγηση αυτού του φαινομένου είναι το γεγονός ότι δεν ήταν ασυνήθιστο τα αγόρια να συμμετείχαν πιο συχνά στον πόλεμο από ότι τα κορίτσια. Εκτός αυτού, ακόμα κι αν δεν εμπλέκονταν άμεσα, η απουσία ενήλικων ανδρών λόγω της συμμετοχής τους στον πόλεμο, επέφερε μεγαλύτερο βάρος στα αγόρια για τη στήριξη των οικογενειών τους, κάτι που μπορεί να εξηγηθεί από τα πολιτιστικά και ηθικά πρότυπα της ελληνικής κοινωνίας κατά τη διάρκεια εκείνων των χρόνων. Επιπλέον, σε σύγκριση με άλλους πληθυσμούς, η συνολικά θετική διαχρονική τάση του αναστήματος του ελληνικού πληθυσμού κατά τη χρονική περίοδο στην οποία εστίασε αυτή η μελέτη παραμένει μέτρια, και μόνο από τη δεκαετία του 1960 και μετά παρατηρείται ταχεία αύξηση του ύψους, η οποία συμπίπτει με την ταχεία εκβιομηχάνιση και την οικονομική ανάπτυξη της Ελλάδας από εκείνη τη δεκαετία και έπειτα. Ως εκ τούτου, μπορούμε να υποστηρίξουμε ότι η αρνητική επίδραση της οικονομικής και διατροφικής στέρησης στο ανάστημα ως αποτέλεσμα του πολέμου ήταν εμφανής στο δείγμα μας, τουλάχιστον για τους άνδρες, ενώ η θετική επίδραση της οικονομικής ανάπτυξης ήταν επίσης εμφανής για ολόκληρο τον πληθυσμό.

6. «Τι και πώς πρέπει να μοιραζόμαστε μετρικά δεδομένα» Μια σύγκριση του σφάλματος μέτρησης κρανιακών αναφορικών σημείων μεταξύ μεθόδων και παρατηρητών.

Η ανάλυση μετρικών χαρακτήρων μπορεί να θεωρηθεί ο ακρογωνιαίος λίθος της οστεολογικής έρευνας. Αποστάσεις και γωνίες, οι οποίες λαμβάνονται με τη βοήθεια παχύμετρου, χρησιμοποιούνται παραδοσιακά στις αναλύσεις μεγέθους και σχήματος (Ross & Williams, 2008). Από την άλλη, η εισαγωγή τις τελευταίες δεκαετίες των τριδιάστατων γεωμετρικών μορφομετρικών μεθόδων βασιζόμενων σε αναφορικά σημεία (3DGM), άνοιξε νέους ορίζοντες για την ανάλυση πιο περίπλοκων μορφολογικών δεδομένων χρησιμοποιώντας ψηφιακά μοντέλα (Shearer et al., 2017). Και οι δύο προσεγγίσεις εμφανίζουν πλεονεκτήματα και μειονεκτήματα (Bertsatos, Papageorgiou, et al., 2018; Spradley & Jantz, 2016), ωστόσο, και οι δύο μοιράζονται μια εγγενή ιδιότητα, η οποία είναι το σφάλμα μέτρησης.

Η δημοτικότητα της γεωμετρικής μορφομετρίας στην οστεολογική έρευνα μαζί με τη διαρκώς αυξανόμενη διαθεσιμότητα ευρέος φάσματος εξοπλισμού για την ψηφιοποίηση οστεολογικών δειγμάτων, καθώς και την ολοένα συχνότερη ανταλλαγή γεωμετρικών μορφομετρικών δεδομένων μεταξύ των ερευνητών (Fruciano et al., 2017), οδήγησε σε μια άλλη «δυσνητικά επιζήμια» πηγή συστηματικού σφάλματος μέτρησης, η οποία προκύπτει από τη χρήση διαφορετικών μεθόδων μέτρησης ή εξοπλισμού σε ομαδοποιημένα δείγματα. Όπως συμπεραίνουν ο Fruciano και οι συνεργάτες του (2017), ένα τέτοιου είδους σφάλμα μέτρησης μπορεί να αποτελέσει πηγή σημαντικής διακύμανσης με απρόσμενη επίδραση στις εκτιμήσεις παραμέτρων κατά τον συνδυασμό διαφορετικών μορφομετρικών δεδομένων. Επιπλέον, αυτές οι ανησυχίες δεν ισχύουν μόνο κατά την ομαδοποίηση δεδομένων από διαφορετικές πηγές, αλλά και κατά τον έλεγχο μεθόδων.

Η παρούσα εργασία αξιολόγησε την ακρίβεια και την ορθότητα των τριδιάστατων ψηφιακών κρανιακών μοντέλων, τα οποία προκύπτουν μέσω της μεθόδου της φωτογραμμετρίας, στην λήψη αναφορικών σημείων. Δεδομένου ότι τα ψηφιακά μοντέλα χρησιμοποιούνται ολοένα και περισσότερο τόσο στη γεωμετρική μορφομετρία όσο και στην λήψη «παραδοσιακών» μετρήσεων όσον αφορά αποστάσεις (ILD) και / ή γωνιών μεταξύ αναφορικών σημείων (ILA), η αξιολόγηση πραγματοποιήθηκε σε δύο άξονες: 1) λαμβάνοντας αποστάσεις και γωνίες από τα ψηφιακά μοντέλα και συγκρίνοντας με αντίστοιχες μετρήσεις που ελήφθησαν είτε με παχύμετρο είτε με το 3D Microscribe, και 2) συγκρίνοντας τριδιάστατες συντεταγμένες διαφορετικών τύπων αναφορικών

σημείων που ελήφθησαν από ψηφιακά μοντέλα (Meshlab) και από πραγματικά οστά (3D Microscribe) καθώς και επαναληπτικές μετρήσεις μεταξύ διαφορετικών παρατηρητών.

6.1. Υλικά και μέθοδοι

Για τις ανάγκες της παρούσας εργασίας ψηφιοποιήθηκαν 50 κρανία με τη μέθοδο της φωτογραμμετρίας. Η επιλογή του δείγματος βασίστηκε στην ίση αντιπροσώπευση των φύλων και των ηλικιακών ομάδων. Στην συνέχεια, ελήφθησαν 56 αναφορικά σημεία (12 στην οβελιαία επιφάνεια και 22 αμφίπλευρα). Αναφορικά σημεία από την ίδια σκελετική συλλογή είχαν επίσης ψηφιοποιηθεί με το 3D Microscribe σε προηγούμενη έρευνα και το υποσύνολο των συντεταγμένων που αντιστοιχούσε στο δείγμα και τα αναφορικά σημεία της παρούσας μελέτης επαναχρησιμοποιήθηκαν για σύγκριση. Επιπλέον, τα αναφορικά σημεία 20 δειγμάτων ελήφθησαν δύο φορές από διαφορετικούς παρατηρητές, προκειμένου να ληφθεί υπόψη η αναπαραγωγικότητα κατά την απόκτηση τριδιάστατων συντεταγμένων αναφορικών σημείων από τα ψηφιακά μοντέλα με το λογισμικό Meshlab. Από τα ίδια 20 δείγματα ελήφθησαν και ένα σύνολο ILDs και ILAs με παχύμετρο για την σύγκριση μεταξύ των μεθόδων σχετικά με τις γραμμικές μετρήσεις.

Η αναπαραγωγικότητα κατά την απόκτηση τριδιάστατων συντεταγμένων αναφορικών σημείων από τα ψηφιακά μοντέλα με το λογισμικό Meshlab υπολογίστηκε, για κάθε τύπο αναφορικών σημείων, ως η ελάχιστη απόκλιση της μέσης τετραγωνικής τιμής των τιμών από το μέσο όρο τους (LRMSD) μεταξύ των ζευγών των διατάξεων των αναφορικών σημείων κατόπιν της υπέρθεσης τους βάσει του αλγόριθμου Kabsch (Kabsch, 1976). Επιπλέον, υπολογίστηκε η μέση απόκλιση κάθε αναφορικού σημείου μεμονωμένα καθώς και τα περιγραφικά στατιστικά στοιχεία των απόλυτων διαφορών μεταξύ όλων των δυνατών ILDs, όπως αυτές προέκυψαν από τα ζεύγη των διατάξεων των αναφορικών σημείων.

Προκειμένου να εκτιμηθεί η ακρίβεια της απόκτησης αναφορικών σημείων από τα ψηφιακά μοντέλα, εφαρμόστηκε η ίδια αναλυτική ροή εργασίας στα ζεύγη των διατάξεων των αναφορικών σημείων, τα οποία προέκυψαν από το σύνολο των 50 κρανίων που ψηφιοποιήθηκαν με τη μέθοδο της φωτογραμμετρίας και το αντίστοιχο σετ δεδομένων που είχε ψηφιοποιηθεί με το 3D Microscribe σε πραγματικά οστά από προηγούμενη έρευνα. Τέλος, αξιολογήθηκε η ακρίβεια της φωτογραμμετρικής μοντελοποίησης με τον υπολογισμό του τεχνικού σφάλματος της μέτρησης (TEM) καθώς και του σχετικού TEM (S. Lewis, 1999) στο σύνολο των 10 ILAs και ILDs που ελήφθησαν από τα 20 κρανία.

6.2. Αποτελέσματα

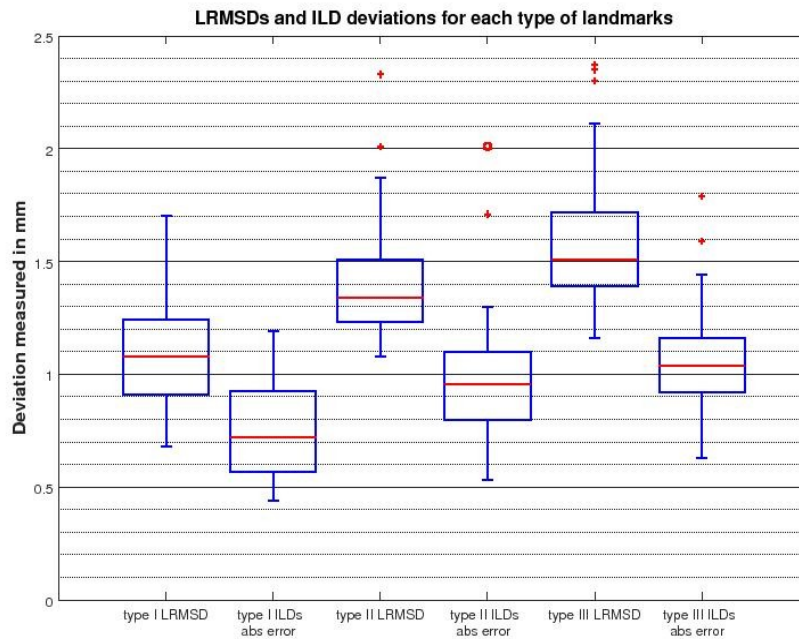
Όσον αφορά την ακρίβεια της φωτογραμμετρικής μοντελοποίησης, τα αποτελέσματα TEM μεταξύ κάθε ζεύγους διαφορετικών μεθόδων μέτρησης παρατίθενται στον Πίνακα 6.1, όπου δίνονται τόσο οι τιμές TEM όσο και οι κλιμακωτές τιμές TEM για κάθε μεμονωμένη μέτρηση. Επιπλέον, στον Πίνακα 6.1 παρατίθενται και οι μέσοι όροι των αποστάσεων και των γωνιών. Μια διαφορά που παρατηρούμε είναι ότι το TEM που σχετίζεται με τις αποστάσεις, είναι χαμηλότερο μεταξύ αυτών που έχουν υπολογιστεί από τα 3D μοντέλα και το Microscribe ή τα 3D μοντέλα και το παχύμετρο σε σύγκριση με τις αποστάσεις που έχουν υπολογιστεί από το Microscribe και το παχύμετρο (διαφορά: ~ 0,3 mm). Αντιθέτως, το TEM που σχετίζεται με τις γωνίες, είναι οριακά χαμηλότερο μεταξύ αυτών που έχουν υπολογιστεί από τα 3D μοντέλα και το παχύμετρο και αυτών που έχουν υπολογιστεί από Microscribe και το παχύμετρο. Παρ' όλα αυτά, δεν υπάρχει στατιστικά σημαντική διαφορά μεταξύ κάθε ζεύγους διαφορετικών μεθόδων μέτρησης σύμφωνα με τα αποτελέσματα της ανάλυσης διακύμανσης των TEM (αποτελέσματα ANOVA για σχετικό TEM: $F(2,18) = 0,126$, $p = 0,883$, $\eta^2 = 0,014$).

Πίνακας 6.1: Αποτελέσματα TEM μεταξύ κάθε ζεύγους διαφορετικών μεθόδων μέτρησης

	D1	D2	D3	D4	D5	D6	A2	A4	A5	A6	Μέσος όρος αποστάσεων	Μέσος όρος γωνιών
TEM ^α : 3d μοντέλα – Microscribe	1,26	1,43	2,42	1,98	1,46	1,28	0,93	2,98	2,21	0,91	1,64	1,75
Σχετικό TEM ^β : 3d μοντέλα – Microscribe	2,5	1,0	1,8	1,3	1,1	0,9	4,0	4,8	3,2	1,3	1,46	3,33
TEM ^α : 3d μοντέλα – παχύμετρο	1,33	1,01	1,86	3,07	0,88	1,62	0,83	2,82	1,43	1,41	1,62	1,62
Σχετικό TEM ^β : 3d μοντέλα – παχύμετρο	2,7	0,7	1,4	2,0	0,7	1,2	3,6	4,6	2,1	2,0	1,45	3,06
TEM ^α : Microscribe – παχύμετρο	0,93	1,18	2,51	3,52	1,28	2,18	0,81	2,56	2,11	0,98	1,93	1,61
Σχετικό TEM ^β : Microscribe – παχύμετρο	1,8	0,9	1,9	2,3	1,0	1,6	3,5	4,0	3,0	1,4	1,57	2,98
^α Αποστάσεις σε mm και γωνίες σε μοίρες ^β Οι τιμές δίνονται σε ποσοστά %												

Τα αποτελέσματα της ακρίβειας της απόκτησης αναφορικών σημείων από τα ψηφιακά μοντέλα για κάθε τύπο αναφορικών σημείων μεμονωμένα συνοψίζονται στην Εικόνα 6.1, η οποία απεικονίζει το σφάλμα μέτρησης μεταξύ των δεδομένων που προέκυψαν από το σύνολο των 50 κρανίων που ψηφιοποιήθηκαν με τη μέθοδο της φωτογραμμετρίας και το αντίστοιχο σετ δεδομένων

που είχε ψηφιοποιηθεί με το 3D Microscribe. Το LRMSD και η διάμεσος των απολύτων αποκλίσεων των ILD από κάθε ζεύγος δειγμάτων εμφανίζονται σε χωριστά θηκογράμματα.



Εικόνα 6.1: Σφάλμα μέτρησης μεταξύ των δεδομένων που προέκυψαν από τα 3D μοντέλα και το Microscribe.

Τα αναφορικά σημεία τύπου I μπορούν να ψηφιοποιηθούν με μεγαλύτερη ακρίβεια (μεταξύ μεθόδων: 1,1 mm, αναπαραγωγιμότητα: 0,47 mm). Τα αναφορικά σημεία τύπου II έδωσαν μέσο όρο LRMSD 1,4 mm μεταξύ των μεθόδων και τα αναφορικά σημεία τύπου III παρουσίασαν τη μικρότερη ακρίβεια (μεταξύ μεθόδων: 1,6 mm), ενώ και οι δύο τύποι αναφορικών σημείων είχαν ουσιαστικά την ίδια αναπαραγωγιμότητα (τύπος II: 0,71 mm · τύπος III: 0,72 mm). Αντίστοιχα αποτελέσματα έδωσαν και οι διάμεσοι των απολύτων αποκλίσεων των ILD αν και με σταθερά μικρότερη διαφορά ανάμεσα στο σφάλμα μέτρησης μεταξύ μεθόδων και στην αναπαραγωγιμότητα. Πιο συγκεκριμένα, το σφάλμα μέτρησης μεταξύ των μεθόδων κυμαινόταν από 0,74 έως 1,05 mm, ενώ η αναπαραγωγιμότητα ήταν μικρότερη από 0,5 mm σε όλες τις περιπτώσεις με τα αναφορικά σημεία τύπου I να αποδίδουν καλύτερα από τα υπόλοιπα.

6.3. Συμπεράσματα

Τα αποτελέσματα της παρούσας μελέτης έδειξαν ότι και οι τρεις μέθοδοι εμφανίζουν παρόμοιο σφάλμα μέτρησης σχετικά με τη λήψη γραμμικών μετρήσεων. Επιπλέον, όσον αφορά τις αποστάσεις, ο μέσος όρος των σχετικών TEM μεταξύ των μεθόδων είναι εντός των αποδεκτών ορίων σφάλματος αναπαραγωγιμότητας μεταξύ έμπειρων ερευνητών (Perini et al., 2005). Αυτό συνεπάγεται ότι ο συνδυασμός δεδομένων γραμμικών μετρήσεων, τα οποία συλλέγονται με

διαφορετικές μεθόδους θα μπορούσε να είναι μια βιώσιμη επιλογή όσον αφορά το σφάλμα μέτρησης που παρατηρείται μεταξύ μεθόδων. Τα αποτελέσματά μας έδειξαν ότι η διακύμανση του σφάλματος μέτρησης σχετίζεται περισσότερο με τη μεταβλητή παρά με τη μέθοδο μέτρησης.

Σχετικά με τη λήψη τριδιάστατων συντεταγμένων αναφορικών σημείων από ψηφιακά μοντέλα, η συγκεκριμένη έρευνα έδειξε ότι το σφάλμα αναπαραγωγιμότητας είναι μικρότερο (ο μέσος όρος ανά τύπο αναφορικών σημείων κυμαίνεται από 0,47mm έως 0,72mm) σε σύγκριση με τη λήψη τριδιάστατων συντεταγμένων αναφορικών σημείων με 3D Microscribe (μεσοσταθμικό σφάλμα αναπαραγωγιμότητας 0,87mm) (Chovaloroulou, Bertatos, et al., 2017). Το μειωμένο σφάλμα μπορεί εν μέρει να αποδοθεί στα πλεονεκτήματα του εντοπισμού αναφορικών σημείων σε ψηφιακά μοντέλα, όπου η ικανότητα περιστροφής και εστίασης σε ένα υψηλής ποιότητας μοντέλο, όσον αφορά την υφή, επιτρέπει πολύ πιο ακριβή τοποθέτηση αναφορικών σημείων. Το γεγονός αυτό γίνεται πιο εμφανές όταν πρόκειται για αναφορικά σημεία τύπου I. Παρόλα αυτά, τα αποτελέσματα της αναπαραγωγιμότητας υποδηλώνουν ότι και οι δύο μέθοδοι απόκτησης αναφορικών σημείων, δηλαδή το Microscribe-3DX και τα ψηφιακά μοντέλα με τη βοήθεια του λογισμικού Meshlab, εμφανίζουν τα ίδια επίπεδα ορθότητας και ακρίβειας με το τελευταίο ίσως να υπερτερεί.

Πολλοί ερευνητές αναφέρουν ότι οι τρεις τύποι αναφορικών σημείων εμφανίζουν διαφορετική αναπαραγωγιμότητα (Barbeito-Andrés et al., 2012; Ross & Williams, 2008; Williams & Richtsmeier, 2003). Τα αποτελέσματα της παρούσας έρευνας έδειξαν ότι τα αναφορικά σημεία τύπου I μπορούν να ψηφιοποιηθούν με την υψηλότερη ακρίβεια. Το γεγονός αυτό συμβαδίζει με προηγούμενα ευρήματα και οφείλεται κυρίως στην υφή υψηλής ποιότητας που είναι εγγενής στα ψηφιακά τριδιάστατα μοντέλα. Όπως και σε προηγούμενες μελέτες, τα αναφορικά σημεία τύπου III εμφάνισαν χαμηλότερη ακρίβεια από τον τύπο I, ωστόσο, η διαφορά στην αναπαραγωγιμότητα μεταξύ των αναφορικών σημείων τύπου II και τύπου III ήταν στατιστικά μη σημαντική τόσο για το LRMSD των διατάξεων των αναφορικών σημείων όσο και για τα ILDs. Από την άλλη πλευρά, η σύγκριση μεταξύ μεθόδων αποκάλυψε στατιστικά σημαντική διαφορά μεταξύ των αναφορικών σημείων τύπου II και τύπου III όσον αφορά το LRMSD των διατάξεων των αναφορικών σημείων, γεγονός που υποδηλώνει ότι τα αναφορικά σημεία τύπου III είναι πιο δύσκολο να ψηφιοποιηθούν με το Microscribe. Αυτό μπορεί να σχετίζεται με τον περιορισμό των τριδιάστατων ψηφιοποιητών ακίδας που απαιτούν τα δείγματα να είναι στατικά καθ' όλη την απόκτηση αναφορικών σημείων, γεγονός που συχνά εμποδίζει την ικανότητα του παρατηρητή να εντοπίζει με ακρίβεια ακραία σημεία όπως αναφορικά σημεία τύπου III. Ωστόσο, αυτός ο περιορισμός δεν φαίνεται να εισάγει περαιτέρω σφάλμα μέτρησης κατά τη χρήση ILDs.

7. Έλεγχος αξιοπιστίας του λογισμικού 3D-ID στην εκτίμηση του φύλου και της καταγωγής με την εφαρμογή σύγχρονου Ελληνικού πληθυσμού.

Η δικαστική ανθρωπολογία μπορεί να οριστεί ως η εφαρμογή της σκελετικής βιολογίας σε νομικά ζητήματα (Komar & Buikstra, 2008). Όταν εντοπίζονται ανθρώπινα υπολείμματα, ο πρωταρχικός στόχος σε περιπτώσεις ατόμων αγνώστων στοιχείων είναι η ταυτοποίηση του βιολογικού προφίλ τους. Η εκτίμηση του φύλου, της καταγωγής, της ηλικίας και του αναστήματος, είναι θεμελιώδης για την ταυτοποίηση του βιολογικού προφίλ ενός ατόμου. Ωστόσο, πρώτα διεξάγεται ο προσδιορισμός του φύλου και της καταγωγής, δεδομένου ότι η εκτίμηση του αναστήματος και της ηλικίας εξαρτώνται από τις δύο προαναφερθείσες παραμέτρους (Ross et al., 2010).

Η ανάπτυξη πολυπαραγοντικών στατιστικών μεθόδων σε συνδυασμό με την αυξανόμενη διαθεσιμότητα μεγάλων σκελετικών συλλογών αναφοράς επέτρεψε την ποσοτική ανάλυση προσδιορισμού του φύλου και της καταγωγής μεταξύ διαφορετικών ομάδων πληθυσμού. Αυτό με τη σειρά του οδήγησε στην ανάπτυξη εξειδικευμένων λογισμικών για την εκτίμηση του φύλου και της καταγωγής αγνώστων στοιχείων κρανιακών υπολειμμάτων, όπως το 3D-ID (Slice & Ross, 2018), το FORDISC 3 (Jantz & Ousley, 2013) και το CRAN-ID (Wright, 2012). Παρόλο που οι ποσοτικές μέθοδοι παρουσιάζουν μεγαλύτερη αντικειμενικότητα σε σύγκριση με τις ποιοτικές (Walker, 2008), η απόδοση αυτών των προγραμμάτων λογισμικού όσον αφορά την επιτυχή ταξινόμηση εξακολουθεί να είναι υπό αμφισβήτηση, με διάφορους ερευνητές να αναφέρουν κακή έως μέτρια απόδοση σε περιπτώσεις όπου το υπό μελέτη δείγμα ανήκει σε πληθυσμιακή ομάδα (**καταγωγή**), η οποία δεν περιλαμβάνεται στη βάση δεδομένων αναφοράς του λογισμικού (Elliott & Collard, 2009; Guyomarc'h & Bruzek, 2011; Ramsthaler et al., 2007; Urbanová et al., 2014). Η εσφαλμένη ταξινόμηση της καταγωγής μπορεί να οδηγήσει με τη σειρά της σε εσφαλμένο προσδιορισμό φύλου, καθώς διάφοροι πληθυσμοί εμφανίζουν διαφορετικά πρότυπα φυλετικού διμορφισμού (Steyn & Patriquin, 2009).

Το 3D-ID λογισμικό είναι το μόνο που χρησιμοποιεί τριδιάστατη γεωμετρική μορφομετρία για την εκτίμηση του φύλου και της καταγωγής άγνωστων ατόμων με βάση την κρανιακή μορφολογία τους. Επιπλέον, το 3D-ID είναι ένα πακέτο πολλαπλών πλατφορμών και είναι το μόνο ελεύθερα διαθέσιμο πρόγραμμα που μπορεί να χειριστεί ελλείπουσες τιμές στα δεδομένα εισόδου. Και οι δύο προαναφερθείσες δυνατότητες παρέχουν σημαντικά πλεονεκτήματα στον χρήστη. Στόχος της παρούσας εργασίας ήταν να αξιολογήσουμε την αξιοπιστία της τελευταίας έκδοσης του λογισμικού

3D-ID (έκδοση 2018) σχετικά με την εκτίμηση του φύλου και της καταγωγής χρησιμοποιώντας διατάξεις αναφορικών σημείων από ένα δείγμα ελληνικού πληθυσμού για το οποίο το φύλο και η καταγωγή των ατόμων είναι γνωστή από τα πιστοποιητικά θανάτου. Το λογισμικό 3D-ID δεν περιλαμβάνει δείγματα ελληνικού πληθυσμού στη βάση δεδομένων αναφοράς του, γεγονός που εμποδίζει την ικανότητά του στον επιτυχή προσδιορισμό της καταγωγής. Ωστόσο, στη βάση δεδομένων αναφοράς του περιλαμβάνεται δείγματα της νοτιοανατολικής Ευρώπης, τα οποία γεωγραφικά, τουλάχιστον, αντιπροσωπεύουν τον υπό μελέτη πληθυσμό.

7.1. Υλικά και μέθοδοι

Για τις ανάγκες της παρούσας εργασίας χρησιμοποιήθηκαν 158 κρανία (80 άνδρες και 78 γυναίκες) ενηλίκων ατόμων από τη σύγχρονη συλλογή αναφοράς. Από τα ψηφιακά μοντέλα με υφή υψηλής ποιότητας, ελήφθησαν οι τριδιάστατες συντεταγμένες είκοσι πέντε αναφορικών σημείων με την χρήση του λογισμικού Meshlab. Προκειμένου να μην επηρεαστούν τα αποτελέσματα της ταξινόμησης του 3D-ID από τυχόν εσφαλμένο εντοπισμό αναφορικών σημείων στα κρανία, η ακεραιότητα όλων των δεδομένων ελέγχθηκε με προσαρμοσμένο αλγόριθμο σε περιβάλλον GNU Octave. Τα αρχεία ".3did" που περιέχουν τις διατάξεις των αναφορικών σημείων κάθε δείγματος επίσης δημιουργήθηκαν με το GNU Octave.

Στην παρούσα μελέτη, η απόδοση του λογισμικού 3D-ID αξιολογήθηκε τόσο όσον αφορά στον προσδιορισμό του φύλου και της καταγωγής συλλογικά όσο και μεμονωμένα. Επιπλέον, η αξιοπιστία του λογισμικού 3D-ID ελέγχθηκε χρησιμοποιώντας μεταβλητές και σχήματος και μορφής. Οι προαναφερθέντες έλεγχοι πραγματοποιήθηκαν σε δύο διαφορετικούς τύπους δεδομένων όσον αφορά τις διατάξεις των αναφορικών σημείων του κάθε δείγματος. Η πρώτη περίπτωση περιελάμβανε διατάξεις 25 αναφορικών σημείων, όπως προτείνεται και από τους συγγραφείς του λογισμικού 3D-ID, ενώ η δεύτερη περίπτωση περιελάμβανε διατάξεις 19 αναφορικών σημείων. Και οι δύο τύποι δεδομένων χρησιμοποιήθηκαν για την αξιολόγηση της ταξινόμησης του φύλου και της καταγωγής συλλογικά (χρησιμοποιώντας την πλήρη βάση δεδομένων αναφοράς), αλλά και της καταγωγής μεμονωμένα, δηλώνοντας το φύλο του κάθε ατόμου (χρησιμοποιώντας την καθορισμένη από το φύλο βάση δεδομένων αναφοράς). Οι διατάξεις των 19 αναφορικών σημείων χρησιμοποιήθηκαν προκειμένου να αυξηθεί το μέγεθος δείγματος αναφοράς των ευρωπαϊκών πληθυσμών σύμφωνα με τις οδηγίες του 3D-ID. Επιπλέον, χρησιμοποιήθηκαν για την αξιολόγηση της απόδοσης του λογισμικού 3D-ID όσον αφορά στον προσδιορισμό του φύλου και της καταγωγής συλλογικά χρησιμοποιώντας μόνο ευρωπαϊκά δείγματα αναφοράς, αλλά και στον προσδιορισμό του φύλου μεμονωμένα καθορίζοντας ότι τα υπό μελέτη δείγματα είναι νοτιοανατολικής ευρωπαϊκής καταγωγής.

Η απόδοση του λογισμικού 3D-ID υπολογίστηκε ως το ποσοστό των επιτυχώς ταξινομημένων δειγμάτων. Επιπλέον, πρέπει να σημειωθεί ότι η ακρίβεια στον προσδιορισμό της καταγωγής του δείγματος του ελληνικού πληθυσμού αξιολογήθηκε τόσο ως προς την ταξινόμησή του στην Νοτιοανατολική Ευρώπη όσο και στην Ευρώπη συνολικά.

7.2. Αποτελέσματα

Τα αποτελέσματα ταξινόμησης, όταν χρησιμοποιείται η πλήρης βάση δεδομένων αναφοράς του 3D-ID, συνοψίζονται στον Πίνακα 7.1. Σχετικά με τον προσδιορισμό της καταγωγής και την ταξινόμηση του ελληνικού δείγματος στην Νοτιοανατολική Ευρώπη, η συνολική ακρίβεια τόσο για μεταβλητές σχήματος όσο και για τις μεταβλητές μορφής βελτιώνεται όταν χρησιμοποιείται η διάταξη των 19 αναφορικών σημείων, καθώς με τη διάταξη των 25 αναφορικών σημείων κανένα δείγμα δεν ταξινομήθηκε επιτυχώς. Κατά την αξιολόγηση της ταξινόμησης του ελληνικού δείγματος στην Ευρώπη συνολικά, η διάταξη των 25 αναφορικών σημείων εμφάνισε ποσοστό 61,4% επιτυχούς ταξινόμησης χρησιμοποιώντας μεταβλητές σχήματος και 63,9% χρησιμοποιώντας μεταβλητές μορφής. Ωστόσο, η διάταξη των 19 αναφορικών σημείων εξακολουθεί να παρουσιάζει υψηλότερα ποσοστά επιτυχούς ταξινόμησης τόσο για τις μεταβλητές σχήματος όσο και για τις μεταβλητές μορφής (70,3% και 70,9 % αντίστοιχα). Αξίζει να σημειωθεί πως οι μεταβλητές μορφής εμφανίζουν οριακά υψηλότερα ποσοστά επιτυχούς ταξινόμησης σε σύγκριση με τις μεταβλητές σχήματος. Σχετικά με τον προσδιορισμό του φύλου, τα αποτελέσματα ήταν ανάλογα με αυτά της καταγωγής με την έννοια ότι τόσο οι διατάξεις των 19 αναφορικών σημείων όσο και οι μεταβλητές μορφής εμφάνισαν υψηλότερα ποσοστά επιτυχούς ταξινόμησης σε σχέση με τις διατάξεις των 25 αναφορικών σημείων και τις μεταβλητές σχήματος αντίστοιχα. Τέλος, όσον αφορά στον προσδιορισμό του φύλου και της καταγωγής συλλογικά, όταν όλες οι ευρωπαϊκές πληθυσμιακές ομάδες θεωρούνται ως ενιαία ομάδα, τα ποσοστά επιτυχούς ταξινόμησης αυξάνονται με τη χρήση των διατάξεων των 19 αναφορικών σημείων τόσο για τις μεταβλητές σχήματος όσο και για τις μεταβλητές μορφής.

Πίνακας 7.1: Αποτελέσματα απόδοσης λογισμικού σχετικά με τον προσδιορισμό του φύλου και της καταγωγής συλλογικά χρησιμοποιώντας την πλήρη βάση δεδομένων αναφοράς.

Πληθυσμιακή ομάδα ^α	Διάταξη 25 αναφορικών σημείων		Διάταξη 19 αναφορικών σημείων	
	Μεταβλητές σχήματος	Μεταβλητές μορφής	Μεταβλητές σχήματος	Μεταβλητές μορφής
Αφρικανική	0	0	0	1 (0,6)
Αφρικανική_Αμερικανική	0	0	2 (1,3)	3 (1,9)
Αφρικανική_Βραζιλιάνικη	0	0	1 (0,6)	1 (0,6)
Βραζιλιάνικη	0	0	11 (7,0)	8 (5,1)
Καραϊβική	0	0	3 (1,9)	5 (3,2)
Κολομβιανή	28 (17,7)	25 (15,8)	11 (7,0)	8 (5,1)
Ασιατική_Ανατολική	7 (4,4)	6 (3,8)	1 (0,6)	1 (0,6)
Ασιατική_Αφρικανική	0	0	0	0
Ευρωπαϊκή_Αμερικανική	10 (6,3)	10 (6,3)	8 (5,1)	10 (6,3)
Ευρωπαϊκή_Κεντρική	0	0	43 (27,2)	43 (27,2)
Ευρωπαϊκή_Ανατολική	0	0	0	0
Ευρωπαϊκή_Νοτιοανατολική	0	0	54 (34,2)	54 (34,2)
Ευρωπαϊκή_Νοτιοδυτική	97 (61,4)	101 (63,9)	14 (8,9)	15 (9,5)
Ιαπωνική_Βραζιλιάνικη	0	0	1 (0,6)	1 (0,6)
Μεσοαμερικανική	0	0	1 (0,6)	1 (0,6)
Νιγηριανή	0	0	0	0
Αμερικανική_Νότια	11 (7,0)	12 (7,6)	7 (4,4)	6 (3,8)
Συριακή	5 (3,2)	4 (2,5)	1 (0,6)	1 (0,6)
Αφρικανική_Δυτική	0	0	0	0
Ταξινόμηση φύλου ^β	117 (74,1)	126 (79,7)	120 (75,9)	135 (85,4)

^α Οι ακέραιοι αριθμοί αντιστοιχούν στον αριθμό ατόμων που ταξινομήθηκαν στην εκάστοτε πληθυσμιακή ομάδα του δείγματος αναφοράς. Οι τιμές στις παρενθέσεις δείχνουν το αντίστοιχο ποσοστό ταξινόμησης του συνολικού υπό μελέτη δείγματος (158 άτομα).

^β Ο αριθμός επιτυχώς ταξινομημένων ατόμων (και το αντίστοιχο ποσοστό τους) αφορά μόνο τον σωστό προσδιορισμό φύλου ανεξαρτήτως της ταξινόμησης σε πληθυσμιακή ομάδα.

Τα αποτελέσματα σχετικά με τον προσδιορισμό της καταγωγής μεμονωμένα, δηλώνοντας δηλαδή το φύλο του κάθε ατόμου και χρησιμοποιώντας την καθορισμένη από το φύλο βάση δεδομένων αναφοράς, παρουσιάζονται στον Πίνακα 7.2 και είναι παρόμοια με αυτά του Πίνακα 7.1 σε σύγκριση με τα ποσοστά επιτυχούς ταξινόμησης του ελληνικού πληθυσμού σε ευρωπαϊκές ομάδες. Η μόνη αξιοσημείωτη διαφορά έγκειται στο διαφορετικό μοτίβο που παρατηρείται στα εσφαλμένα καταχωρημένα δείγματα.

Η απόδοση του λογισμικού βελτιώνεται τόσο στον προσδιορισμό του φύλου όσο και της καταγωγής όταν η ανάλυση περιέχει μόνο ευρωπαϊκά δείγματα αναφοράς. Πιο συγκεκριμένα, όσον αφορά στον προσδιορισμό της καταγωγής, το 43,7% του ελληνικού δείγματος ταξινομείται σωστά στην Νοτιοανατολική Ευρώπη όταν χρησιμοποιούνται μεταβλητές σχήματος και το αντίστοιχο ποσοστό επιτυχούς ταξινόμησης όταν χρησιμοποιούνται μεταβλητές μορφής είναι 46,8% (Πίνακας

7.2). Σχετικά με τον προσδιορισμό φύλου, το ποσοστό επιτυχούς ταξινόμησης για τις μεταβλητές σχήματος είναι 76,6% και για τις μεταβλητές μορφής είναι 86,7%. Τέλος, η χρήση των διατάξεων των 19 αναφορικών σημείων για τον στον προσδιορισμό του φύλου μεμονωμένα, καθορίζοντας δηλαδή ότι τα υπό μελέτη δείγματα είναι νοτιοανατολικής ευρωπαϊκής καταγωγής, εμφάνισε ποσοστά επιτυχούς ταξινόμησης 79,95% για τις μεταβλητές σχήματος και 89,87% για τις μεταβλητές μορφής.

Πίνακας 7.2: Αποτελέσματα απόδοσης λογισμικού σχετικά με τον προσδιορισμό του φύλου και της καταγωγής συλλογικά χρησιμοποιώντας μόνο ευρωπαϊκά δείγματα αναφοράς.

Διάταξη 19 αναφορικών σημείων		
Πληθυσμιακή ομάδα ^α	Μεταβλητές σχήματος	Μεταβλητές μορφής
Ευρωπαϊκή_Κεντρική	51 (32,3)	48 (30,4)
Ευρωπαϊκή_Ανατολική	0	0
Ευρωπαϊκή_Νοτιοανατολική	69 (43,7)	74 (46,8)
Ευρωπαϊκή_Νοτιοδυτική	38 (24,1)	36 (22,8)
Ταξινόμηση φύλου ^β	121 (76,6)	137 (86,7)

^α Οι ακέραιοι αριθμοί αντιστοιχούν στον αριθμό ατόμων που ταξινομήθηκαν στην εκάστοτε πληθυσμιακή ομάδα του δείγματος αναφοράς. Οι τιμές στις παρενθέσεις δείχνουν το αντίστοιχο ποσοστό ταξινόμησης του συνολικού υπό μελέτη δείγματος (158 άτομα).

^β Ο αριθμός επιτυχώς ταξινομημένων ατόμων (και το αντίστοιχο ποσοστό τους) αφορά μόνο τον σωστό προσδιορισμό φύλου ανεξαρτήτως της ταξινόμησης σε πληθυσμιακή ομάδα.

7.3. Συμπεράσματα

Συνολικά, η απόδοση του λογισμικού ήταν μέτρια. Το 3D-ID έχει σχεδιαστεί έτσι ώστε να κατατάσσει οποιοδήποτε άγνωστο δείγμα στην πληθυσμιακή ομάδα από την οποία διαφέρει λιγότερο (Slice & Ross, 2018). Ως εκ τούτου, η κακή έως μέτρια απόδοση του 3D-ID όσον αφορά στον προσδιορισμό της καταγωγής μπορεί να αποδοθεί στο γεγονός ότι δεν συμπεριλαμβάνεται αντιπροσωπευτικό δείγμα αναφοράς του ελληνικού πληθυσμού στη βάση δεδομένων του 3D-ID. Μια παρατήρηση, η οποία αξίζει αναφοράς, είναι ότι η εκ των προτέρων γνώση του φύλου ενός ατόμου δεν βελτιώνει την απόδοση του λογισμικού στον προσδιορισμό της καταγωγής. Τα αποτελέσματα της παρούσας εργασίας ανέδειξαν επίσης ότι οι διατάξεις των 25 και 19 αναφορικών σημείων εμφανίζουν διαφορετικά μοτίβα εσφαλμένης ταξινόμησης σχετικά με την καταγωγή. Λαμβάνοντας υπόψη ότι το 1/6 του δείγματος του ελληνικού πληθυσμού ταξινομείται ως Κολομβιανό, όταν στην ανάλυση χρησιμοποιούνται οι διατάξεις των 25 αναφορικών σημείων, θα μπορούσε κανείς να υποστηρίξει ότι αυτό είναι αποτέλεσμα της φαινοτυπικής βιολογικής ομοιότητας σε κάποιο βαθμό μεταξύ των δύο πληθυσμών. Ωστόσο, όταν στην ανάλυση χρησιμοποιούνται οι διατάξεις των 19 αναφορικών σημείων τα αποτελέσματα διαφέρουν. Κατά

συνέπεια, εγείρεται το ερώτημα εάν μια τέτοιου είδους φαινοτυπική βιολογική ομοιότητα θα μπορούσε να είναι αποτέλεσμα της συγκεκριμένης διάταξης αναφορικών σημείων ή απλώς ένα τεχνικό σφάλμα κατά τη γεωμετρική μορφομετρική ανάλυση, η οποία περιλαμβάνει διατάξεις με μεταβλητό αριθμό αναφορικών σημείων. Σύμφωνα με τις οδηγίες του 3D-ID, η χρήση των διατάξεων των 19 αναφορικών σημείων για την αύξηση της διαθεσιμότητας του μεγέθους δείγματος αναφοράς αφορούν κυρίως τους πληθυσμούς της Ευρώπης και της Βραζιλίας, γεγονός που υποδηλώνει ότι η εσφαλμένη ταξινόμηση σε άλλες πληθυσμιακές ομάδες δεν μπορεί να αποδοθεί στο μικρό μέγεθος του δείγματος αναφοράς. Θα μπορούσε το ίδιο υπό μελέτη δείγμα να ταξινομηθεί διαφορετικά με διαφορετική διάταξη αναφορικών σημείων; Εάν ναι, ποια συγκεκριμένη διάταξη αναφορικών σημείων πρέπει να χρησιμοποιείται;

Θα μπορούσε κανείς να υποστηρίξει ότι ο έλεγχος της αξιοπιστίας του 3D-ID, γνωρίζοντας ότι ο υπό μελέτη πληθυσμός δεν εκπροσωπείται στο δείγμα αναφοράς του λογισμικού, δεν είναι έγκυρος σχεδιασμός μελέτης. Ωστόσο, στόχος της παρούσας μελέτης ήταν να αξιολογήσει την απόδοση του 3D-ID υπό την σκοπιά της χρήσης του ως εργαλείο στην δικαστική ανθρωπολογία, όπου δεν υπάρχει πρότερη γνώση εάν το δείγμα εκπροσωπείται ή όχι στο δείγμα αναφοράς του λογισμικού.

Σχετικά με τον προσδιορισμό του φύλου, η επιτυχής ταξινόμηση του 3D-ID αυξάνεται όταν χρησιμοποιούνται μεταβλητές μορφής αντί για μεταβλητές σχήματος, γεγονός που συμφωνεί με το επιχείρημα ότι η πλειονότητα των διαφορών μεταξύ ανδρών και γυναικών σχετίζεται με το μέγεθος (Lynnerup, 2013). Σύμφωνα με τα αποτελέσματα της συγκεκριμένης μελέτης, το ποσοστό ακρίβειας προσδιορισμού του φύλου ήταν μόλις 76,6% με τη χρήση μεταβλητών σχήματος, το οποίο είναι αρκετά χαμηλό για την δικαστική ανθρωπολογία. Επομένως, η βέλτιστη πρακτική είναι η χρήση μεταβλητών μορφής και η διάταξη των 19 αναφορικών σημείων ανεξαρτήτως καταγωγής του δείγματος, καθώς οι αναλύσεις με αυτού του τύπου τα δεδομένα απέδωσαν περίπου 86% ποσοστό επιτυχούς ταξινόμησης. Επιπλέον, ίδιο ποσοστό επιτυχούς ταξινόμησης φύλου (86,7%) έχει βρεθεί και σε προηγούμενη έρευνα, η οποία εφάρμοσε τριδιάστατη γεωμετρική μορφομετρική ανάλυση στην περιοχή του άνω προσώπου στο ίδιο δείγμα Ελληνικού πληθυσμού (Choualoroulou et al., 2016b). Το γεγονός αυτό, υποδηλώνει ότι το ελληνικό πληθυσμιακό δείγμα αντιπροσωπεύεται καλά από τις ευρωπαϊκές ομάδες της βάσης δεδομένων αναφοράς του 3D-ID, τουλάχιστον όσον αφορά την έκφραση του φυλετικού διμορφισμού. Ωστόσο, η πρότασή μας είναι ότι ο προσδιορισμός του φύλου θα πρέπει, όπου είναι εφικτό, να διασταυρώνεται με μεθόδους, οι οποίες είναι πληθυσμιακά εξειδικευμένες.

8. Προσδιορισμός φύλου του σύγχρονου ελληνικού πληθυσμού από διαστάσεις της κάτω γνάθου.

Η κάτω γνάθος εμφανίζει φυλετικό διμορφισμό από τα πρώιμα στάδια της ζωής (Kano et al., 2015). Πιο συγκεκριμένα, προηγούμενες μελέτες έχουν δείξει ότι από την ηλικία των πέντε μηνών η έκφραση του φυλετικού είναι εμφανής (Loth & Henneberg, 2001; Scheuer, 2002; Schutkowski, 1993). Ωστόσο, ο φυλετικός διμορφισμός της κάτω γνάθου μειώνεται κατά την παιδική ηλικία (μεταξύ των ηλικιών 4 και 14), ενώ γίνεται και πάλι πιο έντονος κατά την ενηλικίωση (Coquerelle et al., 2011). Επιπλέον, έχει αποδειχθεί ότι τα μορφολογικά χαρακτηριστικά που σχετίζονται με το φύλο εμφανίζουν ποικιλομορφία τόσο εντός του ίδιου πληθυσμού όσο και μεταξύ πληθυσμών, η οποία επηρεάζεται από μια σειρά παραγόντων όπως γενετικοί, διατροφικοί, αναπτυξιακοί, οι διαχρονικές τάσεις, οι ασθένειες καθώς και άλλοι γενικοί περιβαλλοντικοί παράγοντες (Ubelaker & DeGaglia, 2017).

Η διακριτική ανάλυση είναι απλή στην εφαρμογή της και ταυτόχρονα παρέχει ακριβή ταξινόμηση για δεδομένα τα οποία ακολουθούν κανονική κατανομή (Dixon & Brereton, 2009). Ως εκ τούτου, χρησιμοποιείται ευρέως στην δικαστική ανθρωπολογία για την ανάπτυξη διακριτικών εξισώσεων προσδιορισμού του φύλου (Dayal et al., 2008; Franklin et al., 2005; Ogawa et al., 2013). Μέχρι σήμερα, δεν είχε πραγματοποιηθεί μελέτη κατασκευής διακριτικών εξισώσεων εκτίμησης φύλου με βάση την κάτω γνάθο για τον ελληνικό πληθυσμό. Η παρούσα εργασία βασίστηκε σε αντιπροσωπευτικό δείγμα του σύγχρονου ελληνικού πληθυσμού και στόχος της ήταν η εύρεση των κατάλληλων συνδυασμών μορφομετρικών μεταβλητών για τη δημιουργία διακριτικών εξισώσεων για εκτίμηση του φύλου από την κάτω γνάθο. Επιπλέον, η εργασία αυτή εξέτασε αν και σε ποιο βαθμό κάποιες από τις υφιστάμενες διακριτικές εξισώσεις από διαφορετικούς πληθυσμούς αναφοράς που αναφέρονται στη διεθνή βιβλιογραφία μπορούν να διαχωρίσουν επιτυχώς τα αρσενικά από τα θηλυκά άτομα στον ελληνικό πληθυσμό. Αυτό αποσκοπούσε στην αξιολόγηση της πληθυσμιακής εξειδίκευσης των υφιστάμενων διακριτικών εξισώσεων και κατ' επέκταση τη διερεύνηση πιθανών ομοιοτήτων και διαφορών στην έκφραση του φυλετικού διμορφισμού της κάτω γνάθου μεταξύ διαφορετικών πληθυσμιακών ομάδων.

8.1. Υλικά και μέθοδοι

Για τις ανάγκες της παρούσας εργασίας χρησιμοποιήθηκαν 194 κάτω γνάθοι, 105 αρσενικών και 89 θηλυκών ενήλικων ατόμων, οι οποίες ήταν ακέραιες και δεν εμφάνιζαν κάποια παθολογικής φύσεως μορφολογική αλλοίωση. Οι γνάθοι είχαν προηγουμένως ψηφιοποιηθεί με τριδιάστατες φωτογραμμετρικές τεχνικές. Από τα τριδιάστατα μοντέλα ελήφθησαν οι τριδιάστατες

συντεταγμένες από 16 αναφορικά σημεία (3 στο οβελιαίο επίπεδο και 13 αμφίπλευρα) με τη χρήση του προγράμματος Meshlab. Εν συνεχεία, από τα αναφορικά σημεία υπολογίστηκαν συνολικά 23 μετρήσεις από τις οποίες 3 αφορούν γωνίες και 11 αφορούν αποστάσεις για την επακόλουθη ανάλυση (Πίνακας 8.1). Οι αμφίπλευρες μετρήσεις υπολογίστηκαν ξεχωριστά για κάθε πλευρά της κάτω γνάθου.

Πίνακας 8.1: Μετρήσεις που χρησιμοποιήθηκαν στην εργασία

Μετρήσεις	Συντομογραφία	Ορισμοί
<i>Bi-condylar breadth</i>	BicBR	Η απόσταση μεταξύ δεξιού και αριστερού προεξέχοντος εξωτερικού σημείου (cdl – cdl).
<i>Bi-gonion breadth</i>	BigBR	Η απόσταση που ορίζεται από το δεξί και αριστερό γόνιο (go – go).
<i>Bi-coronoid breadth</i>	BicrBR	Η απόσταση που ορίζεται από το δεξί και αριστερό κορόνιο (co – co).
<i>Symphysis height</i>	SyH	Η απόσταση που ορίζεται από το γνάθιο και το σημείο ανάμεσα στους δύο κεντρικούς κοπτήρες (gn – id).
<i>Bi-mental foramina breadth</i>	BimfBR	Η απόσταση μεταξύ δεξιού και αριστερού γενειακού τρήματος (mf – mf).
<i>Symphysis breadth</i>	SyBR	Η απόσταση που ορίζεται από το αριστερό και δεξί <i>lateral ifradentale</i> (lid – lid)
<i>Bi-lingula breadth</i>	BilBR	Η απόσταση που ορίζεται από το αριστερό και δεξί <i>lingula mandibulae</i> (lm – lm).
<i>Mandible breadth</i>	MB	Η απόσταση που ορίζεται από το αριστερό και δεξί <i>anterior ramus</i> (ar – ar)
<i>Bi-notch breadth</i>	BinBR	Η απόσταση μεταξύ δεξιάς και αριστερής γναθιαίας εντομής(mn – mn)
<i>Angle formed by bilateral gn-go lines</i>	Agn-go	Η γωνία που σχηματίζεται από το δεξί γόνιο, γνάθιο και το αριστερό γόνιο (go – gn - go)
<i>Angle formed by bilateral gn-co lines</i>	Agn-co	Η γωνία που σχηματίζεται από το δεξί κορόνιο, γνάθιο και το αριστερό κορόνιο (co – gn – co)
<i>Angle formed by bilateral gn-cdl lines</i>	Agn-cdl	Η γωνία που σχηματίζεται από το δεξί προεξέχον εξωτερικό σημείο του κονδύλου, το γνάθιο και το αριστερό προεξέχον σημείο του κονδύλου (cdl – gn – cdl).
<i>Ramus height</i>	RH	Η απόσταση μεταξύ του υψηλότερου σημείου στον κόνδυλο και του γόνιου (cs – go)
<i>Maximum Ramus Height</i>	MRH	Η απόσταση μεταξύ του υψηλότερου σημείου στον κόνδυλο και του <i>mandibular tubercle</i> (cs – mtb)
<i>Projective Height of Ramus</i>	PRH	Η προβολή του ύψους του κλάδου ανάμεσα στο υψηλότερο σημείο του <i>mandibular capitulum</i> και στο χαμηλότερο όριο του οστού.
<i>Coronoid Height</i>	CH	Η κατακόρυφη προβολή ανάμεσα στο κορόνιο και στη βάση του οστού
<i>Maximum Ramus Breadth</i>	MaxRBR	Η απόσταση μεταξύ του πιο μπροστινού σημείου του κλάδου και της γραμμής που συνδέει το πιο οπίσθιο σημείο του κονδύλου με τη γωνία της κάτω γνάθου.
<i>Minimum Ramus Breadth</i>	MinRBR	Η μικρότερη διάμετρος από την πρόσθια ως την οπίσθια πλευρά του κλάδου (ar-pr).
<i>Notch length</i>	NL	Η απόσταση που ορίζεται από το κορόνιο και το υψηλότερο σημείο στον κόνδυλο (co – cs).
<i>Coronion-Gonion length</i>	CGL	Η απόσταση που ορίζεται από το κορόνιο και το γόνιο (co – go).
<i>Maximum Body length</i>	MaxBL	Η απόσταση που ορίζεται από το γόνιο και το πογόνιο (go – pg).
<i>Body length</i>	BL	Η απόσταση που ορίζεται από το γόνιο και το γνάθιο (go – gn).
<i>Maximum Mandible length</i>	MaxML	Η απόσταση ανάμεσα στο υψηλότερο σημείο του κονδύλου και στο πογόνιο (cs – pg).

Η ύπαρξη φυλετικού διμορφισμού διερευνήθηκε μέσω της ανάλυσης διακύμανσης (ANOVA), ενώ ο εντοπισμός εκείνων των μεταβλητών που έχουν τη μεγαλύτερη διαχωριστική ικανότητα για την επιτυχή εκτίμηση του φύλου σε ελληνικό πληθυσμό πραγματοποιήθηκε με μονομεταβλητή διακριτική ανάλυση. Στη συνέχεια, στις μεταβλητές που εμφάνισαν διασταυρούμενη επικύρωση άνω του 75% εφαρμόστηκε πολυπαραγοντική διακριτική ανάλυση τόσο με την κλιμακωτή μέθοδο επιλογής ανεξάρτητων μεταβλητών όσο και με τη μέθοδο χρήσης όλων των ανεξάρτητων μεταβλητών προκειμένου να εντοπιστεί ο βέλτιστος συνδυασμός τους για την εκτίμηση φύλου. Για να ερευνηθεί εάν και σε πιο βαθμό οι διακριτικές εξισώσεις εκτίμησης φύλου με βάση την κάτω γνάθο εμφανίζουν πληθυσμιακή εξειδίκευση, αντίστοιχες διακριτικές εξισώσεις βασιζόμενες σε δείγματα από διαφορετικούς πληθυσμούς αναφοράς κατασκευάστηκαν και για τον ελληνικό πληθυσμό με βάση το υπό μελέτη δείγμα αναφοράς και συγκρίθηκαν τα ποσοστά διασταυρούμενης επικύρωσης. Επιπλέον, έγινε και εφαρμογή των υφιστάμενων διακριτικών εξισώσεων στις γνάθους του ελληνικού δείγματος.

8.2. Αποτελέσματα

Τα αρσενικά άτομα σε σχέση με τα θηλυκά, εμφάνισαν υψηλότερες μέσες τιμές όσον αφορά στις αποστάσεις και χαμηλότερες όσον αφορά στις γωνίες με μοναδική εξαίρεση τη γωνία Agn-go. Η μονομεταβλητή διακριτική ανάλυση αποκάλυψε ότι μόνο πέντε μεταβλητές εμφάνισαν ποσοστό επιτυχούς ταξινόμησης πάνω από 75%. Πρόκειται για τις μεταβλητές RH, MRH, CH, CGL και MaxML της δεξιάς πλευράς, οι οποίες στη συνέχεια χρησιμοποιήθηκαν σε πολυπαραγοντική διακριτική ανάλυση.

Οι Πίνακες 8.2 και 8.3 περιλαμβάνουν τα αποτελέσματα της πολυπαραγοντικής διακριτικής ανάλυσης με τους συνδυασμούς μεταβλητών των διακριτικών εξισώσεων της βιβλιογραφίας καθώς και με τους συνδυασμούς μεταβλητών που προκύπτουν από τις μονομεταβλητές DFA. Οι πρώτες εννέα διακριτικές εξισώσεις σε κάθε πίνακα σχετίζονται με αντίστοιχες εξισώσεις που υπάρχουν σε προηγούμενες δημοσιευμένες μελέτες. Οι υπόλοιπες διακριτικές εξισώσεις αφορούν αυτές που κατασκευάστηκαν με την κλιμακωτή μέθοδο επιλογής ανεξάρτητων μεταβλητών και τη μέθοδο χρήσης όλων των ανεξάρτητων μεταβλητών των πιο φυλετικά διμορφικών μεταβλητών που εντοπίστηκαν στο δείγμα του ελληνικού πληθυσμού.

Πίνακας 8.2: Διακριτικές εξισώσεις βασισμένες στις μετρήσεις της δεξιάς πλευράς της κάτω γνάθου.

Εξίσωση	Wilks' Lambda	Sig.	Μέσες τιμές και σημεία τομής	Ορθώς κατανομημένα	
				Σύνολο δειγματος	Διασταυρωμένη επικύρωση
Εξίσωση 1^α	0,571	<0,001	♀ -0,945 \ [-0,0795] \ ♂ 0,786	83,20%	81,60%
(BicBR x 0,022) + (BigBR x 0,063) + (MinRBR x -0,026) + (BL x -0,043) + (MaxML x 0,183) -26,447					
Εξίσωση 2^β	0,492	<0,001	♀ -1,105 \ [-0,091] \ ♂ 0,923	85,20%	85,20%
(BigBR x 0,052) + (CGL x 0,115) + (MaxML x 0,114) -25,567					
Εξίσωση 3^β	0,381	<0,001	♀ -1,411 \ [-0,141] \ ♂ 1,129	89,90%	84,80%
(BicBR x -0,037) + (BigBR x 0,103) + (BicBR x 0,007) + (SyH x 0,032) + (SyBR x -0,041) + (RH x -0,017) + (CGL x 0,151) + (MaxBL x -0,047) + (MaxML x 0,126) -25,838					
Εξίσωση 4^β	0,481	<0,001	♀ -1,131 \ [-0,0935] \ ♂ 0,944	85,70%	85,70%
(BigBR x 0,064) + (CGL x 0,119) + (MaxBL x -0,063) + (MaxML x 0,139) -24,786					
Εξίσωση 5^γ	0,532	<0,001	♀ -1,015 \ [-0,078] \ ♂ 0,859	82,30%	81,80%
(MRH x 0,214) + (PRH x -0,064) + (CH x 0,09) + (MaxRBR x 0,100) + (MinRBR x -0,124) -16,856					
Εξίσωση 6^γ	0,633	<0,001	♀ -0,824 \ [-0,0635] \ ♂ 0,697	80,20%	79,70%
(CH x 0,180) + (MaxRBR x 0,137) + (MinRBR x -0,133) -12,807					
Εξίσωση 7^δ	0,594	<0,001	♀ -0,894 \ [-0,0685] \ ♂ 0,757	77,10%	77,10%
(MRH x 0,220) + (MaxRBR x 0,029) -16,028					
Εξίσωση 8^δ	0,585	<0,001	♀ -0,91 \ [-0,069] \ ♂ 0,772	80,40%	80,40%
(MRH x 0,247) + (MinRBR x -0,072) -14,495					
Εξίσωση 9^ε	0,519	<0,001	♀ -1,05 \ [-0,0885] \ ♂ 0,873	85,90%	84,90%
(BicBR x 0,008) + (BigBR x 0,049) + (MRH x 0,127) + (MaxML x 0,087) -24,711					
Εξίσωση 10^{στ}	0,498	<0,001	♀ -1,105 \ [-0,101] \ ♂ 0,903	84,10%	84,10%
(RH x 0,097) + (CH x 0,072) + (MaxML x 0,102) -22,935					
Εξίσωση 11^{στ}	0,515	<0,001	♀ -1,068 \ [-0,0975] \ ♂ 0,873	84,10%	83,60%
(MRH x 0,063) + (CH x 0,079) + (MaxML x 0,118) -23,516					
Εξίσωση 12^{στ}	0,494	<0,001	♀ -1,102 \ [-0,091] \ ♂ 0,92	84,10%	84,10%
(RH x 0,081) + (CGL x 0,086) + (MaxML x 0,108) -23,085					
Εξίσωση 13^{στ}	0,498	<0,001	♀ -1,093 \ [-0,0905] \ ♂ 0,912	84,10%	83,10%
(MRH x 0,064) + (CGL x 0,096) + (MaxML x 0,111) -23,417					

^α Εξίσωση βασισμένη στην μελέτη των Steyen και Iscan (1998).

^β Εξίσωση βασισμένη στην μελέτη του Franklin και των συνεργατών του (2008).

^γ Εξίσωση βασισμένη στην μελέτη των Saini και συνεργατών (2011).

^δ Εξίσωση βασισμένη στην μελέτη των Lin και συνεργατών (2014).

^ε Εξίσωση βασισμένη στην μελέτη των Dong και συνεργατών (2015).

^{στ} Συνδυασμός από τις πέντε μεταβλητές με την μεγαλύτερη μονομεταβλητή διακριτική ικανότητα. Η πολυπαραγοντική ανάλυση (DFA) τους με την χρήση της μεθόδου stepwise έδωσε την ίδια εξίσωση και ακρίβεια.

Πίνακας 8.3: Διακριτικές εξισώσεις βασισμένες στις μετρήσεις της αριστερής πλευράς της κάτω γνάθου.

Εξίσωση	Wilks' Lambda	Sig.	Μέσες τιμές και σημεία τομής	Ορθώς καταναμεημένα	
				Σύνολο δείγματος	Διασταυρωμένη επικύρωση
Εξίσωση 1^α	0,587	<0,001	$\varphi -0,915 \setminus [-0,077] \setminus \sigma 0,761$	83,20%	81,60%
(BicBR x 0,028) + (BigBR x 0,065) + (MinRBR x -0,014) + (BL x -0,042) + (MaxML x 0,174) -26,748					
Εξίσωση 2^β	0,506	<0,001	$\varphi -1,074 \setminus [-0,088] \setminus \sigma 0,898$	84,30%	84,30%
(BigBR x 0,056) + (CGL x 0,117) + (MaxML x 0,109) -25,455					
Εξίσωση 3^β	0,399	<0,001	$\varphi -1,36 \setminus [-0,136] \setminus \sigma 1,088$	89,90%	84,80%
(BicBR x -0,017) + (BigBR x 0,101) + (BicrBR x 0,002) + (SyH x 0,024) + (SyBR x -0,020) + (RH x -0,010) + (CGL x 0,154) + (MaxBL x -0,034) + (MaxML x 0,099) -26,007					
Εξίσωση 4^β	0,497	<0,001	$\varphi -1,095 \setminus [-0,0895] \setminus \sigma 0,916$	85,30%	84,80%
(BigBR x 0,069) + (CGL x 0,122) + (MaxBL x -0,056) + (MaxML x 0,128) -24,741					
Εξίσωση 5^γ	0,546	<0,001	$\varphi -0,981 \setminus [-0,0715] \setminus \sigma 0,838$	83,80%	82,20%
(MRH x 0,179) + (PRH x -0,047) + (CH x 0,098) + (MaxRBR x 0,123) + (MinRBR x -0,166) -15,675					
Εξίσωση 6^γ	0,619	<0,001	$\varphi -0,848 \setminus [-0,065] \setminus \sigma 0,718$	79,20%	78,60%
(CH x 0,182) + (MaxRBR x 0,169) + (MinRBR x -0,159) -13,474					
Εξίσωση 7^δ	0,626	<0,001	$\varphi -0,832 \setminus [-0,06] \setminus \sigma 0,712$	77,70%	77,20%
(MRH x 0,211) + (MaxRBR x 0,027) -15,299					
Εξίσωση 8^δ	0,61	<0,001	$\varphi -0,86 \setminus [-0,062] \setminus \sigma 0,736$	80,30%	80,30%
(MRH x 0,245) + (MinRBR x -0,107) -13,339					
Εξίσωση 9^ε	0,545	<0,001	$\varphi -0,997 \setminus [-0,084] \setminus \sigma 0,829$	84,30%	82,70%
(BicBR x 0,014) + (BigBR x 0,055) + (MRH x 0,116) + (MaxML x 0,082) -24,524					
Εξίσωση 10^{στ}	0,525	<0,001	$\varphi -1,029 \setminus [-0,0795] \setminus \sigma 0,87$	83,20%	82,10%
(RH x 0,086) + (CH x 0,079) + (MaxML x 0,100) -22,525					
Εξίσωση 11^ζ	0,538	<0,001	$\varphi -1,003 \setminus [-0,078] \setminus \sigma 0,847$	83,70%	83,20%
(MRH x 0,045) + (CH x 0,088) + (MaxML x 0,118) -22,900					
Εξίσωση 12^η	0,542	<0,001	$\varphi -0,995 \setminus [-0,0775] \setminus \sigma 0,84$	84,20%	83,20%
(CH x 0,110) + (MaxML x 0,138) -23,604					
Εξίσωση 13^{στ}	0,518	<0,001	$\varphi -1,05 \setminus [-0,086] \setminus \sigma 0,878$	83,20%	82,70%
(RH x 0,066) + (CGL x 0,095) + (MaxML x 0,109) -22,949					
Εξίσωση 14^ζ	0,523	<0,001	$\varphi -1,039 \setminus [-0,085] \setminus \sigma 0,869$	84,30%	83,80%
(MRH x 0,041) + (CGL x 0,104) + (MaxML x 0,118) -23,137					
Εξίσωση 15^θ	0,527	<0,001	$\varphi -1,031 \setminus [-0,0845] \setminus \sigma 0,862$	84,80%	84,30%
(CGL x 0,123) + (MaxML x 0,137) -23,774					

^α Εξίσωση βασισμένη στην μελέτη των Steyen και Iscan (1998).

^β Εξίσωση βασισμένη στην μελέτη του Franklin και των συνεργατών του (2008).

^γ Εξίσωση βασισμένη στην μελέτη των Saini και συνεργατών (2011).

^δ Εξίσωση βασισμένη στην μελέτη των Lin και συνεργατών (2014).

^ε Εξίσωση βασισμένη στην μελέτη των Dong και συνεργατών (2015).

^{στ} Συνδυασμός από τις πέντε μεταβλητές με την μεγαλύτερη μονομεταβλητή διακριτική ικανότητα. Η πολυπαραγοντική ανάλυση (DFA) τους με την χρήση της μεθόδου stepwise έδωσε την ίδια εξίσωση και ακρίβεια.

^ζ Συνδυασμός από τις πέντε μεταβλητές με την υψηλότερη μονομεταβλητή διακριτική ικανότητα.

^η Το αποτέλεσμα της stepwise DFA ανάλυσης χρησιμοποιώντας τις μεταβλητές της εξίσωσης 11.

^θ Το αποτέλεσμα της stepwise DFA ανάλυσης χρησιμοποιώντας τις μεταβλητές της εξίσωσης 14.

Όπως διακρίνεται στους πίνακες, η διασταυρωμένη επικύρωση των διακριτικών εξισώσεων κυμαίνεται από 77,1% μέχρι 85,7% με την εξίσωση 7 (η οποία βασίζεται στις μεταβλητές *MRH* και *MinRBR*) να παρουσιάζει την χαμηλότερη ακρίβεια τόσο για τις μεταβλητές της δεξιάς όσο και της αριστερής πλευράς. Αντιθέτως, η διαχωριστική εξίσωση 4 (η οποία βασίζεται στις μεταβλητές *BigBR*, *CGL*, *MaxBL* και *MaxML*) παρουσίασε την υψηλότερη διακριτική ικανότητα σωστής ταξινόμησης. Όσον αφορά στις διακριτικές εξισώσεις που δημιουργήθηκαν από το συνδυασμό των πιο φυλετικά διμορφικών μεταβλητών από το δείγμα του ελληνικού πληθυσμού το ποσοστό επιτυχούς ταξινόμησης έπειτα από επικυρωμένη διασταύρωση ήταν μεγαλύτερο από 82%. Οι εξισώσεις διαχωρισμού που δημιουργήθηκαν με βάση τις μεταβλητές από τη δεξιά πλευρά εμφάνισαν ελαφρώς καλύτερα αποτελέσματα. Εξάιρεση αποτέλεσε η εξίσωση 15, η οποία παρουσίασε το υψηλότερο ποσοστό διασταυρωμένης επικύρωσης (84,3%).

Στον Πίνακα 8.4 συνοψίζονται τα αποτελέσματα της εφαρμογής των διακριτικών εξισώσεων της βιβλιογραφίας στο δείγμα του ελληνικού πληθυσμού. Όπως ήταν αναμενόμενο, τα ποσοστά επιτυχούς ταξινόμησης ήταν χαμηλότερα σε σχέση με τις αντίστοιχες εξισώσεις που προέρχονται από τον ελληνικό πληθυσμό. Πιο συγκεκριμένα, η ακρίβεια της εφαρμογής των εξισώσεων στο ελληνικό δείγμα κυμαίνεται από 63,78% ως 83,84%, όταν χρησιμοποιούνται οι μετρήσεις της αριστερής πλευράς και από 68,11% μέχρι 83,84% όταν χρησιμοποιούνται οι μετρήσεις της δεξιάς πλευράς. Τα αποτελέσματα με τη μεγαλύτερη ομοιότητα μεταξύ του πληθυσμού αναφοράς και του ελληνικού πληθυσμού παρατηρούνται με την εφαρμογή της εξίσωσης 3, η οποία περιλαμβάνει 9 μεταβλητές και το ποσοστό επιτυχούς ταξινόμησης και για τις δύο πλευρές έφτασε το 83,84%.

Πίνακας 8.4: Η ακρίβεια πρόβλεψης φύλου των εξισώσεων της βιβλιογραφίας και τα αποτελέσματα κατόπιν εφαρμογής τους στον ελληνικό πληθυσμό.

Ημ/νια	Ερευνητές	Πληθυσμός	Εξίσωση	Ακρίβεια εξίσωσης	Ακρίβεια της εφαρμογής στο ελληνικό δείγμα
1998	Steyn & Iscan	Λευκοί Νοτιοαφρικανοί	Εξίσωση 1	81.5%	Αριστερή πλευρά : 63,78% Δεξιά πλευρά : 68,11%
2008	Franklin	Γηγενείς Νοτιοαφρικανοί	Εξίσωση 2	81.8%	Αριστερή πλευρά : 72,77% Δεξιά πλευρά : 74,60%
2008	Franklin.	Γηγενείς Νοτιοαφρικανοί	Εξίσωση 3	84,00%	Αριστερή πλευρά : 83,84% Δεξιά πλευρά : 83,84%
2008	Franklin	Γηγενείς Νοτιοαφρικανοί	Εξίσωση 4	82,70%	Αριστερή πλευρά : 78,53% Δεξιά πλευρά : 77,78%
2011	Saini	Ινδοί	Εξίσωση 5	80,20%	Αριστερή πλευρά : 65,45% Δεξιά πλευρά : 68,75%
2011	Saini	Ινδοί	Εξίσωση 6	80,20%	Αριστερή πλευρά : 75,00% Δεξιά πλευρά : 71,88%
2014	Lin	Κορεάτες	Εξίσωση 7	80,80%	Αριστερή πλευρά : 67,36% Δεξιά πλευρά : 68,75%
2014	Lin	Κορεάτες	Εξίσωση 8	80,40%	Αριστερή πλευρά : 68,84% Δεξιά πλευρά : 69,59%

8.3. Συμπεράσματα

Τα αποτελέσματα της παρούσας εργασίας σχετικά με τους πιο φυλετικά διμορφικούς μετρικούς χαρακτήρες της κάτω γνάθου, ήταν σε συμφωνία με προηγούμενες μελέτες του Franklin και των συναδέλφων του (Franklin et al., 2006, 2008) οι οποίοι διαπίστωσαν ότι η κατακόρυφη προβολή ανάμεσα στο κορόνιο και στη βάση του οστού, το ύψος του κλάδου και η απόσταση ανάμεσα στο υψηλότερο σημείο του κονδύλου και στο πογόνιο εμφάνισαν τον πιο έντονο φυλετικό διμορφισμό. Επιπλέον, η διακριτική εξίσωση που προέρχεται από τον ελληνικό πληθυσμό και περιλαμβάνει τις συγκεκριμένες μεταβλητές εμφάνισε διασταυρωμένη επικύρωση 84,1%. Παρόλο που το κρανίο και η λεκάνη παρέχουν πολύ πιο αξιόπιστες εκτιμήσεις φύλου (Bertsatos, Parageorgiou, et al., 2018; Οικοποπούλου et al., 2017), τα αποτελέσματά μας επιβεβαίωσαν ότι η κάτω γνάθος μπορεί να αξιοποιηθεί στην δικαστική ανθρωπολογία στις περιπτώσεις όπου το κρανίο και η λεκάνη δεν είναι διαθέσιμα.

Εκτός από την επαλήθευση του γεγονότος ότι τις περισσότερες φορές οι διακριτικές εξισώσεις που βασίζονται σε ανθρώπινα οστά εμφανίζουν πληθυσμιακή εξειδίκευση, εξ ου και η ανάγκη για μελέτες σε διαφορετικά δείγματα πληθυσμού, το συγκριτικό μέρος αυτής της εργασίας αποκάλυψε και ορισμένες άλλες ενδιαφέρουσες πτυχές. Η παρατήρηση ότι οι ίδιοι συνδυασμοί μεταβλητών αποδίδουν την ίδια ικανότητα διάκρισης φύλου μεταξύ διαφορετικών πληθυσμών, παρά το γεγονός ότι οι αντίστοιχες διακριτικές εξισώσεις τους εμφανίζουν πληθυσμιακή εξειδίκευση, υποδηλώνει ότι το μέγεθος της έκφρασης φυλετικού διμορφισμού ορισμένων μορφομετρικών χαρακτηριστικών της κάτω γνάθου είναι παρόμοιο ανεξαρτήτως πληθυσμού. Ωστόσο, η έκφρασή τους μπορεί να ακολουθεί διαφορετικά μοτίβα σε κάθε πληθυσμό. Τέλος, σχετικά με τη διακριτική εξίσωση 3 και πιο συγκεκριμένα το γεγονός ότι εμφανίζει τα ίδια ποσοστά επιτυχούς ταξινόμησης με αυτά που προκύπτουν από την εφαρμογή της αντίστοιχης εξίσωσης που προέρχεται από αυτόχθονες Νοτιοαφρικανούς στον ελληνικό πληθυσμό, δεν μπορούμε να συμπεράνουμε εάν αυτή η παρατήρηση οφείλεται απλά σε στατιστική τυχαιότητα λόγω δείγματος ή σε κοινούς περιβαλλοντικούς και αναπτυξιακούς παράγοντες μεταξύ των δύο πληθυσμών που διαμόρφωσαν κοινά πρότυπα έκφρασης του φυλετικού διμορφισμού.

9. Μια καινοτόμος μέθοδος για την γεωμετρική ανάλυση της εγκάρσιας τομής της διάφυσης των μακρών οστών

Οι γεωμετρικές παράμετροι της εγκάρσιας τομής της διάφυσης των μακρών οστών (CSG), όπως το εμβαδόν και οι δευτεροβάθμιες ροπές αδράνειας της επιφάνειας έχουν χρησιμοποιηθεί εκτενώς για τα μηχανικά φορτία, τα οποία εφαρμόζονται στις διαφύσεις των μακρών οστών (Cameron & Pfeiffer, 2014; Maggiano et al., 2008; O'Neill & Ruff, 2004; Shackelford et al., 2013; Stock & Pfeiffer, 2001; Zaki et al., 2015). Οι CSG παράμετροι παρέχουν σχετικές με το σχήμα ποσοτικές μετρήσεις της μηχανικής απόδοσης του οστού, και οι οποίες σχετίζονται με την αντοχή του σε δυνάμεις κάμψης (Ruff et al., 1993). Ο σωστός υπολογισμός των CSG παραμέτρων, ο οποίος είναι καθοριστικής σημασίας κατά τη σύγκριση ατόμων (Sparacello & Pearson, 2010) ή τη μελέτη παραγόντων όπως η ανάπτυξη και η γήρανση (Ruff et al., 2013), απαιτεί τον ποσοτικό προσδιορισμό του περιγράμματος τόσο του ενδόστεου όσο και του περιόστεου (O'Neill & Ruff, 2004). Ωστόσο, οι CSG παράμετροι μπορούν να υπολογιστούν κατά προσέγγιση προσδιορίζοντας ποσοτικά μόνο το περίγραμμα της εγκάρσιας τομής του περιόστεου (Macintosh et al., 2013; Sparacello & Pearson, 2010). Επιπλέον, οι CSG παράμετροι που υπολογίζονται βάσει του περιγράμματος της εγκάρσιας τομής του περιόστεου μπορούν να παρέχουν σημαντικές πληροφορίες στη μελέτη παρελθοντικών δραστηριοτήτων (Sparacello & Pearson, 2010), καθώς το περιόστεο επηρεάζεται περισσότερο από μηχανικά ερεθίσματα, ενώ το ενδόστεο από τη διατροφή (Mays et al., 2009; Ruff et al., 2013). Τέλος, οι CSG παράμετροι που βασίζονται στο περίγραμμα της εγκάρσιας τομής του περιόστεου δεν απαιτούν την χρήση αξονικού τομογράφου και επιτρέπουν εναλλακτικές μεθόδους για τον υπολογισμό τους, όπως η λήψη δακτυλίων με οδοντιατρική σιλικόνη (LCM). Παρά τη δημοτικότητα της, η LCM μέθοδος απαιτεί κόπο όσον αφορά την προετοιμασία των δακτυλίων και την προ-επεξεργασία της εικόνας πριν από την ανάλυση και είναι επιρρεπής σε μετρητικά σφάλματα σχετιζόμενα με το χρήστη τόσο κατά την ψηφιοποίηση των δακτυλίων όσο και τον προσανατολισμό της εικόνας.

Η παρούσα εργασία εισήγαγε μια νέα μέθοδο για την γεωμετρική ανάλυση της εγκάρσιας τομής της διάφυσης του βραχιονίου, του μηριαίου και της κνήμης. Αν και η ιδέα της ανάλυσης δεδομένων από σαρωτή λέιζερ για τον υπολογισμό των CSG παραμέτρων της διάφυσης των μακρών οστών δεν είναι καινούργια (Davies et al., 2012), το λογισμικό ανοικτού κώδικα που παρουσιάζεται στην παρούσα εργασία περιλαμβάνει αλγόριθμους στη γλώσσα προγραμματισμού GNU Octave που υπολογίζουν αυτόματα τις CSG παραμέτρους για οποιοδήποτε από τα προαναφερθέντα μακρά οστά. Η πραγματική ωστόσο καινοτομία του λογισμικού που παρουσιάζεται, είναι η πλήρης εξάλειψη μετρητικού σφάλματος σχετιζόμενο με το χρήστη κατά την επεξεργασία των ψηφιακών

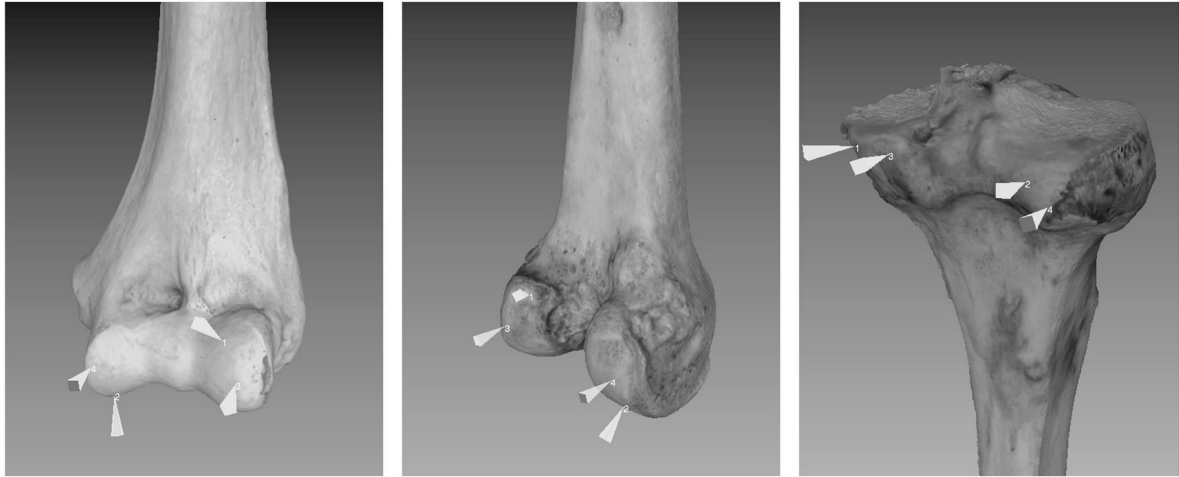
μοντέλων των οστών και επιτρέπει στον χρήστη να ανακτήσει εύκολα τις CSG παραμέτρους του διαθέσιμου σκελετικού δείγματος (Bertsatos & Chovaloroulou, 2018b).

9.1. Προδιαγραφές λογισμικού και περιγραφή αλγορίθμου

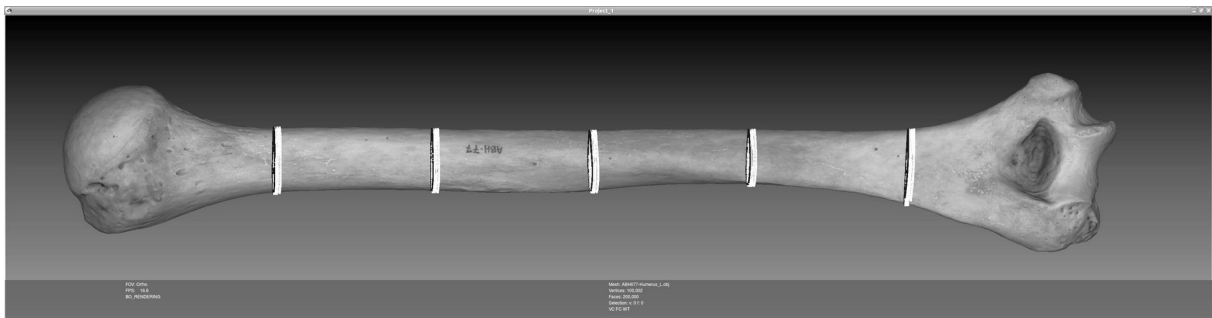
Το λογισμικό ανοικτού κώδικα με την ονομασία long-bone-diaphyseal-CSG-Toolkit (Bertsatos, 2018, 2020) απαιτεί ελάχιστη παρέμβαση από τον χρήστη, ο οποίος χρειάζεται να ορίσει μόνο δύο σημεία στο ψηφιακό μοντέλο του οστού για την αρχική ευθυγράμμιση του σε ανατομική θέση. Αυτό μπορεί να επιτευχθεί εύκολα χρησιμοποιώντας το λογισμικό Meshlab και αποθηκεύοντας τις συντεταγμένες των σημείων στην κατάλληλη μορφή αρχείου (.pp).

Τα σημεία που ορίζονται από τον χρήστη, αν και είναι συγκεκριμένα για κάθε τύπο οστού, δηλαδή για το βραχιόνιο, το μηριαίο και την κνήμη, δεν χρειάζεται να καθοριστούν με ακρίβεια, καθώς το λογισμικό (CSG Toolkit) πραγματοποιεί αυτόματη βελτιστοποίηση της θέσης τους σύμφωνα με τη γεωμετρία κάθε οστού. Παρ' όλα αυτά, για το βραχιόνιο τα δύο σημεία πρέπει να βρίσκονται στην πρόσθια πλευρά στην τροχιλία και στον κόνδυλο αντίστοιχα. Όσον αφορά το μηριαίο και την κνήμη, τα σημεία πρέπει να οριστούν στην οπίσθια πλευρά στον έσω και έξω κόνδυλο. Η σειρά με την οποία ο χρήστης ορίζει τα δύο σημεία δεν επηρεάζει τη λειτουργία του λογισμικού είτε πρόκειται για αριστερά είτε για δεξιά οστά. Το CSG Toolkit βελτιστοποιεί τις θέσεις των δύο σημείων και κατά συνέπεια τον προσανατολισμό των οστών όσον αφορά την ανατομική τους θέση, με την προϋπόθεση ότι τα σημεία που ορίζονται από τον χρήστη βρίσκονται κοντά στις προαναφερθείσες περιοχές. Τα σημεία που προκύπτουν έπειτα από την εφαρμογή του CSG Toolkit ορίζονται ως εξής. Για το βραχιόνιο, τα δύο πιο πρόσθια σημεία της τροχιλίας και του κονδύλου αντίστοιχα. Για το μηριαίο, τα δύο πιο οπίσθια σημεία των έσω και έξω κονδύλων αντίστοιχα. Για την κνήμη, τα δύο πιο οπίσθια σημεία της επιφάνειας του οστού που βρίσκονται στο πίσω και κάτω μέρος από τους έσω και έξω κονδύλους αντίστοιχα (Εικόνα 9.1).

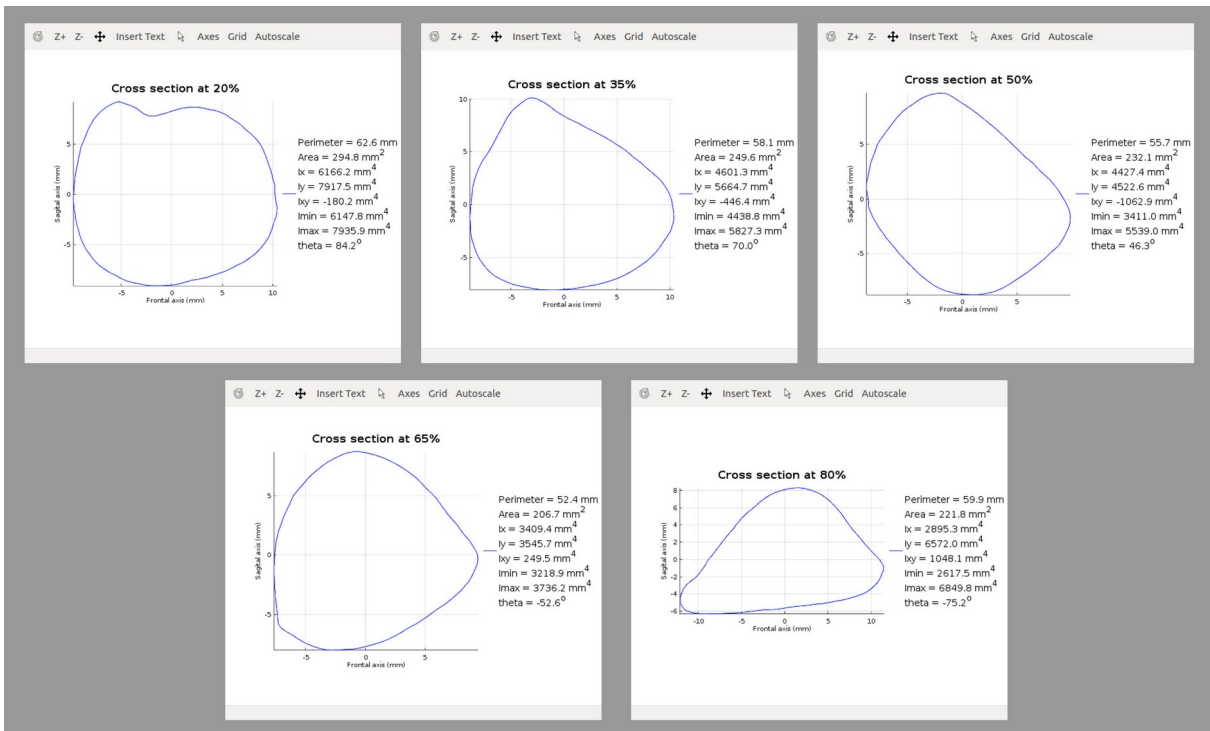
Στην συνέχεια και αφού το CSG Toolkit έχει προσανατολίσει τα οστά στην ανατομική τους θέση με το βέλτιστο δυνατό τρόπο, εκτιμά τις εγκάρσιες τομές της διάφυσης των οστών στο 20%, 35%, 50%, 65% και 80% του μεγίστου μήκους τους (Εικόνα 9.2). Τέλος, το CSG Toolkit υπολογίζει την περίμετρο, το εμβαδόν, τις δευτεροβάθμιες ροπές αδράνειας της επιφάνειας ως προς τους κεντροβαρικούς άξονες I_x και I_y , τις δευτεροβάθμιες ροπές αδράνειας της επιφάνειας ως προς τους κύριους άξονες I_{MAX} και I_{MIN} καθώς και την γωνία θ_m για κάθε εγκάρσια τομή της διάφυσης (Εικόνα 9.3). Επιπλέον δίνονται τα διδιάστατα και τριδιάστατα σημεία, τα οποία περιγράφουν το περίγραμμα κάθε εγκάρσια τομή της διάφυσης. Όλες οι προαναφερθείσες τιμές παρέχονται σε CSV αρχεία.



Εικόνα 9.1: Σημεία που ορίζονται από τον χρήστη (1, 2) και σημεία που προκύπτουν έπειτα από την εφαρμογή του CSG Toolkit στο βραχιόνιο (αριστερά), το μηριαίο (μέση) και την κνήμη (δεξιά).



Εικόνα 9.2: Σημεία περιγράμματος εγκάρσιων διατομών του βραχιονίου.



Εικόνα 9.3: Διαγράμματα περιγράμματος εγκάρσιων διατομών του βραχιονίου και οι γεωμετρικές παράμετροι αυτών.

9.2. Έλεγχος ακρίβειας και αξιοπιστίας του CSG Toolkit

Για τον συγκεκριμένο σκοπό χρησιμοποιήθηκαν τρία τριδιάστατα μοντέλα οστών (ένα βραχιόνιο, ένα μηριαίο και μια κνήμη, τα οποία αποτελούν μέρος της σύγχρονης σκελετικής συλλογής αναφοράς της Αθήνας) και ένα τριδιάστατο μοντέλο με γνωστές γεωμετρικές ιδιότητες, το οποίο δημιουργήθηκε με το λογισμικό CAD.

Ο πρώτος έλεγχος αφορούσε τον αλγόριθμο, ο οποίος σχετίζεται με τη βελτιστοποίηση της θέσης των σημείων που ορίζονται από τον χρήστη. Προκειμένου να πραγματοποιηθεί έλεγχος αξιοπιστίας του αλγόριθμου, το ίδιο τριδιάστατο μοντέλο βραχιονίου, όπως και μηριαίου και κνήμης, αναλύθηκαν επανειλημμένα με το CSG Toolkit ορίζοντας ο χρήστης κάθε φορά διαφορετικό ζεύγος σημείων. Τα σημεία που προέκυψαν έπειτα από την εφαρμογή του CSG Toolkit σε όλες τις δοκιμές ήταν ακριβώς τα ίδια. Η ακρίβεια του αλγόριθμου ελέγχθηκε με το τριδιάστατο μοντέλο του οποίου οι γεωμετρικές ιδιότητες ήταν γνωστές, και τα αποτελέσματα εμφάνισαν απόλυτη ταύτιση.

Ο δεύτερος έλεγχος αφορούσε τον υπολογισμό των CSG παραμέτρων και βασίστηκε στο τριδιάστατο μοντέλο του οποίου οι γεωμετρικές ιδιότητες ήταν γνωστές. Το μοντέλο με γνωστές γεωμετρικές ιδιότητες περιελάμβανε 5 βασικά σχήματα (τετράγωνο, έλλειψη, εξάγωνο, ορθογώνιο και κύκλο). Όσον αφορά τα κλασικά πολύγωνα, όλες οι CSG παράμετροι που υπολογίστηκαν από τον αλγόριθμο ήταν ταυτόσημες με τις αντίστοιχες θεωρητικές τιμές βάση των γνωστών διαστάσεων κάθε πολυγώνου. Μικρές αποκλίσεις παρατηρήθηκαν στον κύκλο και την έλλειψη δεδομένου ότι τα συγκεκριμένα σχήματα στα τριδιάστατα μοντέλα αποτελούν πολύγωνα με πάρα πολύ μεγάλο αριθμό κορυφών. Παρόλα αυτά, το αποτέλεσμα απέδειξαν ότι η εφαρμογή του CSG Toolkit υπολογίζει με ακρίβεια όλες CSG παραμέτρους. Για περαιτέρω έλεγχο, οι εγκάρσιες τομές που προέκυψαν από τα 5 βασικά σχήματα ύστερα από την εφαρμογή του CSG Toolkit αναλύθηκαν και με το λογισμικό ImageJ. Στόχος ήταν να φανεί η συνοχή με μια εναλλακτική μέθοδο υπολογισμού των CSG παραμέτρων, αλλά και να εξεταστεί το σφάλμα που προκύπτει στους υπολογισμούς από την ανάλυση εικόνων. Ο Πίνακας 9.1 συνοψίζει τα σχετικά σφάλματα (ποσοστό απόκλισης από τις θεωρητικές τιμές κάθε σχήματος) των CSG Toolkit και ImageJ. Οι τιμές I_x , I_y και I_{xy} δεν δίνονται στο λογισμικό ImageJ και κατά συνέπεια δεν συμπεριλήφθηκαν στη σύγκριση.

Πίνακας 9.1: Σχετικά σφάλματα στον υπολογισμό των CSG παραμέτρων του τριδιάστατου μοντέλου με γνωστές γεωμετρικές ιδιότητες κατόπιν εφαρμογής του CSG Toolkit και του ImageJ.

CSG παράμετροι	CSG-Toolkit					ImageJ				
	Περίμετρος	Εμβαδόν	I_{MIN}	I_{MAX}	θ^{α}	Περίμετρος	Εμβαδόν	I_{MIN}	I_{MAX}	θ^{α}
Κύκλος	0,27%	0,80%	1,74%	1,42%	-	4,72%	1,75%	5,17%	4,87%	-
Ορθογώνιο	0	0	0	0	0	6,86%	0,44%	2,86%	2,68%	0
Εξάγωνο	0	0	0	0	0,4° ^β	5,33%	1,11%	0,06%	1,23%	1,4°
Έλλειψη	0,10%	0,66%	1,91%	0,72%	0	5,23%	0,58%	4,02%	2,95%	0,1°
Τετράγωνο	0	0	0	0	0	0,33%	1,02%	1,64%	2,35%	1,1°

^α Η γωνία θ περιγράφει το απόλυτο σφάλμα σε μοίρες.

^β Η πολύ μικρή απόκλιση που παρατηρείται στη γωνία θ του εξαγώνου είναι αποτέλεσμα σφάλματος στρογγυλοποίησης λόγω του τρόπου κατάτμησης του τριγωνικού πλέγματος του μοντέλου αναφοράς από το αρχικό τριδιάστατο στερεό στο AutoCAD.

Ο τελευταίος έλεγχος αφορούσε την άμεση σύγκριση των αποτελεσμάτων, όσον αφορά τον υπολογισμό των CSG παραμέτρων, μεταξύ της LCM μεθόδου από πραγματικά οστά μέσω της ανάλυσης των δακτυλίων με το λογισμικό ImageJ και την εφαρμογή του CSG Toolkit στα αντίστοιχα τριδιάστατα μοντέλα. Όπως και με τον προηγούμενο έλεγχο, οι εγκάρσιες τομές των τριδιάστατων μοντέλων αναλύθηκαν με το λογισμικό ImageJ προκειμένου να υπάρχει μέτρο σύγκρισης. Τα αποτελέσματα μεταξύ των δύο μεθόδων ήταν παρόμοια. Πιο συγκεκριμένα, τόσο οι σαρωμένες εικόνες από τους δακτυλίους της LCM μεθόδου όσο και οι εικόνες από τις εγκάρσιες τομές των τριδιάστατων μοντέλων εμφάνισαν παρόμοιο σχετικό σφάλμα σε σχέση με τις τιμές που υπολογίστηκαν με το CSG Toolkit. Επιπλέον, το σχετικό σφάλμα είναι συγκρίσιμο με εκείνο που παρατηρήθηκε στο τριδιάστατο μοντέλο με γνωστές γεωμετρικές ιδιότητες, γεγονός που υποδεικνύει ότι το μεγαλύτερο μέρος του σφάλματος δεν σχετίζεται με την λήψη δακτυλίων με οδοντιατρική σιλικόνη, αλλά με την προ-επεξεργασία της εικόνας (δηλ. κλίμακα εικόνας) και τον υπολογισμό των CSG παραμέτρων μέσω εικόνας. Φυσικά υπάρχει ένα ποσοστό σφάλματος σχετιζόμενο με το χρήστη, το οποίο προκύπτει από τον χειρωνακτικό υπολογισμό των αποστάσεων λήψης δακτυλίων καθώς και το χειρωνακτικό κόψιμο του με κατεύθυνση κάθετη προς τον άξονα της διάφυσης του οστού. Η μόνη εξαίρεση όσον αφορά τα σφάλματα και την ομοιότητά τους μεταξύ των δύο μεθόδων αφορά τη γωνία θ , η οποία εμφάνισε πολύ μεγαλύτερη διακύμανση τιμών μεταξύ των εγκάρσιων τομών του ίδιου οστού με την LCM μέθοδο. Οι τιμές του απόλυτου σφάλματος που σχετίζονται με την γωνία θ συνοψίζονται στον Πίνακα 9.2. Η απόκλιση που παρατηρούμε οφείλεται στην αδυναμία του παρατηρητή σωστής και σταθερής ευθυγράμμισης των δακτυλίων βάσει της ανατομικής θέσης του οστού, γεγονός που τονίζει τα οφέλη της αυτόματης ευθυγράμμισης που επιτυγχάνεται με το CSG Toolkit.

Πίνακας 9.2: Τιμές απόλυτου σφάλματος γωνίας θ μεταξύ των μεθόδων.

	Βραχιόνιο		Μηριαίο		Κνήμη	
	θ_{LCM}^1	θ_{PLOT}^2	θ_{LCM}^1	θ_{PLOT}^2	θ_{LCM}^1	θ_{PLOT}^2
Εγκάρσια τομή στο 20%	2,1	0,1	16,5	0,1	8,2	0,2
Εγκάρσια τομή στο 35%	3,7	0,1	11,2	0,1	2,3	0,1
Εγκάρσια τομή στο 50%	1,6	0	12,9	0,2	4,4	0,1
Εγκάρσια τομή στο 65%	4,1	0,6	7,4	0	3,3	0,1
Εγκάρσια τομή στο 80%	3,1	0,2	6,6	0,1	13,7	0,7

¹ Η γωνία θ_{LCM} περιγράφει το απόλυτο σφάλμα σε μοίρες μεταξύ των σαρωμένων εικόνων από τους δακτυλίους της LCM μεθόδου και τις τιμές που προέκυψαν από την εφαρμογή του CSG Toolkit στα τριδιάστατα μοντέλα.

² Η γωνία θ_{PLOT} περιγράφει το απόλυτο σφάλμα σε μοίρες μεταξύ των εικόνων από τις εγκάρσιες τομές των τριδιάστατων μοντέλων και τις τιμές που προέκυψαν από την εφαρμογή του CSG Toolkit στα τριδιάστατα μοντέλα.

10. Πλεονεκτήματα στον προσδιορισμό του φύλου αξιοποιώντας γεωμετρικές παραμέτρους της εγκάρσιας τομής της διάφυσης των κάτω και άνω άκρων.

Πέρα από τη συζήτηση “Κρανίο ή Μετακρανιακός σκελετός για τον προσδιορισμό του φύλου;” και τη σύγκριση των σκελετικών στοιχείων μεταξύ των ποσοστών ακρίβειας των αντίστοιχων μεθόδων, υπάρχει μια ακόμη σημαντική διάκριση. Πρόσφατες κρανιακές έρευνες έχουν συμβάλει στην πρόοδο της φυλετικής διακριτικής του ικανότητας λόγω των προηγμένων αναλυτικών μεθόδων που χρησιμοποιούν. Αυτές οι μέθοδοι κυμαίνονται από τη γεωμετρική μορφομετρία, την υπολογιστική γεωμετρία και την τεχνητή όραση για την ανάλυση τριδιάστατων μοντέλων έως τη χρήση βαθέων νευρωνικών δικτύων και μηχανών διανυσμάτων υποστήριξης (SVM) για την στατιστική ανάλυση (Bertsatos, Chovalopoulou, Brůžek, et al., 2020; Yang et al., 2019, 2020). Ταυτόχρονα, οι περισσότερες έρευνες που βασίζονται στον μετακρανιακό σκελετό για τον προσδιορισμό φύλου περιορίζονται στην χρήση παραδοσιακών μετρήσεων (Anzellini & Toyne, 2019; Curate et al., 2016; Frutos, 2005; Moore et al., 2016; Šlaus et al., 2013), και λίγες εξαιρέσεις έχουν επικεντρωθεί σε εναλλακτικές μεθόδους (Ammer et al., 2019; Kranioti et al., 2011; Tallman & Blanton, 2020). Αυτό αφορά σε μεγάλο βαθμό τα οστά του κάτω και άνω άκρου, των οποίων ο φυλετικός διμορφισμός αξιολογείται κατά κύριο λόγο βάση τυπικών οστεομετρικών μετρήσεων (Anzellini & Toyne, 2019; Duangto & Mahakkanukrauh, 2020; Frutos, 2005; İscan & Miller-Shaivitz, 1984; Kranioti & Michalodimitrakis, 2009; Kranioti & Tzanakis, 2015; Moore et al., 2016; Robinson & Bidmos, 2011; Ross & Manneschi, 2011; Šlaus et al., 2013; Spradley et al., 2008; Steyn & İscan, 1997). Παρά την ευκολία χρήσης τους, οι παραδοσιακές μετρήσεις δεν παρέχουν απαραίτητως τις πιο φυλετικά διμορφικές μεταβλητές (Bertsatos, Papageorgopoulou, et al., 2018) και επιπλέον θέτουν περαιτέρω ζητήματα σχετικά με το σφάλμα μέτρησης όταν πρόκειται για την εφαρμογή διακριτικών εξισώσεων. Ταυτόχρονα, η υψηλή συσχέτιση των γραμμικών μετρήσεων με το σωματικό μέγεθος και το ανάστημα εντείνει τις επιπτώσεις της πληθυσμιακής εξειδίκευσης, η οποία είναι εγγενές πρόβλημα στην ιατροδικαστική ανθρωπολογία και τη συνεχώς μεταβαλλόμενη δημογραφική διάρθρωση των σύγχρονων πληθυσμών (Spradley et al., 2008).

Η παρούσα εργασία επικεντρώνεται στα μακρά οστά των κάτω (μηριαίο, κνήμη) και άνω άκρων (βραχιόνιο) και εισάγει μια νέα μέθοδο προσδιορισμού του φύλου, η οποία προσπαθεί να ξεπεράσει τους προαναφερθέντες περιορισμούς των παραδοσιακών οστεομετρικών πρωτοκόλλων και των γραμμικών μετρήσεων. Ταυτόχρονα, η προτεινόμενη μέθοδος εξακολουθεί να είναι εύχρηστη για τον προσδιορισμό του φύλου ατόμων αγνώστων στοιχείων από τους δικαστικούς ανθρωπολόγους.

Η μέθοδος χρησιμοποιεί κατά κύριο λόγο τις γεωμετρικές παραμέτρους της εγκάρσιας τομής της διάφυσης των μακρών οστών (CSG), των οποίων τα πλεονεκτήματα όσον αφορά στην αντιστοίχιση αμφίπλευρων οστών από μαζικές ταφές έχουν ήδη αποδειχθεί (Bertsatos & Chovaloroulou, 2020). Αυτή η νέα μέθοδος προσδιορισμού του φύλου βασίζεται στο λογισμικό CSG-Toolkit (Bertsatos & Chovaloroulou, 2019), το οποίο υπολογίζει αυτόματα όλες τις απαραίτητες φυλετικά διμορφικές παραμέτρους από οποιαδήποτε διαθέσιμη τριδιάστατη ψηφιακή αναπαράσταση μιας κνήμης, ενός βραχιονίου ή ενός μηριαίου, και συνοδεύεται από μια εφαρμογή ανοιχτού κώδικα με γραφική διεπαφή χρήστη (GUI).

10.1. Υλικά και μέθοδοι

Η παρούσα εργασία χρησιμοποιεί 1205 οστά των κάτω και άνω άκρων 220 ενήλικων ατόμων γνωστού φύλου και ηλικίας θανάτου από τη σύγχρονη σκελετική συλλογή αναφοράς. Πιο συγκεκριμένα, χρησιμοποιήθηκαν 389 μηριαία (194 αριστερά και 195 δεξιά), τα οποία προέρχονται από 210 άτομα (116 άνδρες και 94 γυναίκες), 412 κνήμες (204 αριστερά και 208 δεξιά), οι οποίες προέρχονται από 214 άτομα (117 άνδρες και 97 γυναίκες) και 404 βραχιόνια (201 αριστερά και 203 δεξιά), τα οποία προέρχονται από 213 άτομα (118 άνδρες και 95 γυναίκες). Όλα τα οστά ψηφιοποιήθηκαν προηγουμένως σε τριδιάστατα μοντέλα πλέγματος με τη μέθοδο της φωτογραμμετρίας.

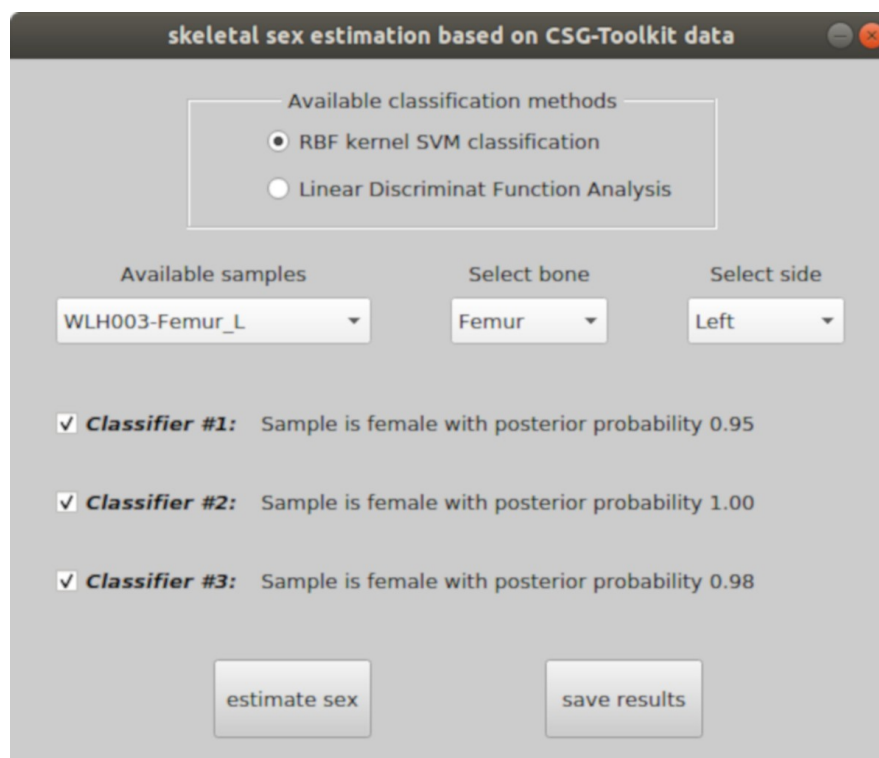
Κάθε οστό αναλύθηκε με την τελευταία έκδοση (v1.1.0) του long-bone-diaphyseal-CSG-Toolkit, το οποίο υπολογίζει αυτόματα τις γεωμετρικές παραμέτρους της εγκάρσιας τομής της διάφυσης των μακρών οστών απευθείας από τα τριδιάστατα μοντέλα. Μετά τον αρχικό έλεγχο για φυλετικό διμορφισμό των 46 μεταβλητών που υπολογίστηκαν από το CSG-Toolkit, αποκλείστηκαν οι παράμετροι I_{XY} και θ_m κάθε εγκάρσιας τομής καθώς και οι γωνίες της κάμψης της διάφυσης. Συμπληρωματικά προς τις υπόλοιπες μεταβλητές, υπολογίστηκαν άλλες τρεις παράμετροι για κάθε εγκάρσια τομή. Η πρώτη παράμετρος είναι το $ArPerIndex$, το οποίο υπολογίζεται βάσει του τύπου

$$\frac{Εμβαδόν \times 4\pi}{Περίμετρος^2}$$

και παρέχει ένα μέτρο σχετικά με το κυκλικό σχήμα και την ομαλότητα της εγκάρσιας τομής (Bertsatos, Chovaloroulou, et al., 2018). Οι άλλες δύο παράμετροι είναι οι λόγοι I_X / I_Y και I_{MAX} / I_{MIN} , οι οποίοι αντικατοπτρίζουν την επιμήκυνση της εγκάρσιας τομής κατά μήκος του μετωπιαίου (στεφανιαίου) άξονα και του κύριου άξονα της δευτεροβάθμιας ροπής επιφάνειας αντίστοιχα. Οι 46 αυτές μεταβλητές υποβλήθηκαν σε περαιτέρω στατιστική επεξεργασία και ανάπτυξη κατάλληλων φυλετικών χαρακτηριστικών για τα οστά του μηριαίου, της κνήμης και του βραχιονίου. Κάθε πλευρά αξιολογήθηκε ξεχωριστά για τη μεγιστοποίηση της αποδοτικότητας των φυλετικών χαρακτηριστικών.

Πριν από τις πολυπαραγοντικές αναλύσεις, σε όλες τις μεταβλητές πραγματοποιήθηκε κανονικοποίηση σε ολόκληρο το δείγμα (σύνολο αρσενικών και θηλυκών ατόμων) και στις επακόλουθες στατιστικές αναλύσεις χρησιμοποιήθηκαν τα αντίστοιχα z-scores. Στις μεταβλητές που εμφάνιζαν φυλετικό διμορφισμό πραγματοποιήθηκε διακριτική ανάλυση (DFA) εφαρμόζοντας την κλιμακωτή μέθοδο επιλογής ανεξάρτητων μεταβλητών για τον προσδιορισμό του υποσυνόλου αυτών που παρουσιάζουν τα βέλτιστα αποτελέσματα ταξινόμησης. Επιπλέον, για κάθε πλευρά και κάθε οστό, τα τρία βέλτιστα χαρακτηριστικά επιλέχθηκαν για περαιτέρω ανάλυση με εποπτευόμενους αλγόριθμους μηχανικής μάθησης και την ανάπτυξη μοντέλων με τη χρήση μηχανών διανυσματικής υποστήριξης (SVM) για τον προσδιορισμό του φύλου. Παρόλο που η μηχανική μάθηση δεν απαιτεί τις αυστηρές στατιστικές παραδοχές του DFA, χρησιμοποιήθηκαν τα ίδια σύνολα δεδομένων (δεν συμπεριλαμβάνονται οι ακραίες τιμές) για τη διευκόλυνση της σύγκρισης των αποτελεσμάτων ταξινόμησης.

Τόσο τα μοντέλα SVM όσο και τα μοντέλα διακριτικής ανάλυσης έχουν ενσωματωθεί σε μια εφαρμογή ανοιχτού κώδικα (GNU Octave), η ονομάζεται "estimate_sex" και επιτρέπει στο χρήστη να εφαρμόσει νέα σετ δεδομένων για τον προσδιορισμό του φύλου μέσω ενός γραφικού περιβάλλοντος διεπαφής με το χρήστη (Εικόνα 10.1). Η εφαρμογή παρέχεται ως λογισμικό ανοιχτού κώδικα (GPLv3) και διατίθεται ελεύθερα στη διεύθυνση <https://github.com/pr0m1th3as/skeletal-sex-estimation>.



Εικόνα 10.1: Το Γραφικό Περιβάλλον Χρήστη της εφαρμογής "estimate_sex".

10.2. Αποτελέσματα

Οι μόνες μεταβλητές, οι οποίες δεν εμφάνισαν σταθερά στατιστικά σημαντικό φυλετικό διμορφισμό είναι οι λόγοι I_X / I_Y και I_{MAX} / I_{MIN} . Αυτό είναι εμφανές σε όλα τα οστά τόσο στις εγκάρσιες τομές των οστών της δεξιάς πλευράς όσο και της αριστερής. Επιπλέον, δεν βρέθηκαν στατιστικά σημαντικά αποτελέσματα για τη μεταβλητή ArPerIndex στην εγκάρσια τομή στο 35% και 65% του δεξιού μηρού καθώς και στο 80% της δεξιάς κνήμης και του βραχιονίου. Όπως ήταν αναμενόμενο, το μέγιστο μήκος κάθε οστού ήταν σταθερά ανάμεσα στα πιο φυλετικά διμορφικά χαρακτηριστικά μαζί με τις μεταβλητές εμβαδόν και περίμετρος σε κάθε εγκάρσιας τομής (20%, 35%, 50%, 65%, 80%). Αξίζει να σημειωθεί πως η περίμετρος, κυρίως στις εγκάρσιες τομές κοντά στο μέσο της διάφυσης (35%, 50%, 65%), εμφάνισε πιο έντονο φυλετικό διμορφισμό σε σύγκριση με το μέγιστο μήκος και ιδιαίτερα στο βραχιόνιο.

Τα ποσοστά επικυρωμένης επιτυχούς ταξινόμησης της διακριτική ανάλυση κλιμακωτής μεθόδου επιλογής ανεξάρτητων μεταβλητών κυμαίνονται από 88,8% έως 97,3%. Όλα τα αποτελέσματα της DFA συνοψίζονται στον Πίνακα 10.1. Πιο συγκεκριμένα, το βραχιόνιο ήταν το περισσότερο φυλετικά διμορφικό μακρύ οστό με τα ποσοστά επικυρωμένης επιτυχούς ταξινόμησης να κυμαίνονται μεταξύ 95,2% και 97,3% και την αριστερή πλευρά να εμφανίζει ελαφρώς βελτιωμένες επιδόσεις. Το μηριαίο επέδειξε ποσοστά επικυρωμένης επιτυχούς ταξινόμησης μεταξύ 91,1% και 94,8%, όπου και σε αυτήν την περίπτωση η αριστερή πλευρά παρουσίασε ελαφρώς βελτιωμένες επιδόσεις. Παραδόξως, η κνήμη εμφάνισε την μεγαλύτερη διαφορά όσον αφορά στην ακρίβεια επιτυχούς ταξινόμησης μεταξύ δεξιάς και αριστερής πλευράς. Τα ποσοστά επικυρωμένης επιτυχούς ταξινόμησης της δεξιάς κνήμης κυμαίνονταν από 94% έως 94,7%, ποσοστά υψηλότερα από αυτά των περισσότερων διακριτικών εξισώσεων των μηριαίων, ενώ της αριστερής κνήμης μεταξύ 88,8% και 90,4% .

Πίνακας 10.1: Βέλτιστες διακριτικές εξισώσεις για το μηριαίο, την κνήμη και το βραχιόνιο.

Μοντέλο	Μεταβλητές	Ορθώς καταναμημένα		
		Σύνολο δείγματος	Διασταυρωμένη επικύρωση	
Αριστερό Μηριαίο #1	1. Μέγιστο Μήκος 3. Εμβαδόν 50% 5. ArPerIndex 80%	2. ArPerIndex 20% 4. Περίμετρος 50% 6. I_X / I_Y 80%	95,4%	94,8%
Αριστερό Μηριαίο #2	1. Μέγιστο Μήκος 3. Εμβαδόν 50% 5. I_X / I_Y 80%	2. ArPerIndex 20% 4. ArPerIndex 80%	93,2%	92,1%
Αριστερό Μηριαίο #3	1. Μέγιστο Μήκος 3. I_{MAX} 50%	2. Εμβαδόν 50% 4. ArPerIndex 80%	92,2%	91,1%
Δεξί Μηριαίο #1	1. Μέγιστο Μήκος 3. Εμβαδόν 50% 5. ArPerIndex 80%	2. ArPerIndex 20% 4. I_{MIN} 65% 6. I_{MIN} 80%	94,3%	92,6%
Δεξί Μηριαίο #2	1. Μέγιστο Μήκος 3. Εμβαδόν 50% 5. I_{MIN} 80%	2. ArPerIndex 20% 4. ArPerIndex 80%	93,8%	92,1%
Δεξί Μηριαίο #3	1. Μέγιστο Μήκος 3. Εμβαδόν 35%	2. ArPerIndex 20% 4. ArPerIndex 80%	91,8%	91,8%
Αριστερή Κνήμη #1	1. Μέγιστο Μήκος 3. ArPerIndex 50% 5. Περίμετρος 65%	2. I_{MAX} 35% 4. I_Y 50%	90,4%	90,4%
Αριστερή Κνήμη #2	1. Μέγιστο Μήκος 3. ArPerIndex 50% 5. Εμβαδόν 65%	2. I_{MAX} 35% 4. I_Y 50%	89,9%	89,9%
Αριστερή Κνήμη #3	1. Μέγιστο Μήκος 3. ArPerIndex 50% 5. Περίμετρος 65%	2. I_Y 35% 4. I_{MAX} 50%	89,9%	88,8%
Δεξιά Κνήμη #1	1. Μέγιστο Μήκος 3. Περίμετρος 35%	2. ArPerIndex 20% 4. I_Y 35%	94,7%	94,7%
Δεξιά Κνήμη #2	1. Μέγιστο Μήκος 3. Περίμετρος 35% 5. I_{MAX} 50%	2. ArPerIndex 20% 4. I_Y 35%	94,0%	94,0%
Δεξιά Κνήμη #3	1. Μέγιστο Μήκος 3. Περίμετρος 35%	2. ArPerIndex 20% 4. I_X 50%	94,0%	93,4%
Αριστερό Βραχιόνιο #1	1. Μέγιστο Μήκος 3. I_X / I_Y 35% 5. Εμβαδόν 65%	2. I_X / I_Y 20% 4. I_X 50% 6. I_{MIN} 80%	97,9%	97,3%
Αριστερό Βραχιόνιο #2	1. Μέγιστο Μήκος 3. I_X / I_Y 35% 5. I_{MAX} 65%	2. I_X / I_Y 20% 4. Εμβαδόν 65%	97,4%	96,8%
Αριστερό Βραχιόνιο #3	1. Μέγιστο Μήκος 3. I_X / I_Y 35% 5. I_{MIN} 80%	2. I_X / I_Y 20% 4. Εμβαδόν 65%	96,8%	96,8%
Δεξί Βραχιόνιο #1	1. Μέγιστο Μήκος 3. Εμβαδόν 65% 5. ArPerIndex 65%	2. I_X / I_Y 35% 4. Περίμετρος 65% 6. Περίμετρος 80%	96,2%	96,2%
Δεξί Βραχιόνιο #2	1. Μέγιστο Μήκος 3. Περίμετρος 65% 5. Περίμετρος 80%	2. I_X 35% 4. ArPerIndex 65%	96,2%	96,2%
Δεξί Βραχιόνιο #3	1. Μέγιστο Μήκος 3. I_X / I_Y 35%	2. I_X / I_Y 20% 4. Εμβαδόν 65%	95,2%	95,2%

Τα αποτελέσματα των SVM μοντέλων, τα οποία βασίστηκαν στα τρία βέλτιστα χαρακτηριστικά κάθε πλευράς και κάθε οστού συνοψίζονται στον Πίνακα 10.2. Τα αποτελέσματα της επικυρωμένης επιτυχούς ταξινόμησης (CV) ακολουθούν το ίδιο μοτίβο με τα αποτελέσματα αρχικής ομάδας (OR), ωστόσο είναι ελαφρώς χαμηλότερα. Επιπλέον, και στην περίπτωση της ανάλυσης με εποπτευόμενους αλγόριθμους μηχανικής μάθησης, το δεξί βραχιόνιο εμφάνισε τα υψηλότερα επίπεδα φυλετικού διμορφισμού, ενώ η αριστερή κνήμη παρουσίασε τη χαμηλότερη επίδοση όσον αφορά στο διαχωρισμό του φύλου.

Πίνακας 10.2: Αποτελέσματα SVM μοντέλων.

Μοντέλο	OR	CV	CI
Left Femur #1	0,931	0,908	0,071
Left Femur #2	0,932	0,911	0,063
Left Femur #3	0,911	0,917	0,071
Right Femur #1	0,949	0,925	0,066
Right Femur #2	0,949	0,926	0,065
Right Femur #3	0,924	0,921	0,070
Left Tibia #1	0,915	0,902	0,070
Left Tibia #2	0,915	0,903	0,072
Left Tibia #3	0,899	0,871	0,078
Right Tibia #1	0,947	0,935	0,074
Right Tibia #2	0,946	0,937	0,055
Right Tibia #3	0,945	0,913	0,089
Left Humerus #1	0,984	0,968	0,037
Left Humerus #2	0,974	0,963	0,045
Left Humerus #3	0,979	0,966	0,048
Right Humerus #1	0,962	0,960	0,050
Right Humerus #2	0,962	0,949	0,057
Right Humerus #3	0,962	0,940	0,069

OR: Σύνολο δείγματος; CV: Διασταυρωμένη επικύρωση; CI: 95% Διάστημα εμπιστοσύνης του CV

10.3. Συμπεράσματα

Η νέα μέθοδος προσδιορισμού φύλου εξαλείφει πλήρως τα σφάλματα που οφείλονται στον παρατηρητή, ενώ διευκολύνει τον ακριβή προσδιορισμό του φύλου από τριδιάστατα ψηφιακά μοντέλα οστών. Τα αποτελέσματά μας έδειξαν ότι ο συνδυασμός παραδοσιακών οστεομετρικών μετρήσεων, όπως το μέγιστο μήκος, με ειδικές μεταβλητές βασισμένες στις γεωμετρικές παραμέτρους της εγκάρσιας τομής της διάφυσης των μακρών οστών, όπως το ArPerIndex, βελτιώνει περαιτέρω τις επιδόσεις ταξινόμησης σε σύγκριση με τις υπάρχουσες μεθόδους. Τα ποσοστά επικυρωμένης επιτυχούς ταξινόμησης στο ελληνικό δείγμα άγγιξαν το 94,8% για το μηριαίο οστό, το 94,7% για την κνήμη και το 97,3% για το βραχιόνιο.

Τέλος, η προτεινόμενη μέθοδος συνοδεύεται από μια εφαρμογή ανοιχτού κώδικα με γραφική διεπαφή χρήστη του περιβάλλοντος GNU Octave, η οποία επιτρέπει στον χρήστη να επιλέξει οποιαδήποτε από τους διαθέσιμους συνδυασμούς μεταβλητών που παρουσιάζονται σε αυτήν τη μελέτη προκειμένου να προσδιορίσει το φύλο από το μηριαίο οστό, την κνήμη ή το βραχιόνιο. Επιπλέον, η εφαρμογή παρέχει και τη σπίνθια πιθανότητα (posterior probability) κάθε εκτίμησης. Δεδομένου ότι υπάρχουν ποικίλοι συνδυασμοί μεταβλητών και κάθε μοντέλο (SVM, DFA) έχει διαφορετικές στατιστικές ιδιότητες, συνιστούμε στον χρήστη την αξιοποίηση όλων για τον προσδιορισμό του φύλου ενός δείγματος αγνώστων στοιχείων και την συνεκτίμηση των οπίσθιων πιθανοτήτων. Τέλος, θα πρέπει να σημειωθεί ότι η ανάπτυξη της προτεινόμενης μεθόδου στηρίχτηκε σε δείγμα σύγχρονου ελληνικού πληθυσμού και κατά συνέπεια η εφαρμογή της σε διαφορετικούς πληθυσμούς ενδέχεται να μην παρουσιάζει την ίδια ακρίβεια.

11. Η επίδραση της δραστηριότητας κατά την εκτίμηση ηλικίας θανάτου από τα οστά της λεκάνης

Παρά τη σημασία της ακριβούς εκτίμησης της ηλικίας-θανάτου από τα οστά και το πλήθος των διαθέσιμων μεθόδων, ο συσχετισμός μεταξύ των σκελετικών ηλικικών δεικτών και της χρονολογικής ηλικίας είναι ατελής, και επιδεινώνεται καθώς η χρονολογική ηλικία ενός ατόμου αυξάνεται. Το συγκεκριμένο πρόβλημα οφείλεται στο γεγονός ότι στους ενήλικες, η ηλικία θανάτου εκτιμάται σε μεγάλο βαθμό βάσει των εκφυλιστικών αλλοιώσεων του σκελετού, ο οποίος ελέγχεται από διάφορους παράγοντες πέρα από την ηλικία (Nawrocki, 2010). Ως αποτέλεσμα, υπάρχει μεγάλη ποικιλότητα μεταξύ ατόμων και μεταξύ πληθυσμών στο βαθμό των σκελετικών αλλοιώσεων και στη σχετική σκελετική γήρανση. Αυτό καθιστά τις σύγχρονες μεθόδους προσδιορισμού της ηλικίας θανάτου ατελείς και περιορίζει την εφαρμογή παγιωμένων ηλικιακών δεικτών σε διαφορετικούς πληθυσμούς (Bocquet-Appel & Masset, 1982).

Στόχος της παρούσας εργασίας είναι να εκτιμήσει κατά πόσον υψηλότερα επίπεδα σωματικής δραστηριότητας οδηγούν σε αύξηση των εκφυλιστικών αλλοιώσεων στην ηβική σύμφυση και στην ωτοειδή επιφάνεια. Το ανώνυμο οστό συσχετίζεται τόσο με την κίνηση όσο και με το βάρος. Κατά συνέπεια, ένας πιο σωματικά δραστήριος τρόπος ζωής αναμένεται να οδηγήσει σε αύξηση της πίεσης, η οποία μπορεί να επηρεάσει τον ρυθμό εκφυλιστικών αλλοιώσεων στην ηβική σύμφυση και στην ωτοειδή επιφάνεια. Η συγκεκριμένη μελέτη βασίζεται σε προηγούμενη έρευνα του Campanacho και των συναδέλφων του (Campanacho et al., 2012), οι οποίοι εξέτασαν την επίδραση της σωματικής δραστηριότητας στην ηβική σύμφυση και στην ωτοειδή επιφάνεια χρησιμοποιώντας μια σειρά σκελετικών δεικτών δραστηριότητας, και συμπεριλαμβάνοντας το φύλο και τη μάζα σώματος μεταξύ των παραγόντων πρόβλεψης.

11.1. Υλικά και μέθοδοι εκτίμησης ηλικίας θανάτου

Υλικό για την παρούσα εργασία αποτέλεσαν 131 σκελετοί από τη σύγχρονη σκελετική συλλογή αναφοράς που διατηρείται στον Τομέα Φυσιολογίας Ζώων και Ανθρώπου του Τμήματος Βιολογίας του Εθνικού Καποδιστριακού Πανεπιστημίου Αθηνών.

Για την εκτίμηση της ηλικίας θανάτου χρησιμοποιήθηκαν τρεις μέθοδοι: 1) των Brooks και Suchey (1990) για την ηβική σύμφυση, 2) του Lovejoy και συνεργατών (Lovejoy et al., 1985) για την ωτοειδή επιφάνεια, 3) Buckberry και Chamberlain (2002), επίσης για την ωτοειδή επιφάνεια. Παρακάτω θα αναφερόμαστε σε αυτές τις μεθόδους ως SB, L, και BC, αντιστοίχως.

Η συλλογή δεδομένων έλαβε χώρα στο πλαίσιο προηγούμενης μελέτης (Xanthoroulou et al., 2018), όπου διαπιστώθηκε ότι η συσχέτιση μεταξύ χρονολογικής ηλικίας και σκελετικών δεικτών

είναι στατιστικά σημαντική τόσο για τα δύο φύλα όσο και για τις τρεις μεθόδους ($r_s = 0,32 - 0,79$, $p\text{-value} < 0,01$; Πίνακας 6 στο Xanthoroulou et al., 2018). Ο βαθμός συμφωνίας μεταξύ της αναπαραγωγιμότητας και της επαναληψιμότητας των αποτελεσμάτων (intra-observer error) βρέθηκε επίσης ότι είναι στατιστικά σημαντικός (Cohen's kappa $> 0,7$, $p\text{-value} < 0,05$). Τέλος, το ποσοστό των ατόμων που ταξινομήθηκαν στη σωστή ηλικιακή ομάδα βάσει των προαναφερθέντων μεθόδων ήταν περίπου 70% και πάνω από 80% για τους άνδρες και τις γυναίκες αντίστοιχα όσον αφορά στη μέθοδο SB, τουλάχιστον 78% και για τα δύο φύλα όσον αφορά στην L μέθοδο, και μόνο 56% όσον αφορά στη μέθοδο BC (Πίνακας 2 στο Xanthoroulou et al., 2018).

11.2. Εκτίμηση του βαθμού σωματικής δραστηριότητας/καταπόνησης

Στη σύγχρονη σκελετική συλλογή αναφοράς, τα πιστοποιητικά θανάτου αναφέρουν συνήθως μόνο το τελευταίο δηλωθέν επάγγελμα και πολλά επαγγέλματα είναι ασαφή, όπως «δημόσιος υπάλληλος». Οι συγκεκριμένοι περιορισμοί δεν επιτρέπουν την εκτίμηση του βαθμού σωματικής δραστηριότητας/καταπόνησης των ατόμων. Ως εκ τούτου, τα επίπεδα δραστηριότητας του κάθε ατόμου καθορίστηκαν σκελετικά τόσο με τη χρήση γεωμετρικών χαρακτηριστικών της διατομής του μηριαίου (CSG δείκτες) όσο και μέσω της εξέτασης διάφορων χαρακτηριστικών αλλοίωσης των ινοchonδρικών ενθέσεων αυτού.

Οι CSG δείκτες που χρησιμοποιήθηκαν περιλαμβάνουν τη συνολική υποπεριοστική επιφάνεια (TA), η οποία αποτελεί μια μέτρηση της ευρωστίας του οστού, συγκεκριμένα σχετικά με την ασκούμενη μηχανική πίεση ή τη συμπίεση, καθώς και τους λόγους I_x/I_y και I_{MAX}/I_{MIN} , όπου I_x , I_y , I_{MAX} και I_{MIN} εκφράζουν την αντίσταση σε φορτία κάμψης και οι λόγοι τους είναι ενδεικτικοί της διατομής, αφού αντιπροσωπεύουν την κατανομή του φλοιώδους οστού μέσα στην τομή (Ruff, 2008). Προς αποφυγήν εισαγωγής υπερβολικά πολλών αλληλοεπηρεαζόμενων παραμέτρων στα στατιστικά τεστ, χρησιμοποιήθηκαν μόνο τα χαρακτηριστικά που υπολογίστηκαν στο 50% του συνολικού μήκους του οστού. Αυτή η τεχνική χρήσης του συγκεκριμένου μήκους ως επαρκής μέτρηση, η οποία εκφράζει την αντίσταση της διάφυσης ενάντια στις μηχανικές πιέσεις, έχει υιοθετηθεί από αρκετούς επιστήμονες (Sparacello & Pearson, 2010; Stock & Shaw, 2007).

Οι ενθέσεις που εξετάστηκαν στο μηριαίο περιλάμβαναν αυτή του ελάσσων γλουτιαίος μυ (*gluteus minimus*), του μείζων ψοΐτη μυ (*psoas major*), του λαγόνιου μυ (*iliacus*), και του μέσου γλουτιαίου μυ (*gluteus medius*). Όλες οι ενθέσεις καταγράφηκαν στα κάτω άκρα της αριστερής πλευράς, εκτός και αν αυτά ήταν απόντα. Η μέθοδος που χρησιμοποιήθηκε για την εξέταση των αλλοιώσεων στις ενθέσεις ήταν η New Coimbra Method (Henderson et al., 2016). Η καταγραφή των ενθέσεων σε ένα δείγμα 22 μηριαίων πραγματοποιήθηκε δύο φορές από τον ίδιο παρατηρητή με διάστημα τριών εβδομάδων μεταξύ των δύο καταγραφών. Ο έλεγχος επαναληψιμότητας για

κάθε μεταβλητή ξεχωριστά βασίστηκε στη διαφορά μεταξύ των δύο μετρήσεων και έδειξε ότι για όλες τις εξεταζόμενες ενθέσεις η μέση τιμή είναι 99,02%.

11.3. Εκτίμηση σωματικής μάζας

Η σωματική μάζα υπολογίστηκε από τη διάμετρο της κεφαλής του μηριαίου και τη χρήση εξισώσεων παλινδρόμησης που έχουν αναπτυχθεί από τη Νικήτα και τη Χοβαλοπούλου (Nikita & Chovalopoulou, 2017) με βάση την ίδια τη σύγχρονη σκελετική συλλογή αναφοράς. Οι συγκεκριμένες εξισώσεις εμφάνισαν χαμηλό τυχαίο και συστηματικό σφάλμα. Σε αυτό το σημείο πρέπει να τονίσουμε ότι μια πρόσφατη μελέτη διαπίστωσε ότι η εκτίμηση της σωματικής μάζας σώματος είναι ακριβής μόνο όσον αφορά στη μέση σωματική μάζα ενός πληθυσμού και όχι όσον αφορά στα άτομα (Lacoste Jeanson et al., 2017). Παρόλα αυτά, η μάζα σώματος στην συγκεκριμένη εργασία χρησιμοποιείται μόνο ως προσεγγιστική μεταβλητή προκειμένου να εξεταστεί εάν άτομα με μεγαλύτερη μάζα σώματος εμφανίζουν εντονότερες εκφυλιστικές αλλοιώσεις στην ηβική σύμφυση και στην ωτοειδή επιφάνεια και της ηβικής σύμφυσης.

11.4. Στατιστική ανάλυση

Αρχικά, αξιολογήθηκε το αν υπάρχει σημαντική αμφίπλευρη ασυμμετρία στις τιμές των σκελετικών δεικτών ηλικίας, ώστε να ελεγχθεί αν είναι αναγκαίο να εξεταστεί η σχέση δραστηριότητας και σκελετικής ηλικίας σε κάθε πλευρά ξεχωριστά ή όχι. Προς εκτίμηση αυτής της αμφίπλευρης ασυμμετρίας χρησιμοποιήθηκε ο έλεγχος προσημασμένης διάταξης Wilcoxon (Wilcoxon signed ranks tests). Ακολούθως, η συσχέτιση μεταξύ σκελετικών δεικτών ηλικίας και δεικτών CSG ελέγχθηκε με τον συντελεστή συσχέτισης Spearman (Spearman correlations). Η μηδενική υπόθεση ελέγχθηκε με Monte-Carlo με αντιμεταθέσεις (1000 επαναλήψεις). Στη συνέχεια ακολούθησαν μερικές συσχετίσεις Spearman (partial Spearman correlations) ώστε να ελεγχθεί και πάλι η συσχέτιση μεταξύ των σκελετικών δεικτών ηλικίας και των δεικτών CSG, αυτή τη φορά όμως ελέγχοντας για την χρονολογική ηλικία θανάτου και τη σωματική μάζα. Ο συντελεστής Spearman χρησιμοποιήθηκε επίσης για τον έλεγχο της συσχέτισης μεταξύ των σκελετικών ηλικιακών δεικτών και των αθροιστικών αποτελεσμάτων για τις αλλοιώσεις στις ενθέσεις, ενώ με μερικές συσχετίσεις Spearman αξιολογήθηκε αυτή η συσχέτιση ελέγχοντας για την επίδραση της χρονολογικής ηλικίας θανάτου και της σωματικής μάζας. Οι μερικές συσχετίσεις Spearman με τη χρονολογική ηλικία και τη σωματική μάζα ως συμμεταβλητές χρησιμοποιήθηκαν και πάλι για τις αλλοιώσεις στις ενθέσεις, αυτή τη φορά για κάθε χαρακτηριστικό αλλοίωσης ξεχωριστά, και όχι για τα αθροιστικά αποτελέσματα, ώστε να εκτιμηθεί αν παραβλέφθηκε κάποιο σημαντικό μοτίβο συσχέτισης λόγω της αθροιστικής διαδικασίας. Επιπλέον, χρησιμοποιήθηκαν γενικευμένα

γραμμικά μοντέλα (GLM) για να εξεταστεί ταυτόχρονα η επίδραση του φύλου, της χρονολογικής ηλικίας θανάτου, της σωματικής μάζας, των αθροιστικών αλλοιώσεων των ενθέσεων και των δεικτών CSG στους σκελετικούς ηλικιακούς δείκτες. Τα GLM πραγματοποιήθηκαν δύο φορές, στην μία περίπτωση τα αθροίσματα των αλλοιώσεων διαχειρίστηκαν ως διατάξιμες μεταβλητές, ενώ στην άλλη ως συνεχείς μεταβλητές.

Λαμβάνοντας υπόψη ότι η ακρίβεια των μεθόδων εκτίμησης ηλικίας θανάτου τείνει να μειώνεται σε ηλικιωμένα άτομα, ως ένα πρόσθετο μέσο διερεύνησης της επίδρασης της δραστηριότητας κατά την εκτίμηση ηλικίας θανάτου, χωρίσαμε το δείγμα σε άτομα ηλικίας κάτω των 50 ετών και άνω των 50 ετών. Στη συνέχεια, ταξινομήθηκε κάθε άτομο σε μία από τις τρεις κατηγορίες: υποεκτίμηση ηλικίας, σωστή εκτίμηση ηλικίας και υπερεκτίμηση ηλικίας, βάση της επίδοσης κάθε μεθόδου εκτίμησης ηλικίας θανάτου. Πιο συγκεκριμένα, η ταξινόμηση βασίστηκε στο κατά πόσον η εκτιμώμενη ηλικία ενός ατόμου εμπίπτει στο εύρος του μέσου όρου ηλικίας \pm την τυπική απόκλιση που παρέχεται στη βιβλιογραφία για τις μεθόδους SB και BC και εντός των ηλικιακών ευρών που προτείνονται για τη μέθοδο L. Στη συνέχεια, οι αλλοιώσεις των ενθέσεων (EC) και οι δείκτες CSG συγκρίθηκαν μεταξύ των τριών κατηγοριών με τη μέθοδο Kruskal-Wallis προκειμένου να αξιολογηθεί εάν οι συγκεκριμένοι δείκτες σωματικής δραστηριότητας διαφοροποιούνται σημαντικά σε άτομα για τα οποία η εκτίμηση ηλικίας θανάτου παρουσιάζει αποκλίσεις από την χρονολογική ηλικία θανάτου.

11.5. Αποτελέσματα και Συμπεράσματα

Χρησιμοποιήθηκε μόνο η αριστερή πλευρά της λεκάνης για τη μελέτη της επίδρασης της δραστηριότητας στους ηλικιακούς δείκτες, καθώς η διαφορά στην αμφίπλευρη έκφραση των εξεταζόμενων σκελετικών ηλικιακών δεικτών βρέθηκε στατιστικά μη σημαντική (SB: Wilcoxon Z = -0,218, p-value = 0,827; L: Wilcoxon Z = -1,753, p-value = 0,080; BC: Wilcoxon Z = -0,865, p-value = 0,387).

Αρχικά, ελέγχθηκε η συσχέτιση μεταξύ των σκελετικών δεικτών ηλικίας και των δεικτών CSG με τον συντελεστή συσχέτισης Spearman (Spearman correlations). Σύμφωνα με τα αποτελέσματα, όταν εξετάζονται τα δύο φύλα συγκεντρωτικά, η συσχέτιση είναι πολύ μικρή (τιμή rho από -0,225 έως 0,149). Παρ' όλα αυτά, η συσχέτιση μεταξύ του λόγου I_x/I_y και των τριών σκελετικών δεικτών είναι στατιστικά σημαντική (p-values < 0,05) και οι αρνητικοί συντελεστές υποδηλώνουν ότι όσο η σκελετική αλλοίωση αυξάνεται, η αντίσταση του σκελετού σε καμπτικές δυνάμεις, η οποία εκφράζεται ως η κατανομή του φλοιώδους οστού μεταξύ των πρόσθιων-οπίσθιων και μεσοπλευρικών αξόνων του επιπέδου της διατομής, γενικά μειώνεται, γεγονός που είναι αναμενόμενο. Αντίστοιχα, μείωση στην αντίσταση του σκελετού σε δυνάμεις συμπίεσης,

εφελκυσμού και διάτμησης όσο αυξάνεται η σκελετική αλλοίωση υποδεικνύεται από τον αρνητικό συντελεστή ρ , σχετικά με το δείκτη TA. Ωστόσο, η παρατήρηση αυτή είναι στατιστικά σημαντική μόνο στην ηβική σύμφυση (p -value = 0,03). Οι θετικές τιμές του συντελεστή ρ που παρατηρούνται για το λόγο I_{MAX}/I_{MIN} υποδηλώνουν μία αύξηση στην αντίσταση των οστών καθώς αυξάνεται η σκελετική αλλοίωση, γεγονός που φαινομενικά αντικρούει τα αποτελέσματα του I_X/I_Y . Ωστόσο, δεδομένου ότι τα I_{MAX} και I_{MIN} αντιπροσωπεύουν την κάθετη μέγιστη και ελάχιστη καμπτική αντίσταση κατά μήκος ενός αυθαίρετου άξονα του επιπέδου της διατομής του οστού και ο δείκτης TA μειώνεται με την αύξηση της σκελετικής ηλικίας, το αποτέλεσμα αυτό μπορεί να δείχνει πως η περιοχή της διατομής γίνεται πιο επιμήκης, υποδηλώνοντας ότι η μηριαία διάφυση εξασθενεί ακόμα περισσότερο κατά μήκος κάποιου συγκεκριμένου προσανατολισμού της καμπτικής δύναμης. Παρόλα αυτά, αυτή η συσχέτιση ήταν στατιστικά σημαντική μόνο για τη μέθοδο BC (p -value = 0,04). Οι μερικές συσχετίσεις Spearman (partial Spearman correlations) για τον έλεγχο της συσχέτισης μεταξύ των σκελετικών δεικτών ηλικίας και των δεικτών CSG, ελέγχοντας για την χρονολογική ηλικία θανάτου και τη σωματική μάζα, έδειξαν ότι η μόνη στατιστικά σημαντική συσχέτιση σχετικά με τα συγκεντρωτικά αποτελέσματα των δύο φύλων είναι μεταξύ του δείκτη TA και της μεθόδου BC, ενώ οι συντελεστές ρ προσεγγίζουν το μηδέν. Κατά συνέπεια, η συσχέτιση μεταξύ δραστηριότητας και σκελετικής αλλοίωσης στην ηβική σύμφυση και την ωτοειδή επιφάνεια είναι περιορισμένη. Σύμφωνα με τα αποτελέσματα των αντίστοιχων αναλύσεων για τα αρσενικά και θηλυκά άτομα ξεχωριστά, παρατηρείται σημαντική μείωση του αριθμού των στατιστικά σημαντικών συσχετίσεων μεταξύ των δεικτών CSG και των σκελετικών δεικτών ηλικίας. Πιο συγκεκριμένα, κατά τις αναλύσεις των μερικών συσχετίσεων, εντοπίστηκαν μόνο δύο στατιστικά σημαντικές συσχετίσεις (αρσενικά άτομα: {BC – I_{MAX}/I_{MIN} }: $\rho = 0,306$, p -value = 0,004 ~ θηλυκά άτομα: {BC – TA}: $\rho = -0,239$, p -value = 0,014). Η μείωση του αριθμού των στατιστικά σημαντικών συσχετίσεων ίσως αποδίδεται στο μικρότερο, διαχωρισμένο πλέον στα δύο φύλα, μέγεθος του δείγματος. Σε κάθε περίπτωση όμως, δεν φαίνεται να υπάρχει σαφής φυλετικός διμορφισμός όσον αφορά στην συσχέτιση μεταξύ των σκελετικών δεικτών ηλικίας και των δεικτών CSG.

Στη συνέχεια, ελέγχθηκε η συσχέτιση μεταξύ των αθροιστικών αποτελεσμάτων των αλλοιώσεων των ενθέσεων και των σκελετικών δεικτών ηλικίας με τον συντελεστή συσχέτισης Spearman (Spearman correlations), όπου τα συγκεντρωτικά δεδομένα των δύο φύλων δίνουν μια εικόνα όπου η συσχέτιση είναι πάντα θετική (οι συντελεστές ρ κυμαίνονται από 0,156 έως 0,642) και στατιστικά σημαντική. Αυτό υποδεικνύει ότι τα γηραιότερα άτομα παρουσιάζουν πιο έντονες αλλοιώσεις στις ενθέσεις. Ωστόσο, σύμφωνα με τα αποτελέσματα των μερικών συσχετίσεων Spearman, οι συντελεστές συσχέτισης μειώνονται αισθητά (-0,145 σε 0,165) και η συσχέτιση

μεταξύ των αλλοιώσεων των ενθέσεων και των σκελετικών αλλοιώσεων στην ηβική σύμφυση και την ωτοειδή επιφάνεια είναι κατά κύριο λόγο μη στατιστικά σημαντική. Το γεγονός αυτό υποδηλώνει ότι η στατιστικά σημαντική συσχέτιση μεταξύ των αλλοιώσεων των ενθέσεων και των σκελετικών δεικτών ηλικίας οφείλεται κυρίως στην επίδραση της χρονολογικής ηλικίας (και σε κάποιο βαθμό και στη σωματική μάζα). Η συσχέτιση μεταξύ δραστηριότητας (αυξημένες τιμές στις ενθέσεις) και σκελετικών ηλικιακών δεικτών απουσιάζει από τη στιγμή που απομακρύνονται οι επιδράσεις της χρονολογικής ηλικίας και της σωματικής μάζας. Έχοντας στο νου ότι η παρούσα εργασία διερευνά το εάν η αυξημένη δραστηριότητα μπορεί να επιταχύνει τη σκελετική φθορά στην ηβική σύμφυση και την ωτοειδή επιφάνεια, συμπεραίνουμε πως τα αυξημένα επίπεδα φυσικής δραστηριότητας, όπως αυτά έχουν εκτιμηθεί μέσω των αλλοιώσεων των μηριαίων ενθέσεων, δεν επιταχύνουν τις αλλοιώσεις στις συγκεκριμένες περιοχές, το οποίο ισχύει και για τα δύο φύλα.

Τα αποτελέσματά των GLM σύμφωνα με τα οποία πραγματοποιήθηκε η ταυτόχρονη εξέταση της επίδρασης του φύλου, της σωματικής μάζας, της χρονολογικής ηλικίας θανάτου, των αλλοιώσεων των ενθέσεων και των δεικτών CSG στους σκελετικούς ηλικιακούς δείκτες, συνηγορούν στο ότι η ηλικία είναι ο μόνος παράγοντας ο οποίος συστηματικά έχει σημαντική επίδραση στην έκφραση των ηλικιακών σκελετικών δεικτών. Στην ηβική σύμφυση, το φύλο εμφανίζεται να παίζει σημαντικό ρόλο στην έκφραση των ηλικιακών σκελετικών δεικτών ανεξάρτητα από τη δραστηριότητα, τη σωματική μάζα και τη χρονολογική ηλικία. Πιθανόν αυτό να οφείλεται σε ορμονικές διαφορές (Mays, 2015).

Τέλος, τα αποτελέσματα των αναλύσεων Kruskal-Wallis (Πίνακες 11.1 – 11.2) επιβεβαιώνουν σε μεγάλο βαθμό τα παραπάνω ευρήματα ενώ ταυτόχρονα προσφέρουν και κάποιες επιπλέον πληροφορίες. Όσον αφορά στους CSG δείκτες (Πίνακας 11.1), μόνο ο TA εμφανίζει, και αυτό σπανίως, στατιστικά σημαντικές διαφορές μεταξύ των τριών κατηγοριών, υποστηρίζοντας την έλλειψη επίδρασης της δραστηριότητας, όπως αυτή περιγράφεται μέσω των CSG δεικτών, στις εκφυλιστικές αλλοιώσεις της ηβικής σύμφυσης και της ωτοειδούς επιφάνειας.

Πίνακας 11.1: Αποτελέσματα ανάλυσης Kruskal-Wallis για τους CSG δείκτες.

Μέθοδος εκτίμησης ηλικίας θανάτου	Ομάδες	TA				Ix/Iy				Imax/Imin			
		Μέσο της κατάταξης				Μέσο της κατάταξης				Μέσο της κατάταξης			
		1	2	3	p-value	1	2	3	p-value	1	2	3	p-value
SB	Συνολικό	73,7	66,5	40,6	0,019	60,0	67,33	76,5	0,332	66,1	68,0	54,7	0,483
	Αρσενικά	45,5	35,6	32,2	0,166	34,4	40,1	43,4	0,521	41,2	37,9	31,0	0,615
	Θηλυκά	29,4	31,0	15,6	0,037	25,3	27,2	35,7	0,270	25,4	31,2	22,6	0,260
	<=50 ετών	29,7	29,8	19,3	0,093	14,7	28,6	25,6	0,299	12,7	29,7	23,2	0,105
	>50 ετών	42,5	36,7	-	0,254	41,1	38,0	-	0,549	40,2	38,8	-	0,780
	A <=50 yrs	18	17,1	16,2	0,972	10,5	17,7	16	0,580	8,5	18,2	14,4	0,320
	A >50yrs	24,5	19,4	-	0,181	21,7	22,3	-	0,884	23,7	20,2	-	0,356
	Θ <=50 yrs	13,0	11,3	9,3	0,701	2,0	11,3	10,6	0,325	5,0	12,8	8,6	0,188
	Θ >50yrs	16,8	19,0	-	0,529	19,2	17,0	-	0,529	16,8	19,1	-	0,508
L	Συνολικό	73,7	61,8	66,6	0,340	59,6	63,8	76,4	0,158	57,1	71,9	63,6	0,174
	Αρσενικά	40,4	36,9	39,0	0,843	31,3	39,4	44,4	0,136	30,9	44,4	37,5	0,084
	Θηλυκά	29,2	29,7	22,5	0,403	26,0	27,1	32,2	0,583	28,5	28,4	26,6	0,942
	<=50 ετών	26,5	24,0	28,5	0,698	22,0	28,3	28,6	0,443	18,7	30,1	29,2	0,103
	>50 ετών	47,8	35,7	44,0	0,107	38,9	40,2	34,4	0,851	37,9	42,1	19,6	0,099
	A <=50 yrs	18,3	12,0	18,7	0,243	11,8	14,6	19,6	0,165	11,5	17,6	18,5	0,299
	A >50yrs	22,0	21,8	24,0	0,958	20,9	23,9	12,7	0,310	18,2	25,6	13,3	0,086
	Θ <=50 yrs	11,7	11,3	9,5	0,747	10,3	13,5	9,4	0,502	7,2	13,8	11,2	0,197
	Θ >50yrs	20,4	17,2	23,5	0,597	12,4	18,7	22,5	0,368	25,8	17,5	5,5	0,051
BC	Συνολικό	85,5	69,5	59,5	0,149	59,2	61,0	73,2	0,185	86,2	63,1	67,5	0,337
	Αρσενικά	67,3	39,2	34,8	0,049	34,0	35,2	43,5	0,267	62,0	35,6	40,2	0,115
	Θηλυκά	37,0	30,8	23,6	0,162	27,0	24,9	31,7	0,312	28,0	28,4	27,6	0,985
	<=50 ετών	-	34,1	25,4	0,107	-	25,6	27,3	0,750	-	20,4	28,5	0,134
	>50 ετών	49,8	38,4	40,1	0,494	41,2	39,7	37,5	0,936	49,7	38,5	39,6	0,514
	A <=50 yrs	-	18,0	16,6	0,716	-	14,0	18,1	0,275	-	12,9	18,5	0,135
	A >50yrs	37,0	20,8	21,3	0,100	21,0	22,1	22,1	0,990	33,0	20,8	23,0	0,265
	Θ <=50 yrs	-	15,0	10,3	0,435	-	15,0	10,3	0,435	-	10,0	10,5	0,931
	Θ >50yrs	22,3	17,4	18,4	0,732	20,7	17,9	16,8	0,872	18,0	18,3	16,2	0,913

Υπόμνημα: 1 = υποεκτίμηση ηλικίας, 2 = σωστή εκτίμηση ηλικίας, 3 = υπερεκτίμηση ηλικίας

Πίνακας 11.2: Αποτελέσματα ανάλυσης Kruskal-Wallis για τα αθροιστικά αποτελέσματα των αλλοιώσεων των ενθέσεων

Μέθοδος εκτίμησης ηλικίας θανάτου	Ομάδες	Ενθέσεις ελάσσων γλουτιαίου μ				Ενθέσεις μείζων φοΐτη μ				Ενθέσεις λαγόνιου μ				Ενθέσεις μέσου γλουτιαίου μ			
		Μέσο της κατάταξης				Μέσο της κατάταξης				Μέσο της κατάταξης				Μέσο της κατάταξης			
		1	2	3	p-value	1	2	3	p-value	1	2	3	p-value	1	2	3	p-value
SB	Συνολικό	91,8	57,4	37,0	0,000	87,7	57,7	37,0	0,000	80,2	59,2	56,4	0,006	85,8	59,1	35,5	0,000
	Αρσενικά	53,4	32,4	24,5	0,000	49,5	32,6	24,3	0,002	46,1	33,7	38,4	0,059	49,5	33,9	21,1	0,002
	Θηλυκά	40,5	25,7	11,9	0,000	40,0	26,1	11,6	0,000	34,7	26,2	21,4	0,061	37,1	25,9	14,3	0,001
	<=50 ετών	21,5	27,4	27,3	0,780	22,7	25,9	28,8	0,722	32,5	25,9	28,6	0,649	22,7	28,2	23,1	0,438
	>50 ετών	46,9	32,5	-	0,004	44,3	33,8	-	0,034	43,6	34,5	-	0,060	44,6	33,5	-	0,025
	A <=50 yrs	17,8	16,9	17,5	0,979	14,3	16,4	17,7	0,890	19,3	16,3	19,8	0,673	17,3	17,5	14,1	0,710
	A >50yrs	25,5	18,4	-	0,057	22,9	19,9	-	0,417	23,9	18,9	-	0,164	23,2	19,6	-	0,321
	Θ <=50 yrs	4,0	12,2	9,4	0,251	9,0	11,2	9,9	0,833	14,0	10,1	10,6	0,778	6,0	11,6	8,9	0,379
	Θ >50yrs	23,0	13,8	-	0,007	22,9	13,8	-	0,007	20,3	16,1	-	0,195	22,2	14,5	-	0,024
L	Συνολικό	71,7	75,7	42,3	0,000	66,2	79,4	37,6	0,000	71,5	72,5	46,7	0,001	67,9	74,4	44,9	0,001
	Αρσενικά	46,8	40,4	27,4	0,009	41,4	44,8	23,6	0,001	45,8	39,9	27,6	0,011	42,5	43,3	26,0	0,006
	Θηλυκά	26,4	33,3	15,4	0,003	25,5	33,8	14,9	0,001	24,3	32,4	19,8	0,029	26,4	30,8	19,6	0,092
	<=50 ετών	28,5	21,8	28,5	0,337	32,1	28,1	23,6	0,193	26,3	33,5	24,6	0,177	29,8	25,5	25,7	0,655
	>50 ετών	43,8	39,4	22,8	0,169	33,5	41,7	34,5	0,315	44,3	38,4	23,2	0,124	37,1	40,1	36,2	0,828
	A <=50 yrs	17,8	11,9	18,9	0,143	18,3	17,5	15,6	0,758	20,2	19,2	15,1	0,344	21,1	15,3	16,4	0,411
	A >50yrs	25,1	21,4	10,0	0,139	19,3	22,9	22,0	0,654	24,8	20,2	14,2	0,248	20,0	23,0	18,0	0,648
	Θ <=50 yrs	10,7	11,6	10,0	0,868	13,2	10,9	8,8	0,278	8,0	15,4	10,1	0,096	10,1	10,6	9,7	0,951
	Θ >50yrs	24,8	17,2	12,8	0,225	16,7	18,7	11,5	0,583	20,0	18,3	9,3	0,383	19,6	17,6	19,0	0,913
BC	Συνολικό	112,8	78,4	45,1	0,000	113,3	79,4	40,5	0,000	98,0	70,1	55,9	0,006	118,3	76,4	44,6	0,000
	Αρσενικά	63,5	44,5	28,0	0,001	64,5	44,9	24,1	0,000	66,3	39,8	32,6	0,018	70,0	44,7	26,1	0,000
	Θηλυκά	50,8	35,1	16,8	0,000	49,7	36,0	15,9	0,000	34,5	30,9	23,8	0,173	49,0	32,4	18,7	0,000
	<=50 ετών	-	18,3	29,0	0,028	-	24,9	26,9	0,689	-	24,5	27,6	0,534	-	29,9	25,7	0,375
	>50 ετών	60,6	40,6	23,3	0,003	61,8	40,3	19,6	0,000	53,1	38,6	33,7	0,184	66,6	38,7	26,6	0,001
	A <=50 yrs	-	13,4	18,3	0,138	-	16,5	16,5	1,000	-	16,1	17,3	0,725	-	18,9	16,3	0,422
	A >50yrs	31,2	23,0	13,2	0,065	32,8	22,5	10,3	0,017	35,5	21,1	16,6	0,066	37,0	21,4	15,3	0,029
	Θ <=50 yrs	-	4,0	10,8	0,217	-	16,5	10,2	0,252	-	5,0	10,8	0,292	-	13,5	9,8	0,473
	Θ >50yrs	30,8	18,0	10,4	0,022	29,8	18,3	9,4	0,018	19,2	17,9	17,7	0,975	30,2	17,8	11,9	0,043

Υπόμνημα: 1 = υποεκτίμηση ηλικίας, 2 = σωστή εκτίμηση ηλικίας, 3 = υπερεκτίμηση ηλικίας

Σχετικά με τις αλλοιώσεις των ενθέσεων (Πίνακας 11.2), όταν το δείγμα δεν χωρίζεται σε ηλικιακές ομάδες, εμφανίζουν σχεδόν πάντα στατιστικά σημαντικές διαφορές μεταξύ των τριών κατηγοριών (υποεκτίμηση ηλικίας, σωστή εκτίμηση ηλικίας και υπερεκτίμηση ηλικίας). Ωστόσο, όταν το δείγμα είναι χωρισμένο σε ηλικιακές ομάδες, στατιστικά σημαντικές διαφορές παρατηρούνται σχεδόν αποκλειστικά στα άτομα άνω των 50 ετών. Το γεγονός αυτό μας οδηγεί στο συμπέρασμα ότι οι διαφορές αυτές πρέπει να σχετίζονται πρωτίστως με την καλά τεκμηριωμένη επίδραση που έχει η χρονολογική ηλικία στην έκφραση των αλλοιώσεων των ενθέσεων παρά στην επίδραση των τελευταίων στις εκφυλιστικές αλλοιώσεις των πυελικών αρθρώσεων. Είναι ενδιαφέρον να σημειωθεί ότι το παραπάνω μοτίβο δεν ήταν ορατό στην L μέθοδο εκτίμησης ηλικίας θανάτου, σε αντίθεση με τη μέθοδο BC, όπου είναι πολύ σαφές. Παρόλο που και οι δύο μέθοδοι επικεντρώνονται στην ίδια ανατομική περιοχή, υιοθετούν διαφορετικούς τρόπους καταγραφής μορφολογικών αλλαγών και συσχετίζουν διαφορετικά ηλικιακά διαστήματα με διαφορετικές μορφολογικές αλλαγές (ιδιαίτερως, η μέθοδος L τοποθετεί όλα τα άτομα άνω των 60 ετών σε μία ηλικιακή ομάδα). Επιπλέον, παρατηρούμε ότι όταν εφαρμόζεται η SB μέθοδος εκτίμησης ηλικίας θανάτου, στατιστικά σημαντικές διαφορές στην έκφραση των αλλοιώσεων των ενθέσεων μεταξύ των τριών κατηγοριών εντοπίζονται μόνο στα θηλυκά άτομα άνω των 50 ετών αλλά όχι στα αρσενικά άτομα. Αυτό το φυλετικά διμορφικό μοτίβο μπορεί να συνδέεται με ορμονικές αλλαγές που σχετίζονται με την εμμηνόπαυση ή άλλες σχετιζόμενες με το φύλο φυσιολογικές και λειτουργικές διαφορές, όπως ο τοκετός (Becker et al., 2010; Ji & Yu, 2015).

Τα αποτελέσματά της παρούσας εργασίας στο σύνολό τους υποδηλώνουν ότι υπάρχει σημαντικός συσχετισμός μεταξύ των σκελετικών δεικτών δραστηριότητας/καταπόνησης και των ηλικιακών σκελετικών δεικτών. Αυτός ο συσχετισμός είναι ιδιαίτερα εμφανής στο λόγο Ix/Iy και στο TA όσον αφορά στους CSG δείκτες, καθώς και σε όλες τις αλλοιώσεων των ενθέσεων. Ωστόσο, όταν “απομακρύνεται” η επίδραση της χρονολογικής ηλικίας και της σωματικής μάζας, ο συσχετισμός μειώνεται, υποδηλώνοντας ότι αυτή η συσχέτιση αποδίδεται άμεσα στη χρονολογική ηλικία των ατόμων. Παρομοίως, οι διαφορές στους CSG δείκτες μεταξύ των τριών κατηγοριών εκτίμησης ηλικίας είναι σπάνιες, ενώ για τις αλλοιώσεις των ενθέσεων εντοπίζονται συστηματικά διαφορές μόνο σε άτομα ηλικίας άνω των 50 ετών, υποστηρίζοντας τον βασικό ρόλο της χρονολογικής ηλικίας στην έκφραση των μορφολογικών αλλαγών τόσο στην ηβική σύμφυση όσο και στις ενθέσεις. Ο συνδυασμός αυτών των αποτελεσμάτων υποδεικνύει ότι δεν υπάρχει άμεση επίδραση της δραστηριότητας στις εκφυλιστικές αλλοιώσεις της ηβικής σύμφυσης και της ωτοειδούς επιφάνειας.

Ένα σημείο που αξίζει να επισημανθεί είναι η πιθανή επίδραση της μάζας σώματος. Κάποιος θα μπορούσε να υποστηρίξει ότι τα βαρύτερα άτομα αναμένεται να εμφανίζουν υψηλότερες τιμές στους ηλικιακούς δείκτες σε μικρότερη ηλικία από την κανονική, υπό την έννοια ότι τα αυξημένα μηχανικά φορτία της μάζας σώματος στη λεκάνη αναμένεται να έχουν ως αποτέλεσμα ταχύτερο εκφυλισμό των πυελικών αρθρώσεων σε μεγαλύτερα άτομα. Ωστόσο, δεν βρέθηκε σημαντική επίδραση η μάζας σώματος στα μοντέλα GLM. Η παρατήρηση αυτή είναι σε συμφωνία με προηγούμενες μελέτες. Οι Wescott και Drew (2015), οι οποίοι μελέτησαν σύγχρονες σκελετικές συλλογές των ΗΠΑ, διαπίστωσαν ότι τα παχύσαρκα άτομα είχαν εντονότερες εκφυλιστικές αλλοιώσεις από τις αναμενόμενες βάσει χρονολογικής ηλικίας στην ωτοειδή επιφάνεια αλλά όχι στην ηβική σύμφυση, ενώ η Merritt (2015) διαπίστωσε ότι τα άτομα με μικρό μέγεθος σώματος τείνουν να γερνούν νωρίτερα σύμφωνα με τις εκφυλιστικές αλλοιώσεις της ηβικής σύμφυσης, της ωτοειδούς επιφάνειας και του στερνικού άκρου των πλευρών.

12. Προηγμένη μέθοδος εκτίμησης φύλου βάση φυλετικά διμορφικών κρανιακών μορφολογικών χαρακτηριστικών

Η εκτίμηση του φύλου από το κρανίο πραγματοποιείται συνήθως είτε με οπτική παρατήρηση είτε με διάφορες μορφομετρικές μεθόδους. Η οπτική παρατήρηση βασίζεται στην υποκειμενική αξιολόγηση των φυλετικά διμορφικών περιοχών του κρανίου, όπως το μεσόφρυο (glabella) ή η μαστοειδής απόφυση. Οι περιοχές αυτές σκοράρονται εμπειρικά σύμφωνα με τυποποιημένα πρωτόκολλα και στη συνέχεια χρησιμοποιούνται για την εκτίμηση φύλου (Godde et al., 2018; Langley et al., 2018; Walker, 2008). Οι μορφομετρικές μέθοδοι αντιθέτως, μπορεί να περιλαμβάνουν μια σειρά τεχνικών μέτρησης, όπως οι τυπικές μετρήσεις με παχύμετρο (Casado, 2017; Franklin et al., 2013; Kranioti et al., 2008), η γεωμετρική μορφομετρία με αναφορικά σημεία (Bigoni et al., 2010; Choaloroulou et al., 2013) ή ακόμη και η ανάλυση επιφάνειας τριδιάστατων μοντέλων, η οποία εστιάζει είτε σε συγκεκριμένα φυλετικά διμορφικά χαρακτηριστικά (Bejdoná et al., 2018; Čechová et al., 2019; Garvin & Ruff, 2012; Schlager & Rüdell, 2017) είτε σε ολόκληρο το κρανίο (Arigbabu et al., 2017). Λαμβάνοντας υπόψη τα πλεονεκτήματα και τα μειονεκτήματα της κάθε μεθόδου, η οπτική παρατήρηση, μολονότι είναι ευκολότερη και ταχύτερη, τείνει να βασίζεται στην εμπειρία του παρατηρητή και έχει επικριθεί για υψηλή υποκειμενικότητα και φτωχή αναπαραγωγικότητα (C. J. Lewis & Garvin, 2016; Walrath et al., 2004). Αντιθέτως, οι μορφομετρικές μέθοδοι είναι συχνά πιο ακριβείς και δεν έχουν έντονα προβλήματα αναπαραγωγικότητας (Fruciano, 2016; Shearer et al., 2017). Ωστόσο, είναι πιο περίπλοκες και συχνά αποτυγχάνουν να αποτυπώσουν επαρκώς τις ανεπαίσθητες φυλετικά διμορφικές διακυμάνσεις (Walker, 2008). Επιπλέον, οι διακριτικές εξισώσεις εκτίμησης φύλου που προκύπτουν από τις συγκεκριμένες μεθόδους είναι εγγενώς τείνουν να εμφανίζουν πληθυσμιακή εξειδίκευση (Bertsatos, Athanasoroulou, et al., 2019; Noble et al., 2019).

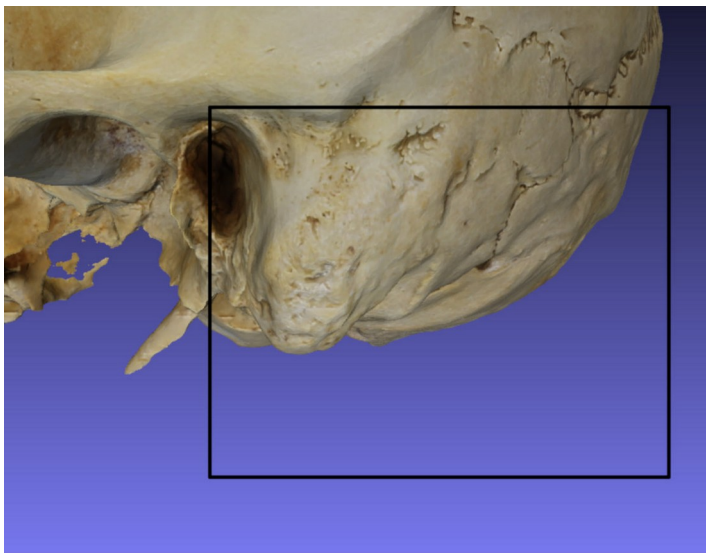
Ο συνδυασμός των βέλτιστων χαρακτηριστικών των προαναφερθέντων τεχνικών προσβλέπει στην ανάπτυξη νέων μεθόδων εκτίμησης φύλου που θα έχουν την ακρίβεια και την αναπαραγωγικότητα των μορφομετρικών μεθόδων, διατηρώντας παράλληλα την ευελιξία ενός ιδιαίτερα εκπαιδευμένου παρατηρητή για την οπτική αξιολόγηση των ανεπαίσθητων φυλετικά διμορφικών μορφολογικών διακυμάνσεων. Για την ακρίβεια, κατά την τελευταία δεκαετία, αρκετές μελέτες έχουν προτείνει προσαρμοσμένες μορφομετρικές μεταβλητές για συγκεκριμένα φυλετικά διμορφικά χαρακτηριστικά, όπως η μαστοειδής απόφυση (Kanchan et al., 2013; Petaros et al., 2015) και το υπερόφρυο τόξο (Shearer et al., 2012). Μετά τις πρόσφατες καινοτομίες όσον αφορά στην εκτίμηση φύλου από το κρανίο και λαμβάνοντας υπόψη την αυξανόμενη διαθεσιμότητα τριδιάστατων σκελετικών δειγμάτων καθώς και την εξέλιξη των τεχνολογιών τριδιάστατης

σάρωσης, η παρούσα μελέτη προτείνει μια νέα μέθοδο για την εκτίμηση του φύλου από το κρανίο, η οποία βασίζεται σε πλήρως αυτοματοποιημένη και ακριβή εξαγωγή μορφομετρικών χαρακτηριστών από τριδιάστατα μοντέλα. Αυτό επιτυγχάνεται με ελάχιστες παρεμβάσεις από τον χρήστη και ταυτόχρονα, ενσωματωμένες ρουτίνες βελτιστοποίησης εξασφαλίζουν ότι το σφάλμα μέτρησης καθορίζεται ουσιαστικά μόνο από την ακρίβεια των τριδιάστατων μοντέλων. Η προτεινόμενη μέθοδος εισάγει επίσης την έννοια των εικόνων χαρτών ύψους (HMI) ως μορφομετρικά χαρακτηριστικά, οι οποίες μπορούν να θεωρηθούν ως ο αντίστροφος μετασχηματισμός της τριδιάστατης κρανιακής επιφάνειας συγκεκριμένων φυλετικά διμορφικών περιοχών σε διδιάστατες εικόνες κλίμακας του γκρι. Τα HMIs υποβάλλονται σε επεξεργασία μέσω αλγορίθμων μηχανικής όρασης και μάθησης. Επιπλέον, η παρούσα μελέτη αξιοποιεί διεξοδικά δύο δείγματα ευρωπαϊκών πληθυσμών για την αποφυγή της υπερπροσαρμογής (over-fitting) των στατιστικών μοντέλων και της πληθυσμιακής εξειδίκευσης προκειμένου να εξασφαλιστεί ότι η δηλωθείς ακρίβεια εκτίμησης φύλου αντικατοπτρίζει την πραγματική ισχύ της προτεινόμενης μεθόδου σε άγνωστα δείγματα.

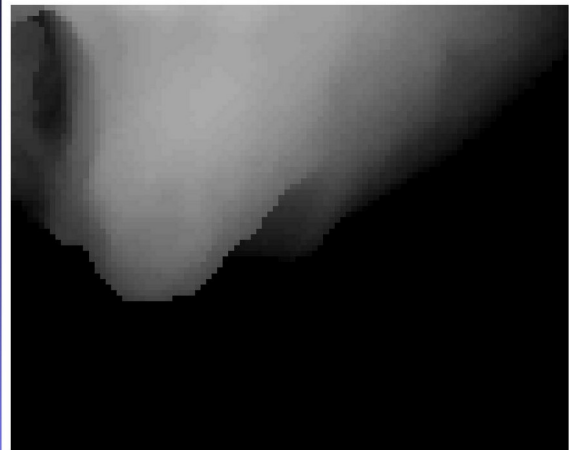
12.1. Υλικά και μέθοδοι

Το δείγμα που χρησιμοποιήθηκε στην παρούσα μελέτη αποτελείται από 324 τριδιάστατα μοντέλα κρανίων που αντιπροσωπεύουν δύο σύγχρονους ευρωπαϊκούς πληθυσμούς. Το ελληνικό δείγμα αποτελείται από 154 (79 αρσενικά και 75 θηλυκά) κρανία από την σύγχρονη σκελετική συλλογή αναφοράς. Το τσέχικο δείγμα αποτελείται από 170 μοντέλα (102 άνδρες και 68 γυναίκες, ηλικίας 21 έως 87 ετών), τα οποία ανακατασκευάστηκαν από ανώνυμες αξονικές τομογραφίες (αρχεία DICOM) από το Νοσοκομείο Na Homolce στην Πράγα.

Και τα 324 μοντέλα αναλύθηκαν με το λογισμικό "skullanalyzer" (Bertsatos, 2019). Ένα πρόγραμμα Η/Υ ανοικτού κώδικα που εξάγει αυτόματα μορφομετρικά χαρακτηριστικά κρανιακών περιοχών με έντονο φυλετικό διμορφισμό από τριδιάστατα μοντέλα και διατίθεται δωρεάν στη διεύθυνση <https://github.com/pr0m1th3as/skullanalyzer>. Το skullanalyzer εξάγει συγκεκριμένα HMIs από το ινιακό έπαρμα, το υπερόφρυο τόξο και τις μαστοειδείς αποφύσεις, καθώς και τη διδιάστατη προβολή του εξωκράνιου τμήματος nasion-bregma σε μορφή polyline για το οποίο υπολογίζονται αυτόματα οι αντίστοιχοι ελλειπτικοί περιγραφείς Fourier (EFDs). Η εικόνα 12.1 απεικονίζει την αντιστοιχία μεταξύ του HMI της πλευρικής όψης της αριστερής μαστοειδούς απόφυσης και της αντίστοιχης τριδιάστατης ψηφιακής αναπαράστασης του κρανίου. Η ευθυγράμμιση του τριδιάστατου μοντέλου πραγματοποιήθηκε χειροκίνητα στο Meshlab.



Left mastoid process lateral-view HMI



Εικόνα 12.1: Παράδειγμα αντιστοίχισης του HMI με την τριδιάστατη ψηφιακή αναπαράσταση της αριστερής μαστοειδούς απόφυσης

Ο σχεδιασμός της ανάλυσης δεδομένων στοχεύει στην αξιολόγηση των προαναφερθέντων φυλετικά διμορφικών κρανιακών χαρακτηριστικών υπό το πρίσμα της πληθυσμιακής γενίκευσης έναντι της πληθυσμιακής εξειδίκευσης. Όλη η ανάλυση, η οποία περιλαμβάνει χειρισμό δειγμάτων, ανάλυση εικόνας, εποπτευόμενη μηχανική μάθηση, υποτύπωση αποτελεσμάτων και στατιστική μοντελοποίηση για την εκτίμηση του φύλου, πραγματοποιήθηκε με προσαρμοσμένες ρουτίνες στην Python και στην GNU Octave. Ολόκληρη η βάση κώδικα ανάλυσης δεδομένων που εφαρμόζεται σε αυτήν την μελέτη είναι διαθέσιμη στη διεύθυνση <https://github.com/pr0m1th3as/cranial-sex-estimation-code>.

12.1.1. Ανάλυση δεδομένων - διαχωρισμός δειγμάτων

Καθένα από τα δύο πληθυσμιακά δείγματα χωρίστηκε σε δύο τμήματα, ένα κύριο και ένα μικρότερο. Αυτός ο διαχωρισμός υιοθετήθηκε για τη δημιουργία ενός μεγάλου ομαδοποιημένου δείγματος, το οποίο απαρτίζεται από ολόκληρο το δείγμα του ενός πληθυσμού και το μικρότερο τμήμα του δεύτερου πληθυσμού. Από το ομαδοποιημένο αυτό δείγμα με τυχαία μετάθεση προκύπτουν τα δεδομένα εκπαίδευσης της μηχανικής εκμάθησης (machine learning) και το δείγμα ελέγχου των στατιστικών μοντέλων εκτίμησης φύλου. Στο σύνολο, πραγματοποιήθηκαν 100 διαδοχικές επαναλήψεις τυχαίων μεταθέσεων. Το κύριο τμήμα του δεύτερου πληθυσμού, το οποίο παραμένει το ίδιο και στις 100 διαδοχικές επαναλήψεις, χρησιμοποιείται για επιπλέον διασταυρούμενη επικύρωση της επιτυχούς ταξινόμησης του φύλου για κάθε στατιστικό μοντέλο. Αυτή η επαναλαμβανόμενη προσέγγιση πραγματοποιήθηκε δύο φορές, μία με τον ελληνικό πληθυσμό να αποτελεί σχεδόν το σύνολο του ομαδοποιημένου δείγματος και το κύριο τμήμα του

τσέχικου πληθυσμού να χρησιμοποιείται για επιπλέον διασταυρούμενη επικύρωση και μια δεύτερη φορά με τα δύο πληθυσμιακά δείγματα να έχουν αντιστραφεί.

12.1.2. Ανάλυση δεδομένων - μηχανική μάθηση

Η αξιολόγηση κάθε μορφομετρικού χαρακτήρα ως προς την εκτίμηση φύλου πραγματοποιήθηκε με εποπτευόμενη μηχανική εκμάθηση, η οποία αξιοποίησε τα δεδομένα εκπαίδευσης, τα δείγματα ελέγχου και τα δείγματα διασταυρούμενης επικύρωσης. Τα έξι HMI, που αντιστοιχούν σε διαφορετικούς μορφομετρικούς χαρακτήρες, αρχικά υποβλήθηκαν σε επεξεργασία με τον αλγόριθμο ιστόγραμμα προσανατολισμένων κλίσεων (HOGs). Στη συνέχεια, τα HOGs χρησιμοποιήθηκαν ως τα βασικά δεδομένα μηχανικής εκμάθησης για μηχανές διανυσμάτων υποστήριξης (SVM) εφαρμόζοντας ακτινικές συναρτήσεις βάσης (RBF), γνωστές και ως πυρήνας, με τις εξής παραμέτρους: $C = 20$, $\gamma = 1 / \text{αριθμός στοιχείων HOG}$. Οι τελικές παράμετροι τόσο για τον πυρήνα RBF όσο και για τα HOGs βελτιστοποιήθηκαν μέσω αλληπάληλων δοκιμών. Οι ελλειπτικοί περιγραφείς Fourier (EFDs) της διδιάστατης προβολής του εξωκράνιου τμήματος nasion-bregma αναλύθηκαν αρχικά με ανάλυση γραμμικής διακρίνουσας (LDA) και μόνο οι επτά πιο διακριτικοί συντελεστές ($a_1, a_2, a_3, a_4, a_7, c_2, c_3$) χρησιμοποιήθηκαν στον τελικό ταξινομητή γραμμικής διακριτικής ανάλυσης (LDA Classifier).

12.1.3. Ανάλυση δεδομένων – υποτύπωση αποτελεσμάτων

Κάθε ένας από τους επτά ταξινομητές αξιολογήθηκε ανεξάρτητα τόσο βάσει του δείγματος ελέγχου των στατιστικών μοντέλων εκτίμησης φύλου όσο και βάσει του δείγματος διασταυρούμενης επικύρωσης. Η αξιολόγηση πραγματοποιήθηκε συγκρίνοντας το αποτέλεσμα της κάθε συνάρτησης του αντίστοιχου ταξινομητή με το φύλο του εκάστοτε δείγματος. Δεδομένου ότι οι ταξινομητές SVM που εφαρμόζουν πυρήνα RBF δεν παρέχουν άμεσες εκτιμήσεις των οπίσθιων πιθανοτήτων, η επιτυχής ταξινόμηση αξιολογήθηκε σύμφωνα με το καθορισμένο κριτήριο φραγμού (ορίζεται στο μηδέν) κάθε ταξινομητή καθώς και βάσει διαφορετικών ευρών κατωφλίου (thresholds) αυξανόμενων κατά $\pm 0,1$ έως το $\pm 1,4$. Σε κάθε περίπτωση, τα δείγματα τα αποτελέσματα των οποίων βρίσκονταν εκτός του αντίστοιχου εύρους κατωφλίου θεωρούνταν έγκυρα, ενώ τα δείγματα τα αποτελέσματα των οποίων βρίσκονταν εντός του εύρους κατωφλίου θεωρούνταν απροσδιόριστα και απορρίπτονταν. Η ορθή ταξινόμηση υπολογίστηκε ως η ποσοστιαία αναλογία των επιτυχώς ταξινομημένων δειγμάτων προς των έγκυρων δειγμάτων.

12.1.4. Ανάλυση δεδομένων – συνδυασμός μοντέλων

Μετά την αξιολόγηση της απόδοσης κάθε ταξινομητή με την προαναφερθείσα επαναλαμβανόμενη διαδικασία, οι παράμετροι των καλύτερων ταξινομητών κάθε μορφομετρικού χαρακτηριστικού χρησιμοποιήθηκαν περαιτέρω συνδυαστικά για την εκτίμηση του φύλου. Όλα τα διαθέσιμα μορφομετρικά χαρακτηριστικά από κάθε κρανιακό δείγμα (324 άτομα) αξιολογήθηκαν με τους αντίστοιχους ταξινομητές και στη συνέχεια κάθε δείγμα ταξινομήθηκε σύμφωνα με το σταθμισμένο άθροισμα των αποτελεσμάτων των ταξινομητών αυτών. Και σε αυτήν την περίπτωση, η επιτυχής ταξινόμηση του κάθε δείγματος αξιολογήθηκε σύμφωνα με το καθορισμένο κριτήριο φραγμού καθώς και βάσει διαφορετικών ευρών κατωφλίου αυξανόμενων κατά $\pm 0,1$ έως το $\pm 1,4$. Σε πρώτο στάδιο, η προαναφερθείσα διαδικασία καθώς και η καταγραφή των αποτελεσμάτων πραγματοποιήθηκε ξεχωριστά για τα δύο διαχωρισμένα δείγματα. Ωστόσο, στη συνέχεια, και οι 14 ταξινομητές (2 για κάθε χαρακτηριστικό, 7 από κάθε διαχωρισμένο δείγμα) εφαρμόστηκαν συλλογικά σε ολόκληρο το δείγμα και έγινε καταγραφή των αποτελεσμάτων για κάθε υποομάδα των δύο πληθυσμών.

12.2. Αποτελέσματα

Τα αποτελέσματα ταξινόμησης του πρώτου και του δεύτερου διαχωρισμένου δείγματος συνοψίζονται επιλεκτικά στους Πίνακες 12.1 και 12.2 αντίστοιχα. Τα αποτελέσματα αυτά αφορούν τα ποσοστά επιτυχούς ταξινόμησης και απόρριψης της διάμεσης και μέγιστης τιμής των 100 επαναλήψεων στο καθορισμένο κριτήριο φραγμού (ορίζεται στο μηδέν) καθώς και εύρος κατωφλίου $\pm 0,2$.

Πίνακας 12.1: Αποτελέσματα ταξινόμησης και διασταυρούμενη επικύρωση όσον αφορά στο διαχωρισμένο δείγμα # 1.

Ομάδες αξιολόγησης	Ομαδοποιημένο δείγμα (ABH, WLH, MOTOL)				Διασταυρούμενη επικύρωση (HOMOLKA)			
	Διάμεσος		Μέγιστη τιμή		Διάμεσος		Μέγιστη τιμή	
Κατώφλι ^β	Μηδέν ^γ	± 0,2 ^δ	Μηδέν ^γ	± 0,2 ^δ	Μηδέν ^γ	± 0,2 ^δ	Μηδέν ^γ	± 0,2 ^δ
Ινιακό έπαρμα ΗΜΙ	71,7	75,4 / 20,7	83,3	82,5 / 14,2	67,0	70,5 / 14,6	73,8	76,5 / 17,5
Υπερόφρυο τόξο ΗΜΙ	76,5	80,0 / 13,9	86,3	89,5 / 14,4	80,6	86,5 / 15,2	85,4	89,9 / 16,5
Αριστερή μαστοειδής απόφυση – κάτω όψη ΗΜΙ	66,2	68,4 / 19,7	78,5	82,2 / 30,8	70,6	74,6 / 20,0	77,5	78,6 / 19,6
Αριστερή μαστοειδής απόφυση – πλευρική όψη ΗΜΙ	80,0	84,3 / 12,7	90,8	91,4 / 10,8	79,4	84,1 / 15,8	84,3	86,7 / 18,6
Δεξιά μαστοειδής απόφυση – κάτω όψη ΗΜΙ	67,2	72,2 / 19,5	79,7	85,4 / 25,0	70,6	74,3 / 14,6	77,5	80,7 / 13,7
Δεξιά μαστοειδής απόφυση – πλευρική όψη ΗΜΙ	82,8	86,6 / 12,7	92,2	93,3 / 7,0	78,4	80,6 / 11,2	82,4	82,2 / 9,3
<i>nasion-bregma</i> EFDs	85,2	89,8 / 18,6	96,7	98,1 / 14,8	82,5	92,7 / 23,4	85,4	92,9 / 26,9

^α Η στατιστική ταξινόμηση αφορά τη διάμεσης και μέγιστης τιμής των 100 επαναλήψεων.

^β Το κατώφλι για την αξιολόγηση των αποτελεσμάτων κάθε ταξινομητή.

^γ Καθορισμένο κριτήριο φραγμού (ορίζεται στο μηδέν). Τα αποτελέσματα, τα οποία παρουσιάζονται σε ποσοστά αφορούν το σύνολο του δείγματος.

^δ Στο εύρος κατωφλίου τόσο τα αποτελέσματα επιτυχούς ταξινόμησης όσο και απόρριψης δίνονται ως ποσοστά (επιτυχής ταξινόμηση / απορριπτέα). Το ποσοστό απόρριψης αφορά το τμήμα του δείγματος, το οποίο βάσει του εύρους κατωφλίου κρίνεται απροσδιόριστο.

Πίνακας 12.2: Αποτελέσματα ταξινόμησης και διασταυρούμενη επικύρωση όσον αφορά στο διαχωρισμένο δείγμα # 2.

Ομάδες αξιολόγησης	Ομαδοποιημένο δείγμα (HOMOLKA, WLH, MOTOL)				Διασταυρούμενη επικύρωση (ABH)			
	Διάμεσος		Μέγιστη τιμή		Διάμεσος		Μέγιστη τιμή	
Κατώφλι ^β	Μηδέν ^γ	± 0,2 ^δ	Μηδέν ^γ	± 0,2 ^δ	Μηδέν ^γ	± 0,2 ^δ	Μηδέν ^γ	± 0,2 ^δ
Ινιακό έπαρμα ΗΜΙ	66,1	68,4 / 19,4	80,6	80,0 / 11,3	72,3	78,6 / 24,0	80,2	84,0 / 25,7
Υπερόφρουο τόξο ΗΜΙ	78,6	82,1 / 11,2	89,3	90,8 / 9,5	74,5	78,1 / 15,5	80,6	85,5 / 15,3
Αριστερή μαστοειδής απόφυση – κάτω όψη ΗΜΙ	67,6	69,9 / 17,7	83,8	84,9 / 22,1	62,6	65,7 / 21,2	71,7	72,0 / 27,8
Αριστερή μαστοειδής απόφυση – πλευρική όψη ΗΜΙ	80,9	83,7 / 14,4	89,7	91,1 / 17,6	79,8	84,2 / 16,8	86,9	87,5 / 11,1
Δεξιά μαστοειδής απόφυση – κάτω όψη ΗΜΙ	69,1	73,6 / 20,0	79,4	84,0 / 23,5	70,4	73,6 / 19,9	80,6	79,7 / 19,4
Δεξιά μαστοειδής απόφυση – πλευρική όψη ΗΜΙ	76,5	80,6 / 15,8	85,3	87,7 / 12,3	86,7	91,2 / 12,9	90,8	92,7 / 11,9
<i>nasion-bregma</i> EFDs	84,1	90,9 / 22,2	93,7	96,2 / 17,5	83,2	86,8 / 16,1	86,1	88,2 / 15,8

^α Η στατιστική ταξινόμηση αφορά τη διάμεσης και μέγιστης τιμές των 100 επαναλήψεων.

^β Το κατώφλι για την αξιολόγηση των αποτελεσμάτων κάθε ταξινομητή.

^γ Καθορισμένο κριτήριο φραγμού (ορίζεται στο μηδέν). Τα αποτελέσματα, τα οποία παρουσιάζονται σε ποσοστά αφορούν το σύνολο του δείγματος.

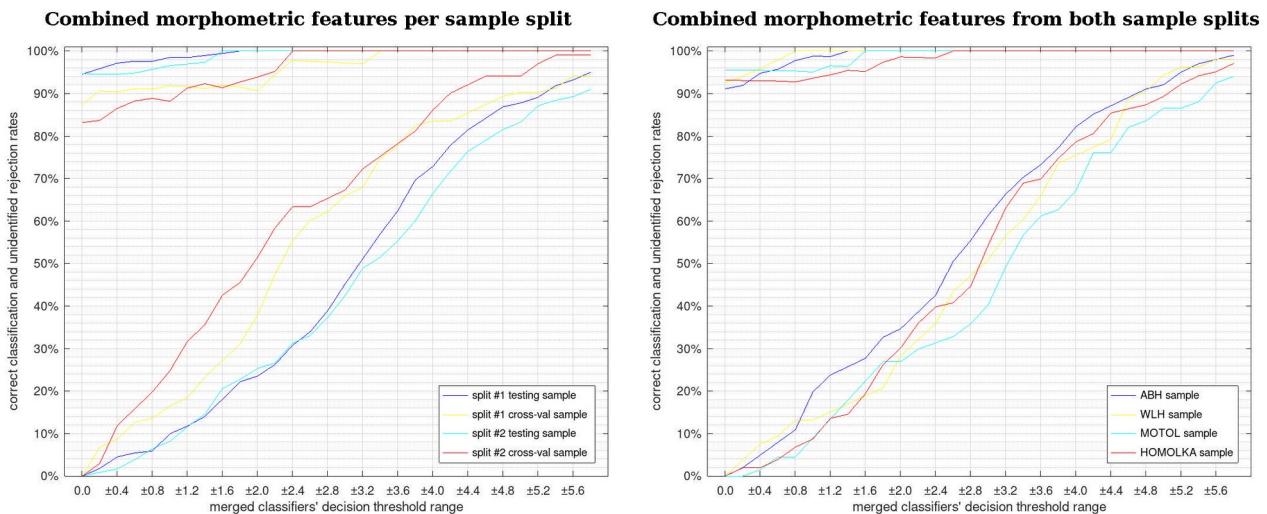
^δ Στο εύρος κατωφλίου τόσο τα αποτελέσματα επιτυχούς ταξινόμησης όσο και απόρριψης δίνονται ως ποσοστά (επιτυχής ταξινόμηση / απορριπτέα). Το ποσοστό απόρριψης αφορά το τμήμα του δείγματος, το οποίο βάσει του εύρους κατωφλίου κρίνεται απροσδιόριστο.

Οι μορφομετρικοί χαρακτήρες με περιορισμένη διακριτική ικανότητα ως προς το φύλο είναι τα ΗΜΙs της κάτω όψης της μαστοειδής απόφυσης, ακολουθεί το ινιακό έπαρμα και τα ΗΜΙs του υπερόφρουου τόξου. Αντιθέτως, τα EFDs τη διδιάστατης προβολής του εξωκράνιου τμήματος *nasion-bregma* εμφανίζουν την βέλτιστη διακριτική ικανότητα ως προς το φύλο και ακολουθούν τα ΗΜΙs της πλευρικής όψης των μαστοειδών αποφύσεων.

Εκτός από τις αξιοσημείωτες διαφορές στα αποτελέσματα ταξινόμησης μεταξύ των αξιολογούμενων μορφομετρικών χαρακτήρων, οι συγκρίσεις ανά ζεύγη του δείγματος ελέγχου και του δείγματος που χρησιμοποιείται για διασταυρούμενη επικύρωση αποκαλύπτουν επίσης πληθυσμιακές διαφορές στην ακρίβεια ταξινόμησης. Όπως ήταν αναμενόμενο, η επίδοση της ταξινόμησης και για τα δύο δείγματα που χρησιμοποιούνται για διασταυρούμενη επικύρωση είναι σε γενικές γραμμές χαμηλότερη σε σύγκριση με την αντίστοιχη επίδοση που επιδεικνύουν τα δειγμάτων ελέγχου, δεδομένου ότι τα πρώτα δεν αντιπροσωπεύονται ικανοποιητικά στο δείγμα που απαρτίζει τα δεδομένα εκπαίδευσης της μηχανικής εκμάθησης των στατιστικών μοντέλων εκτίμησης φύλου. Ωστόσο, υπάρχουν ορισμένες εξαιρέσεις σε αυτήν τη γενική τάση. Η πιο αξιοσημείωτη εξαίρεση αφορά το HMI της πλευρικής όψης της δεξιάς μαστοειδούς απόφυσης, όπου στο διαχωρισμένο δείγμα # 1, η επίδοση της ταξινόμησης του δείγματος που χρησιμοποιείται για διασταυρούμενη επικύρωση (HOMOLKA) είναι χαμηλότερη συγκρινόμενη με αυτήν του δείγματος ελέγχου (μέση τιμή 78,4% - μέγιστο 82,4% και 82,8% - 92,2%, αντίστοιχα), ενώ στο διαχωρισμένο δείγμα # 2 παρατηρείται το ακριβώς αντίθετο. Η υποομάδα ABH του ελληνικού πληθυσμού που χρησιμοποιήθηκε για διασταυρούμενη επικύρωση εμφάνισε πολύ υψηλότερη επίδοση ταξινόμησης (διάμεσος 86,7% - μέγιστο 90,9%) σε αντίθεση με το αντίστοιχο δείγμα ελέγχου (76,5% - 85,3%).

Τα αποτελέσματα ταξινόμησης του φύλου βάσει πολλαπλών μορφομετρικών χαρακτήρων παρουσιάζονται στα δύο διαγράμματα της εικόνας 12.2. Το αριστερό διάγραμμα δείχνει τα ποσοστά επιτυχούς ταξινόμησης και τα ποσοστά απροσδιόριστων ατόμων όσον αφορά στην εφαρμογή των 7 καλύτερων ταξινομητών κάθε μορφομετρικού χαρακτήρα στα δύο διαχωρισμένα δείγματα ξεχωριστά. Το διάγραμμα στη δεξιά πλευρά της εικόνας 12.2 δείχνει τα αντίστοιχα ποσοστά που αφορούν στην εφαρμογή των 14 ταξινομητών στην κάθε υποομάδα των δύο πληθυσμών. Οι γραμμές στο πάνω μέρος κάθε διαγράμματος αντιστοιχούν στα ποσοστά επιτυχούς ταξινόμησης και τα ίδια χρώματα αντίστοιχα στο κάτω μέρος δείχνουν τα ποσοστά απροσδιόριστων ατόμων σύμφωνα με τα διαφορετικά εύρη κατωφλίου, τα οποία ορίζονται στον x άξονα. Παρατηρούμε ότι η ακρίβεια της ταξινόμησης είναι πολύ υψηλή όσον αφορά στην εφαρμογή των 7 καλύτερων ταξινομητών κάθε μορφομετρικού χαρακτηριστικού στα δύο διαχωρισμένα δείγματα ξεχωριστά. Πιο συγκεκριμένα, η ακρίβεια είναι πάνω από 94% και για τα δύο δείγματα ελέγχου στο καθορισμένο κριτήριο φραγμού (ορίζεται στο μηδέν) και αγγίζει τον πλήρη διαχωρισμό των φύλων μεταξύ των έγκυρων δειγμάτων λίγο πριν τα ποσοστά απροσδιόριστων ατόμων φτάσουν το 20%. Τα ποσοστά επιτυχούς ταξινόμησης όσον αφορά στα δείγματα που χρησιμοποιούνται για διασταυρούμενη επικύρωση είναι χαμηλότερα (περίπου 87% και 83% για το διαχωρισμένο δείγμα # 1 και # 2 αντίστοιχα) και επίσης αυξάνονται με χαμηλότερο ρυθμό σε σχέση με την αύξηση των

ποσοστών απροσδιόριστων ατόμων. Η εφαρμογή των 14 ταξινομητών στην κάθε υποομάδα των δύο πληθυσμών βελτιώνει περαιτέρω την επίδοση ταξινόμησης, η οποία υπερβαίνει το 91% στο καθορισμένο κριτήριο φραγμού για όλες τις υποομάδες των δύο πληθυσμών και αυξάνεται ταχύτερα σε σχέση με τα ποσοστά απροσδιόριστων ατόμων.



Εικόνα 12.2: Επίδοση ταξινόμησης βάσει πολλαπλών χαρακτηριστικών.

Οι βέλτιστοι συνδυασμοί των στατιστικών μοντέλων έχουν ενσωματωθεί σε μια ξεχωριστή συνάρτηση GNU Octave, που ονομάζεται “estimate_sex.m”, η οποία μπορεί να χρησιμοποιηθεί για την εκτίμηση του φύλου από τα μορφομετρικά χαρακτηριστικά που είναι αποθηκευμένα στο σχετικό αρχείο .mat που δημιουργήθηκε με το skullanalyzer. Η συνάρτηση, είναι επίσης διαθέσιμη στη διεύθυνση <https://github.com/pr0m1th3as/cranial-sex-estimation-code>.

12.3. Συμπεράσματα

Ο σκοπός αυτής της μελέτης ήταν να δημιουργήσει μια αποτελεσματική αλλά και αξιόπιστη μέθοδο για την εκτίμηση του φύλου από κρανία αγνώστων στοιχείων. Ο στόχος για αποδοτικότητα επιτεύχθηκε μέσω της υλοποίησης λογισμικού, δηλαδή του προγράμματος H/Y **skullanalyzer** και της συνάρτησης του GNU Octave **estimate_sex.m**, τα οποία συνδυαστικά μπορούν να αναλύσουν ένα κρανίο και γίνει εκτίμηση του φύλου σε λίγα λεπτά. Σε γενικές γραμμές, έχοντας στη διάθεση μας το τριδιάστατο μοντέλο ενός κρανίου, ολόκληρη η διαδικασία εκτίμησης φύλου, συμπεριλαμβανομένης του αρχικού εντοπισμού αναφορικών σημείων στο Meshlab, μπορεί να πραγματοποιηθεί σε οποιονδήποτε σύγχρονο υπολογιστή μέσα σε λίγα λεπτά. Φυσικά αυτό προϋποθέτει τη διαθεσιμότητα ενός τριδιάστατου μοντέλου, αλλά με τις ραγδαίες εξελίξεις στις τριδιάστατες τεχνολογίες και ιδιαίτερα τους σαρωτές δομημένου φωτός, αυτή είναι μια αρκετά προσιτή διαδικασία, η οποία αναμένεται να γίνει ολοένα και φθηνότερη και πιο αποτελεσματική. Η

αξιοπιστία, η οποία είναι ίσως ακόμη πιο σημαντική, έχει επιτευχθεί σε μεγάλο βαθμό με δύο τρόπους. Αφενός, η ενσωμάτωση αλγορίθμων βελτιστοποίησης των αναφορικών σημείων για την ελαχιστοποίηση του σφάλματος του παρατηρητή καθιστά τα μορφομετρικά χαρακτηριστικά, και κατά συνέπεια και την εκτίμηση φύλου, εξαιρετικά αναπαραξίμη. Από την άλλη πλευρά, ο συγκεκριμένος σχεδιασμός στατιστικής ανάλυσης όχι μόνο διασφαλίζει ότι η αναφερόμενη ακρίβεια δεν υπερεκτιμάται από τους περιορισμούς των δειγμάτων, αλλά παρέχει επίσης την οπίσθια πιθανότητα της εκτίμησης του φύλου που την καθιστά ακόμη πιο αξιόπιστη στις περιπτώσεις δειγμάτων αγνώστων στοιχείων.

Τα αποτελέσματα της παρούσας εργασίας όσον αφορά στην ακρίβεια ταξινόμησης της προτεινόμενης μεθόδου, παρόλο που δεν είναι τα βέλτιστα λόγω της στρατηγικής δειγματοληψίας και των σχετικών περιορισμών που έχουν προαναφερθεί, εξακολουθούν να βρίσκονται στα ίδια επίπεδα με πρόσφατες εργασίες που περιλαμβάνουν άλλες καινοτόμες προσεγγίσεις (Nikita, 2019; Yang et al., 2019). Ωστόσο, αξίζει να σημειωθεί πως οι εργασίες αυτές αξιολογήθηκαν σε δείγματα μεμονωμένων πληθυσμών, γεγονός που μπορεί να ενισχύει τα αποτελέσματα ταξινόμησης. Η προτεινόμενη μέθοδος φέρει μια σειρά από καινοτομίες στην εφαρμογή της, με ίσως την πιο σημαντική τα HMIs, που επιτρέπουν την αξιολόγηση της τριδιάστατης κρανιακής γεωμετρίας από αλγόριθμους μηχανικής όρασης που εστιάζουν σε 2D εικόνες. Απαιτείται περαιτέρω μελέτη με μεγαλύτερα δείγματα και περισσότερους πληθυσμούς για τη βελτίωση της απόδοσης της προτεινόμενης μεθόδου. Για το σκοπό αυτό, το σύνολο του κώδικα προγραμματισμού (σε GNU Octave, python, C++) της παρούσας εργασίας, συμπεριλαμβανομένου του λογισμικού skullanalyzer και του συνόλου δεδομένων μορφομετρικών χαρακτηριστικών, έχει διατεθεί στο κοινό με κατάλληλη άδεια ανοικτού κώδικα (GNU GPLv3) για τη διευκόλυνση της μελλοντικής έρευνας και της δημιουργικής συνεργασίας.

Απολογισμός έργου

Συνοψίζοντας το ερευνητικό έργο των τελευταίων 50 μηνών, που διήρκησε η εκπόνηση της παρούσας διδακτορικής διατριβής, τα ερευνητικά αποτελέσματα και συμπεράσματα καθώς και οι μεθοδολογικές καινοτομίες που παρουσιάστηκαν στα προηγούμενα κεφάλαια ανέδειξαν και τον όγκο της δουλειάς που ολοκληρώθηκε αλλά και την σημασία της στον εμπλουτισμό της επιστημονικής γνώσης. Το εύρος των διαφορετικών ερωτημάτων που απαντήθηκαν προέκυψαν σε μεγάλο βαθμό από ανθρωπολογική περιέργεια διερευνώντας τα όρια της οστεολογικής μελέτης και τι δύναται να μας πληροφορήσει αλλά και κατά κάποιο τρόπο κατ' ανάγκη καθώς η χρονοβόρα υλοποίηση της ψηφιακής τεκμηρίωσης ολόκληρης της σύγχρονης σκελετικής συλλογής έθεσε το χρονικό πλαίσιο και τους ρυθμούς εξέλιξης του ερευνητικού έργου. Με την περάτωση της ψηφιοποίησης δόθηκε και η δυνατότητα για την ανάπτυξη και υλοποίηση νέων μεθοδολογιών, οι οποίες ανοίγουν νέους ορίζοντες στην μορφομετρική σκελετική ανάλυση συνδυάζοντας πλήθος εργαλείων από την υπολογιστική γεωμετρία έως την μηχανική εκμάθηση και την υπολογιστική όραση.

Κομβικό σημείο της παρούσας διατριβής είναι φυσικά η σύγχρονη σκελετική συλλογή αναφοράς, η οποία είναι μία από τις πληρέστερες και καλύτερα διατηρημένες σε ολόκληρη την Ευρώπη. Το ερευνητικό έργο των τελευταίων ετών αλλά και η ολοκληρωμένη ψηφιακή τεκμηρίωσή της έχουν θέσει γερά θεμέλια τόσο ως προς της βέλτιστη διατήρηση της συλλογής αλλά και την περαιτέρω ανάδειξή της σε διεθνές επίπεδο ως ένα εξαιρετικό ερευνητικό εργαλείο. Μένοντας στη συλλογή πρέπει να επισημανθεί ότι στο σύνολο της παρούσας διατριβής αλλά και στις πρωτότυπες δημοσιεύσεις, που παρατίθενται στο τέλος, γίνεται συστηματική χρήση των “φυλάσσεται” είτε “διατηρείται” είτε “συντηρείται” έναντι του “ανήκει” σε σχέση με τον Τομέα Φυσιολογίας Ζώων και Ανθρώπου, ο οποίος φέρει την ευθύνη ρύθμισης όλων των θεμάτων που αφορούν τη συλλογή. Βασιζόμενος στις αρχές της ακαδημαϊκής και επιστημονικής βιοηθικής και ανεξαρτήτως της πραγματικής έλλειψης κανονιστικού πλαισίου για τα σκελετικά κατάλοιπα, αυτή είναι μία συνειδητή επιλογή, η οποία αρνείται να αναγνωρίσει ή να προσδώσει ιδιότητες εμπράγματος δικαιώματος στα ανθρώπινα σκελετικά κατάλοιπα είτε ως ιδιοκτησία είτε ως εμπορικό προϊόν. Εν κατακλείδι, είναι σαφές ότι αυτή η επιλογή φρασεολογίας δεν αφήνει κανένα περιθώριο προς οποι@δήποτε να επικαλεστεί ιδιοκτησία οποιασδήποτε μορφής πάνω στη σύγχρονη σκελετική συλλογή αναφοράς που φυλάσσεται και διατηρείται στον Τομέα Φυσιολογίας Ζώων και Ανθρώπου του Τμήματος Βιολογίας του ΕΚΠΑ.

A.M.

Βιβλιογραφία

- Acar, N., Karaarslan, A., Karakasli, A., & Erduran, M. (2017). Femoral Head Fovea Capitis Variant Configurations and Age Related Changes- A Radiological Study. *Iranian Journal of Radiology, InPress(InPress)*. <https://doi.org/10.5812/iranjradiol.41130>
- Ammer, S., d'Oliveira Coelho, J., & Cunha, E. M. (2019). Outline Shape Analysis on the Trochlear Constriction and Olecranon Fossa of the Humerus: Insights for Sex Estimation and a New Computational Tool. *Journal of Forensic Sciences, 64(6)*, 1788–1795. <https://doi.org/10.1111/1556-4029.14096>
- Anzellini, A., & Toyne, J. M. (2019). Estimating sex using isolated appendicular skeletal elements from Chachapoyas, Peru. *International Journal of Osteoarchaeology, 29(6)*, 961–973. <https://doi.org/10.1002/oa.2810>
- Arigbabu, O. A., Liao, I. Y., Abdullah, N., & Mohamad Noor, M. H. (2017). Computer vision methods for cranial sex estimation. *IPSS Transactions on Computer Vision and Applications, 9(1)*, 19. <https://doi.org/10.1186/s41074-017-0031-6>
- Barbeito-Andrés, J., Anzelmo, M., Ventrice, F., & Sardi, M. L. (2012). Measurement error of 3D cranial landmarks of an ontogenetic sample using Computed Tomography. *Journal of Oral Biology and Craniofacial Research, 2(2)*, 77–82. <https://doi.org/10.1016/j.jobcr.2012.05.005>
- Becker, I., Woodley, S. J., & Stringer, M. D. (2010). The adult human pubic symphysis: A systematic review: The pubic symphysis. *Journal of Anatomy, 217(5)*, 475–487. <https://doi.org/10.1111/j.1469-7580.2010.01300.x>
- Bejdová, Š., Dupej, J., Krajíček, V., Velemínská, J., & Velemínský, P. (2018). Stability of upper face sexual dimorphism in central European populations (Czech Republic) during the modern age. *International Journal of Legal Medicine, 132(1)*, 321–330. <https://doi.org/10.1007/s00414-017-1625-3>
- Beltran, L. S., Mayo, J. D., Rosenberg, Z. S., De Tuesta, M. D., Martin, O., Neto, L. P., & Bencardino, J. T. (2012). Fovea Alta on MR Images: Is It a Marker of Hip Dysplasia in Young Adults? *American Journal of Roentgenology, 199(4)*, 879–883. <https://doi.org/10.2214/AJR.11.8193>
- Bertsatos, A. (2018). *long-bone-diaphyseal-CSG-Toolkit: A GNU Octave CSG Toolkit (v1.0.1)* [Computer software]. Zenodo. <https://doi.org/10.5281/zenodo.1467045>

- Bertsatos, A. (2019). *skullanalyzer: A concrete way of extracting cranial morphometric features* (v1.0) [Computer software]. Zenodo. <https://doi.org/10.5281/zenodo.3594565>
- Bertsatos, A. (2020). *long-bone-diaphyseal-CSG-Toolkit: A GNU Octave CSG Toolkit* (v1.1.0) [Computer software]. Zenodo. <https://doi.org/10.5281/zenodo.3813770>
- Bertsatos, A., Athanasopoulou, K., & Chovalopoulou, M.-E. (2019). Estimating sex using discriminant analysis of mandibular measurements from a modern Greek sample. *Egyptian Journal of Forensic Sciences*, 9(25), 1–12. <https://doi.org/10.1186/s41935-019-0133-7>
- Bertsatos, A., & Chovalopoulou, M.-E. (2018a). Secular change in adult stature of modern Greeks. *American Journal of Human Biology*, 30(2), e23077. <https://doi.org/10.1002/ajhb.23077>
- Bertsatos, A., & Chovalopoulou, M.-E. (2018b). *Validation study of correct operation for the “long-bone-diaphyseal-CSG-Toolkit.”* Zenodo. <https://doi.org/10.5281/zenodo.1466135>
- Bertsatos, A., & Chovalopoulou, M.-E. (2019). A novel method for analyzing long bone diaphyseal cross-sectional geometry. A GNU Octave CSG Toolkit. *Forensic Science International*, 297, 65–71. <https://doi.org/10.1016/j.forsciint.2019.01.041>
- Bertsatos, A., & Chovalopoulou, M.-E. (2020). Advances in Osteometric Sorting: Utilizing Diaphyseal CSG Properties for Lower Limb Skeletal Pair-Matching: Skeletal Pair-Matching with CSG Properties. *Journal of Forensic Sciences*. <https://doi.org/10.1111/1556-4029.14480>
- Bertsatos, A., Chovalopoulou, M.-E., Boskovits, N.-M., Garoufi, N., & Nikita, E. (2020). The impact of activity on pelvic age-at-death estimation. *International Journal of Osteoarchaeology*. <https://doi.org/10.1002/oa.2941>
- Bertsatos, A., Chovalopoulou, M.-E., Brůžek, J., & Bejdová, Š. (2020). Advanced procedures for skull sex estimation using sexually dimorphic morphometric features. *International Journal of Legal Medicine*. <https://doi.org/10.1007/s00414-020-02334-9>
- Bertsatos, A., Chovalopoulou, M.-E., Giannaki, K., & Valakos, E. D. (2018). Morphological variation of the femoral head fovea capitis. *European Journal of Anatomy*, 22(5), 397–402.
- Bertsatos, A., Christaki, A., & Chovalopoulou, M.-E. (2019). Testing the reliability of 3D-ID software in sex and ancestry estimation with a modern Greek sample. *Forensic Science International*, 297, 132–137. <https://doi.org/10.1016/j.forsciint.2019.02.004>

- Bertsatos, A., Garoufi, N., & Chovalopoulou, M.-E. (2020). Advancements in sex estimation using the diaphyseal cross-sectional geometric properties of the lower and upper limbs. *International Journal of Legal Medicine*. <https://doi.org/10.1007/s00414-020-02437-3>
- Bertsatos, A., Gkaniatsou, E., Papageorgopoulou, C., & Chovalopoulou, M.-E. (2019). “What and how should we share?” An inter-method inter-observer comparison of measurement error with landmark-based craniometric datasets. *Anthropologischer Anzeiger*, 109–120. <https://doi.org/10.1127/anthranz/2019/1047>
- Bertsatos, A., Papageorgopoulou, C., Valakos, E. D., & Chovalopoulou, M.-E. (2018). Investigating the sex-related geometric variation of the human cranium. *International Journal of Legal Medicine*, 132(5), 1505–1514. <https://doi.org/10.1007/s00414-018-1790-z>
- Bigoni, L., Velemínská, J., & Brůžek, J. (2010). Three-dimensional geometric morphometric analysis of cranio-facial sexual dimorphism in a Central European sample of known sex. *HOMO*, 61(1), 16–32. <https://doi.org/10.1016/j.jchb.2009.09.004>
- Bocquet-Appel, J.-P., & Masset, C. (1982). Farewell to paleodemography. *Journal of Human Evolution*, 11(4), 321–333. [https://doi.org/10.1016/S0047-2484\(82\)80023-7](https://doi.org/10.1016/S0047-2484(82)80023-7)
- Boskey, A. L., & Coleman, R. (2010). Aging and Bone. *Journal of Dental Research*, 89(12), 1333–1348. <https://doi.org/10.1177/0022034510377791>
- Brooks, S., & Suchey, J. M. (1990). Skeletal age determination based on the os pubis: A comparison of the Acsádi-Nemeskéri and Suchey-Brooks methods. *Human Evolution*, 5(3), 227–238. <https://doi.org/10.1007/BF02437238>
- Bruzek, J. (2002). A method for visual determination of sex, using the human hip bone. *American Journal of Physical Anthropology*, 117(2), 157–168. <https://doi.org/10.1002/ajpa.10012>
- Brůžek, J., Santos, F., Dutailly, B., Murail, P., & Cunha, E. (2017). Validation and reliability of the sex estimation of the human os coxae using freely available DSP2 software for bioarchaeology and forensic anthropology: BRŮŽEK et al. *American Journal of Physical Anthropology*, 164(2), 440–449. <https://doi.org/10.1002/ajpa.23282>
- Buckberry, J. L., & Chamberlain, A. T. (2002). Age estimation from the auricular surface of the ilium: A revised method. *American Journal of Physical Anthropology*, 119(3), 231–239. <https://doi.org/10.1002/ajpa.10130>

- Cameron, M. E., & Pfeiffer, S. (2014). Long bone cross-sectional geometric properties of Later Stone Age foragers and herder–foragers. *South African Journal of Science*, *110*(9/10), 1–11. <https://doi.org/10.1590/sajs.2014/20130369>
- Campanacho, V., Santos, A. L., & Cardoso, H. F. V. (2012). Assessing the influence of occupational and physical activity on the rate of degenerative change of the pubic symphysis in portuguese males from the 19th to 20th century. *American Journal of Physical Anthropology*, *148*(3), 371–378. <https://doi.org/10.1002/ajpa.22059>
- Carrivick, J., Smith, M., & Quincey, D. (2016). *Structure from motion in the geosciences*. Wiley, Blackwell.
- Casado, A. M. (2017). Quantifying Sexual Dimorphism in the Human Cranium: A Preliminary Analysis of a Novel Method. *Journal of Forensic Sciences*, *62*(5), 1259–1265. <https://doi.org/10.1111/1556-4029.13441>
- Čechová, M., Dupej, J., Brůžek, J., Bejdová, Š., Horák, M., & Velemínská, J. (2019). Sex estimation using external morphology of the frontal bone and frontal sinuses in a contemporary Czech population. *International Journal of Legal Medicine*, *133*(4), 1285–1294. <https://doi.org/10.1007/s00414-019-02063-8>
- Chovalopoulou, M.-E., Bertsatos, A., & Papageorgopoulou, C. (2017). Age-related changes in the craniofacial region in a modern Greek population sample of known age and sex. *International Journal of Legal Medicine*, *131*(4), 1103–1111. <https://doi.org/10.1007/s00414-016-1470-9>
- Chovalopoulou, M.-E., Papageorgopoulou, C., & Bertsatos, A. (2017). Cranium asymmetry in a modern Greek population sample of known age and sex. *International Journal of Legal Medicine*, *131*(3), 803–812. <https://doi.org/10.1007/s00414-016-1430-4>
- Chovalopoulou, M.-E., Valakos, E. D., & Manolis, S. K. (2013). Sex determination by three-dimensional geometric morphometrics of the palate and cranial base. *Anthropologischer Anzeiger*, *70*(4), 407–425. <https://doi.org/10.1127/0003-5548/2013/0363>
- Chovalopoulou, M.-E., Valakos, E. D., & Manolis, S. K. (2016a). Sex determination by three-dimensional geometric morphometrics of the vault and midsagittal curve of the neurocranium in a modern Greek population sample. *HOMO*, *67*(3), 173–187. <https://doi.org/10.1016/j.jchb.2015.09.007>

- Chovalopoulou, M.-E., Valakos, E. D., & Manolis, S. K. (2016b). Sex determination by three-dimensional geometric morphometrics of craniofacial form. *Anthropologischer Anzeiger*, 73(3), 195–206. <https://doi.org/10.1127/anthranz/2016/0470>
- Cignoni, P., Callieri, M., Corsini, M., Dellepiane, M., Ganovelli, F., & Ranzuglia, G. (2008). MeshLab: An Open-Source Mesh Processing Tool. *Eurographics Italian Chapter Conference*, 8 pages. <https://doi.org/10.2312/LOCALCHAPTEREVENTS/ITALCHAP/ITALIANCHAPCONF2008/129-136>
- Coquerelle, M., Bookstein, F. L., Braga, J., Halazonetis, D. J., Weber, G. W., & Mitteroecker, P. (2011). Sexual dimorphism of the human mandible and its association with dental development. *American Journal of Physical Anthropology*, 145(2), 192–202. <https://doi.org/10.1002/ajpa.21485>
- Curate, F., Coelho, J., Gonçalves, D., Coelho, C., Ferreira, M. T., Navega, D., & Cunha, E. (2016). A method for sex estimation using the proximal femur. *Forensic Science International*, 266, 579.e1-579.e7. <https://doi.org/10.1016/j.forsciint.2016.06.011>
- Davies, T. G., Shaw, C. N., & Stock, J. T. (2012). A test of a new method and software for the rapid estimation of cross-sectional geometric properties of long bone diaphyses from 3D laser surface scans. *Archaeological and Anthropological Sciences*, 4(4), 277–290. <https://doi.org/10.1007/s12520-012-0101-8>
- Dayal, M. R., Spocter, M. A., & Bidmos, M. A. (2008). An assessment of sex using the skull of black South Africans by discriminant function analysis. *HOMO*, 59(3), 209–221. <https://doi.org/10.1016/j.jchb.2007.01.001>
- Dixon, S. J., & Brereton, R. G. (2009). Comparison of performance of five common classifiers represented as boundary methods: Euclidean Distance to Centroids, Linear Discriminant Analysis, Quadratic Discriminant Analysis, Learning Vector Quantization and Support Vector Machines, as dependent on data structure. *Chemometrics and Intelligent Laboratory Systems*, 95(1), 1–17. <https://doi.org/10.1016/j.chemolab.2008.07.010>
- Duangto, P., & Mahakkanukrauh, P. (2020). Sex estimation from upper limb bones in a Thai population. *Anatomy & Cell Biology*, 53(1), 36. <https://doi.org/10.5115/acb.19.179>
- Eaton, J. W., Bateman, D., Hauberg, S., & Wehbring, R. (2019). *GNU Octave version 5.1.0 manual: A high-level interactive language for numerical computations*. <https://www.gnu.org/software/octave/doc/v5.1.0/>

- Elliott, M., & Collard, M. (2009). FORDISC and the determination of ancestry from cranial measurements. *Biology Letters*, 5(6), 849–852. <https://doi.org/10.1098/rsbl.2009.0462>
- Franklin, D., Cardini, A., Flavel, A., & Kuliukas, A. (2013). Estimation of sex from cranial measurements in a Western Australian population. *Forensic Science International*, 229(1–3), 158.e1-158.e8. <https://doi.org/10.1016/j.forsciint.2013.03.005>
- Franklin, D., Freedman, L., & Milne, N. (2005). Sexual dimorphism and discriminant function sexing in indigenous South African crania. *HOMO*, 55(3), 213–228. <https://doi.org/10.1016/j.jchb.2004.08.001>
- Franklin, D., O’Higgins, P., Oxnard, C. E., & Dadour, I. (2006). Determination of Sex in South African Blacks by Discriminant Function Analysis of Mandibular Linear Dimensions: A Preliminary Investigation Using the Zulu Local Population. *Forensic Science, Medicine and Pathology*, 2(4), 263–268. <https://doi.org/10.1385/FSMP:2:4:263>
- Franklin, D., O’Higgins, P., Oxnard, C. E., & Dadour, I. (2007). Sexual dimorphism and population variation in the adult mandible: Forensic applications of geometric morphometrics. *Forensic Science, Medicine, and Pathology*, 3(1), 15–22. <https://doi.org/10.1385/FSMP:3:1:15>
- Franklin, D., O’Higgins, P., Oxnard, C. E., & Dadour, I. (2008). Discriminant function sexing of the mandible of Indigenous South Africans. *Forensic Science International*, 179(1), 84.e1-84.e5. <https://doi.org/10.1016/j.forsciint.2008.03.014>
- Fruciano, C. (2016). Measurement error in geometric morphometrics. *Development Genes and Evolution*, 226(3), 139–158. <https://doi.org/10.1007/s00427-016-0537-4>
- Fruciano, C., Celik, M. A., Butler, K., Dooley, T., Weisbecker, V., & Phillips, M. J. (2017). Sharing is caring? Measurement error and the issues arising from combining 3D morphometric datasets. *Ecology and Evolution*, 7(17), 7034–7046. <https://doi.org/10.1002/ece3.3256>
- Frutos, L. R. (2005). Metric determination of sex from the humerus in a Guatemalan forensic sample. *Forensic Science International*, 147(2–3), 153–157. <https://doi.org/10.1016/j.forsciint.2004.09.077>
- Fthenakis, W. (1970). Größe und Gewicht griechischer Kinder und Jugendlicher unter Berücksichtigung der Akzeleration. *Anthropologischer Anzeiger*, 32(1), 24–45.
- Garvin, H. M., & Ruff, C. B. (2012). Sexual dimorphism in skeletal browridge and chin morphologies determined using a new quantitative method. *American Journal of Physical Anthropology*, 147(4), 661–670. <https://doi.org/10.1002/ajpa.22036>

- Godde, K., Thompson, M. M., & Hens, S. M. (2018). Sex estimation from cranial morphological traits: Use of the methods across American Indians, modern North Americans, and ancient Egyptians. *HOMO*, 69(5), 237–247. <https://doi.org/10.1016/j.jchb.2018.09.003>
- Gómez-Valdés, J. A., Quinto-Sánchez, M., Menéndez Garmendia, A., Velemínska, J., Sánchez-Mejorada, G., & Bruzek, J. (2012). Comparison of methods to determine sex by evaluating the greater sciatic notch: Visual, angular and geometric morphometrics. *Forensic Science International*, 221(1–3), 156.e1-156.e7. <https://doi.org/10.1016/j.forsciint.2012.04.027>
- Gonzalez, P. N., Bernal, V., & Perez, S. I. (2009). Geometric morphometric approach to sex estimation of human pelvis. *Forensic Science International*, 189(1–3), 68–74. <https://doi.org/10.1016/j.forsciint.2009.04.012>
- Guyomarc'h, P., & Bruzek, J. (2011). Accuracy and reliability in sex determination from skulls: A comparison of Fordisc® 3.0 and the discriminant function analysis. *Forensic Science International*, 208(1–3), 180.e1-180.e6. <https://doi.org/10.1016/j.forsciint.2011.03.011>
- Hasluck, M. M., & Morant, G. M. (1929). Measurements of Macedonian men. *Biometrika*, 21(1–4), 322–336. <https://doi.org/10.1093/biomet/21.1-4.322>
- 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 Enteseal Changes: The New 'Coimbra Method.' *International Journal of Osteoarchaeology*, 26(5), 925–932. <https://doi.org/10.1002/oa.2477>
- İşcan, M. Y., & Miller-Shaivitz, P. (1984). Determination of sex from the Tibia. *American Journal of Physical Anthropology*, 64(1), 53–57. <https://doi.org/10.1002/ajpa.1330640104>
- Jantz, R. L., & Ousley, S. D. (2013). *FORDISC 3: Computerized Forensic Discriminant Functions (Version 3.0.)* [Computer software]. The University of Tennessee.
- Ji, M.-X., & Yu, Q. (2015). Primary osteoporosis in postmenopausal women. *Chronic Diseases and Translational Medicine*, 1(1), 9–13. <https://doi.org/10.1016/j.cdtm.2015.02.006>
- Kabsch, W. (1976). A solution for the best rotation to relate two sets of vectors. *Acta Crystallographica Section A*, 32(5), 922–923. <https://doi.org/10.1107/S0567739476001873>
- Kanchan, T., Gupta, A., & Krishan, K. (2013). Estimation of sex from mastoid triangle – A craniometric analysis. *Journal of Forensic and Legal Medicine*, 20(7), 855–860. <https://doi.org/10.1016/j.jflm.2013.06.016>

- Kano, T., Oritani, S., Michiue, T., Ishikawa, T., Hishmat, A. M., Sogawa, N., Inamori-Kawamoto, O., & Maeda, H. (2015). Postmortem CT morphometry with a proposal of novel parameters for sex discrimination of the mandible using Japanese adult data. *Legal Medicine*, *17*(3), 167–171. <https://doi.org/10.1016/j.legalmed.2014.12.009>
- Keep, R., & Bogin, B. (1999). Eight thousand years of economic and political history in Latin America revealed by anthropometry. *Annals of Human Biology*, *26*(4), 333–351. <https://doi.org/10.1080/030144699282651>
- Komar, D. A., & Buikstra, J. E. (2008). *Forensic anthropology: Contemporary theory and practice*. Oxford University Press.
- Komlos, J. (2009). Anthropometric history: An overview of a quarter century of research. *Anthropologischer Anzeiger*, *67*(4), 341–356. <https://doi.org/10.1127/0003-5548/2009/0027>
- Kranioti, E. F., İşcan, M. Y., & Michalodimitrakis, M. (2008). Craniometric analysis of the modern Cretan population. *Forensic Science International*, *180*(2–3), 110.e1-110.e5. <https://doi.org/10.1016/j.forsciint.2008.06.018>
- Kranioti, E. F., & Michalodimitrakis, M. (2009). Sexual Dimorphism of the Humerus in Contemporary Cretans-A Population-Specific Study and a Review of the Literature. *Journal of Forensic Sciences*, *54*(5), 996–1000. <https://doi.org/10.1111/j.1556-4029.2009.01103.x>
- Kranioti, E. F., Nathana, D., & Michalodimitrakis, M. (2011). Sex estimation of the Cretan humerus: A digital radiometric study. *International Journal of Legal Medicine*, *125*(5), 659–667. <https://doi.org/10.1007/s00414-010-0470-4>
- Kranioti, E. F., & Tzanakis, N. (2015). Estimation of Sex from the Upper Limb in Modern Cretans with the Aid of ROC-Analysis: A Technical Report. *Foresic Research & Criminology International Journal*, *1*(2). <https://doi.org/10.15406/frcij.2015.01.00008>
- Krishan, K., Chatterjee, P. M., Kanchan, T., Kaur, S., Baryah, N., & Singh, R. K. (2016). A review of sex estimation techniques during examination of skeletal remains in forensic anthropology casework. *Forensic Science International*, *261*, 165.e1-165.e8. <https://doi.org/10.1016/j.forsciint.2016.02.007>
- Lacoste Jeanson, A., Santos, F., Villa, C., Dupej, J., Lynnerup, N., & Brůžek, J. (2017). Body mass estimation from the skeleton: An evaluation of 11 methods. *Forensic Science International*, *281*, 183.e1-183.e8. <https://doi.org/10.1016/j.forsciint.2017.10.026>

- Langley, N. R., Dudzik, B., & Cloutier, A. (2018). A Decision Tree for Nonmetric Sex Assessment from the Skull. *Journal of Forensic Sciences*, 63(1), 31–37. <https://doi.org/10.1111/1556-4029.13534>
- Lango Allen, H., Estrada, K., Lettre, G., Berndt, S. I., Weedon, M. N., Rivadeneira, F., Willer, C. J., Jackson, A. U., Vedantam, S., Raychaudhuri, S., Ferreira, T., Wood, A. R., Weyant, R. J., Segrè, A. V., Speliotes, E. K., Wheeler, E., Soranzo, N., Park, J.-H., Yang, J., ... Hirschhorn, J. N. (2010). Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature*, 467(7317), 832–838. <https://doi.org/10.1038/nature09410>
- Lewis, C. J., & Garvin, H. M. (2016). Reliability of the Walker Cranial Nonmetric Method and Implications for Sex Estimation. *Journal of Forensic Sciences*, 61(3), 743–751. <https://doi.org/10.1111/1556-4029.13013>
- Lewis, S. (1999). Quantifying measurement error. *Current and Recent Research in Osteoarchaeology. 2: Proc. of the 4th, 5th and 6th Meetings of the Osteoarchaeological Research Group*.
- Loth, S. R., & Henneberg, M. (2001). Sexually dimorphic mandibular morphology in the first few years of life. *American Journal of Physical Anthropology*, 115(2), 179–186. <https://doi.org/10.1002/ajpa.1067>
- Lovejoy, C. O., Meindl, R. S., Pryzbeck, T. R., & Mensforth, R. P. (1985). Chronological metamorphosis of the auricular surface of the ilium: A new method for the determination of adult skeletal age at death. *American Journal of Physical Anthropology*, 68(1), 15–28. <https://doi.org/10.1002/ajpa.1330680103>
- Lyman, R. L. (1994). *Vertebrate Taphonomy* (1st ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9781139878302>
- Lynnerup, N. (2013). Forensic anthropology and human identification. *Scandinavian Journal of Forensic Science*, 19(1), 16–38. <https://doi.org/10.2478/sjfs-2013-0005>
- Macintosh, A. A., Davies, T. G., Ryan, T. M., Shaw, C. N., & Stock, J. T. (2013). Periosteal versus true cross-sectional geometry: A comparison along humeral, femoral, and tibial diaphyses. *American Journal of Physical Anthropology*, 150(3), 442–452. <https://doi.org/10.1002/ajpa.22218>
- Maggiano, I. S., Schultz, M., Kierdorf, H., Sosa, T. S., Maggiano, C. M., & Tiesler Blos, V. (2008). Cross-sectional analysis of long bones, occupational activities and long-distance trade of the

- Classic Maya from Xcambó—Archaeological and osteological evidence. *American Journal of Physical Anthropology*, 136(4), 470–477. <https://doi.org/10.1002/ajpa.20830>
- Mahakkanukrauh, P., Sinthubua, A., Prasitwattanaseree, S., Ruengdit, S., Singsuwan, P., Praneatpolgrang, S., & Duangto, P. (2015). Craniometric study for sex determination in a Thai population. *Anatomy & Cell Biology*, 48(4), 275. <https://doi.org/10.5115/acb.2015.48.4.275>
- Manolis, S., Neroutsos, A., Zafeiratos, C., & Pentzou-Daponte, A. (1995). Secular changes in body formation of greek students. *Human Evolution*, 10(3), 199–204. <https://doi.org/10.1007/BF02438972>
- Manthey, L., Jantz, R. L., Bohnert, M., & Jellinghaus, K. (2017). Secular change of sexually dimorphic cranial variables in Euro-Americans and Germans. *International Journal of Legal Medicine*, 131(4), 1113–1118. <https://doi.org/10.1007/s00414-016-1469-2>
- Marinescu, M., Panaitescu, V., Rosu, M., Maru, N., & Punga, A. (2014). Sexual dimorphism of crania in a Romanian population: Discriminant function analysis approach for sex estimation. *Romanian Journal of Legal Medicine*, 22(1). <https://doi.org/10.4323/rjlm.2014.21>
- Mays, S. (2015). The effect of factors other than age upon skeletal age indicators in the adult. *Annals of Human Biology*, 42(4), 332–341. <https://doi.org/10.3109/03014460.2015.1044470>
- Mays, S., Ives, R., & Brickley, M. (2009). The effects of socioeconomic status on endochondral and appositional bone growth, and acquisition of cortical bone in children from 19th century Birmingham, England. *American Journal of Physical Anthropology*, 140(3), 410–416. <https://doi.org/10.1002/ajpa.21076>
- Merritt, C. E. (2015). The influence of body size on adult skeletal age estimation methods: Influence of Body Size on Age Estimation. *American Journal of Physical Anthropology*, 156(1), 35–57. <https://doi.org/10.1002/ajpa.22626>
- Meydenbauer, A. (1867). Die Photometrographie. *Wochenblatt Des Architektenvereins Zu Berlin*, 1(14), 125–126.
- Michaels, G., & Matles, A. L. (1970). 21 The Role of the Ligamentum Teres in Congenital Dislocation of the Hip: *Clinical Orthopaedics and Related Research*, 71(1), 199–201. <https://doi.org/10.1097/00003086-197007000-00023>

- Moore, M. K., DiGangi, E. A., Niño Ruíz, F. P., Hidalgo Davila, O. J., & Sanabria Medina, C. (2016). Metric sex estimation from the postcranial skeleton for the Colombian population. *Forensic Science International*, 262, 286.e1-286.e8. <https://doi.org/10.1016/j.forsciint.2016.02.018>
- Muñoz-Muñoz, F., Quinto-Sánchez, M., & González-José, R. (2016). Photogrammetry: A useful tool for three-dimensional morphometric analysis of small mammals. *Journal of Zoological Systematics and Evolutionary Research*, 54(4), 318–325. <https://doi.org/10.1111/jzs.12137>
- Murton, N., Borrini, M., & Eliopoulos, C. (2015). Sexual Dimorphism of the Fovea Capitis Femoris in a Medieval Population from Gloucester, England. *Global Journal of Anthropology Research*, 2(2), 9–14. <https://doi.org/10.15379/2410-2806.2015.02.02.02>
- Nawrocki, S. P. (2010). The nature and sources of error in the estimation of age at death from the skeleton. In K. E. Latham & M. Finnegan (Eds.), *Age estimation of the human skeleton* (pp. 79–101). Charles C Thomas.
- Nikita, E. (2019). Quantitative Sex Estimation Based on Cranial Traits Using R Functions. *Journal of Forensic Sciences*, 64(1), 175–180. <https://doi.org/10.1111/1556-4029.13833>
- Nikita, E., & Chovalopoulou, M.-E. (2017). Regression equations for the estimation of stature and body mass using a Greek documented skeletal collection. *HOMO*, 68(6), 422–432. <https://doi.org/10.1016/j.jchb.2017.11.002>
- Nikita, E., Xanthopoulou, P., Bertsatos, A., Chovalopoulou, M., & Hafez, I. (2019). A three-dimensional digital microscopic investigation of enthesal changes as skeletal activity markers. *American Journal of Physical Anthropology*, ajpa.23850. <https://doi.org/10.1002/ajpa.23850>
- Noble, J., Flavel, A., Aviv, R., & Franklin, D. (2019). Forensic anthropological standards for cranial sex estimation in Canada: Preliminary results. *Australian Journal of Forensic Sciences*, 51(sup1), S1–S4. <https://doi.org/10.1080/00450618.2019.1569720>
- Ogawa, Y., Imaizumi, K., Miyasaka, S., & Yoshino, M. (2013). Discriminant functions for sex estimation of modern Japanese skulls. *Journal of Forensic and Legal Medicine*, 20(4), 234–238. <https://doi.org/10.1016/j.jflm.2012.09.023>
- Oikonomopoulou, E.-K., Valakos, E., & Nikita, E. (2017). Population-specificity of sexual dimorphism in cranial and pelvic traits: Evaluation of existing and proposal of new

functions for sex assessment in a Greek assemblage. *International Journal of Legal Medicine*, 131(6), 1731–1738. <https://doi.org/10.1007/s00414-017-1655-x>

- 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(4), 221–235. <https://doi.org/10.1016/j.jhevol.2004.07.002>
- Perini, T. A., Oliveira, G. L. de, Ornellas, J. dos S., & Oliveira, F. P. de. (2005). Technical error measurement in anthropometry. *Revista Brasileira de Medicina Do Esporte*, 11(1), 81–85. <https://doi.org/10.1590/S1517-86922005000100009>
- Perumal, V., Woodley, S. J., & Nicholson, H. D. (2017). The morphology and morphometry of the fovea capitis femoris. *Surgical and Radiologic Anatomy*, 39(7), 791–798. <https://doi.org/10.1007/s00276-016-1810-y>
- Petaros, A., Sholts, S. B., Slaus, M., Bosnar, A., & Wärmländer, S. K. T. S. (2015). Evaluating sexual dimorphism in the human mastoid process: A viewpoint on the methodology. *Clinical Anatomy*, 28(5), 593–601. <https://doi.org/10.1002/ca.22545>
- Phenice, T. W. (1969). A newly developed visual method of sexing the os pubis. *American Journal of Physical Anthropology*, 30(2), 297–301. <https://doi.org/10.1002/ajpa.1330300214>
- Ramsthaler, F., Kreutz, K., & Verhoff, M. A. (2007). Accuracy of metric sex analysis of skeletal remains using Fordisc® based on a recent skull collection. *International Journal of Legal Medicine*, 121(6), 477–482. <https://doi.org/10.1007/s00414-007-0199-x>
- Roberts, D. F. (1954). The Cretans. A Geographical Analysis of Some Aspects of Their Physical Anthropology. *The Journal of the Royal Anthropological Institute of Great Britain and Ireland*, 84(1/2), 145. <https://doi.org/10.2307/2844007>
- Robinson, M. S., & Bidmos, M. A. (2011). An assessment of the accuracy of discriminant function equations for sex determination of the femur and tibia from a South African population. *Forensic Science International*, 206(1–3), 212.e1–212.e5. <https://doi.org/10.1016/j.forsciint.2010.12.009>
- Ross, A. H., & Manneschi, M. J. (2011). New identification criteria for the Chilean population: Estimation of sex and stature. *Forensic Science International*, 204(1–3), 206.e1–206.e3. <https://doi.org/10.1016/j.forsciint.2010.07.028>

- Ross, A. H., Slice, D. E., & Williams, S. E. (2010). *Geometric Morphometric Tools for the Classification of Human Skulls* (No. 231195; pp. 1–59).
<https://www.ncjrs.gov/pdffiles1/nij/grants/231195.pdf>
- Ross, A. H., & Williams, S. E. (2008). Testing Repeatability and Error of Coordinate Landmark Data Acquired from Crania. *Journal of Forensic Sciences*, 53(4), 782–785.
<https://doi.org/10.1111/j.1556-4029.2008.00751.x>
- 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). Wiley.
- Ruff, C. B., Garofalo, E., & Holmes, M. A. (2013). Interpreting skeletal growth in the past from a functional and physiological perspective. *American Journal of Physical Anthropology*, 150(1), 29–37. <https://doi.org/10.1002/ajpa.22120>
- Ruff, C. B., Holt, B. M., Niskanen, M., Sladěk, V., Berner, M., Garofalo, E., Garvin, H. M., Hora, M., Maijanen, H., Niinimäki, S., Salo, K., Schuplerová, E., & Tompkins, D. (2012). Stature and body mass estimation from skeletal remains in the European Holocene. *American Journal of Physical Anthropology*, 148(4), 601–617. <https://doi.org/10.1002/ajpa.22087>
- Ruff, C. B., Trinkaus, E., Walker, A., & Larsen, C. S. (1993). Postcranial robusticity in Homo. I: Temporal trends and mechanical interpretation. *American Journal of Physical Anthropology*, 91(1), 21–53. <https://doi.org/10.1002/ajpa.1330910103>
- Scheuer, L. (2002). A blind test of mandibular morphology for sexing mandibles in the first few years of life. *American Journal of Physical Anthropology*, 119(2), 189–191.
<https://doi.org/10.1002/ajpa.10098>
- Schlager, S., & Rüdell, A. (2017). Sexual Dimorphism and Population Affinity in the Human Zygomatic Structure-Comparing Surface to Outline Data: Analyzing zygomatic shape. *The Anatomical Record*, 300(1), 226–237. <https://doi.org/10.1002/ar.23450>
- Schutkowski, H. (1993). Sex determination of infant and juvenile skeletons: I. Morphognostic features. *American Journal of Physical Anthropology*, 90(2), 199–205.
<https://doi.org/10.1002/ajpa.1330900206>
- Shackelford, L., Marshall, F., & Peters, J. (2013). Identifying donkey domestication through changes in cross-sectional geometry of long bones. *Journal of Archaeological Science*, 40(12), 4170–4179. <https://doi.org/10.1016/j.jas.2013.06.006>

- Shearer, B. M., Cooke, S. B., Halenar, L. B., Reber, S. L., Plummer, J. E., Delson, E., & Tallman, M. (2017). Evaluating causes of error in landmark-based data collection using scanners. *PLOS ONE*, *12*(11), e0187452. <https://doi.org/10.1371/journal.pone.0187452>
- Shearer, B. M., Sholts, S. B., Garvin, H. M., & Wärmländer, S. K. T. S. (2012). Sexual dimorphism in human browridge volume measured from 3D models of dry crania: A new digital morphometrics approach. *Forensic Science International*, *222*(1–3), 400.e1–400.e5. <https://doi.org/10.1016/j.forsciint.2012.06.013>
- Shen, S. (2013). Accurate Multiple View 3D Reconstruction Using Patch-Based Stereo for Large-Scale Scenes. *IEEE Transactions on Image Processing*, *22*(5), 1901–1914. <https://doi.org/10.1109/TIP.2013.2237921>
- Siebenrock, K. A., Steppacher, S. D., Albers, C. E., Haefeli, P. C., & Tannast, M. (2013). Diagnosis and Management of Developmental Dysplasia of the Hip from Triradiate Closure Through Young Adulthood: *The Journal of Bone and Joint Surgery*, *95*(8), 749–755. <https://doi.org/10.2106/00004623-201304170-00012>
- Šlaus, M., Bedić, Ž., Strinović, D., & Petrovečki, V. (2013). Sex determination by discriminant function analysis of the tibia for contemporary Croats. *Forensic Science International*, *226*(1–3), 302.e1–302.e4. <https://doi.org/10.1016/j.forsciint.2013.01.025>
- Slice, D., & Ross, A. (2018). *3D-ID: geometric morphometric classification of crania for forensic scientists* (Version 2018-05-27) [Computer software]. <http://www.3d-id.org>
- Sparacello, V. S., & Pearson, O. M. (2010). The importance of accounting for the area of the medullary cavity in cross-sectional geometry: A test based on the femoral midshaft. *American Journal of Physical Anthropology*, *143*(4), 612–624. <https://doi.org/10.1002/ajpa.21361>
- Spradley, K. M. (2016). Metric Methods for the Biological Profile in Forensic Anthropology: Sex, Ancestry, and Stature. *Academic Forensic Pathology*, *6*(3), 391–399. <https://doi.org/10.23907/2016.040>
- Spradley, K. M., & Jantz, R. L. (2011). Sex Estimation in Forensic Anthropology: Skull Versus Postcranial Elements. *Journal of Forensic Sciences*, *56*(2), 289–296. <https://doi.org/10.1111/j.1556-4029.2010.01635.x>

- Spradley, K. M., & Jantz, R. L. (2016). Ancestry Estimation in Forensic Anthropology: Geometric Morphometric versus Standard and Nonstandard Interlandmark Distances. *Journal of Forensic Sciences*, 61(4), 892–897. <https://doi.org/10.1111/1556-4029.13081>
- Spradley, K. M., Jantz, R. L., Robinson, A., & Peccerelli, F. (2008). Demographic Change and Forensic Identification: Problems in Metric Identification of Hispanic Skeletons. *Journal of Forensic Sciences*, 53(1), 21–28. <https://doi.org/10.1111/j.1556-4029.2007.00614.x>
- Steckel, R. H. (2008). Biological Measures of the Standard of Living. *Journal of Economic Perspectives*, 22(1), 129–152. <https://doi.org/10.1257/jep.22.1.129>
- Steyn, M., & İşcan, M. Y. (1997). Sex determination from the femur and tibia in South African whites. *Forensic Science International*, 90(1–2), 111–119. [https://doi.org/10.1016/S0379-0738\(97\)00156-4](https://doi.org/10.1016/S0379-0738(97)00156-4)
- Steyn, M., & Patriquin, M. L. (2009). Osteometric sex determination from the pelvis—Does population specificity matter? *Forensic Science International*, 191(1–3), 113.e1–113.e5. <https://doi.org/10.1016/j.forsciint.2009.07.009>
- Stock, J. T., & Pfeiffer, S. (2001). Linking structural variability in long bone diaphyses to habitual behaviors: Foragers from the southern African Later Stone Age and the Andaman Islands. *American Journal of Physical Anthropology*, 115(4), 337–348. <https://doi.org/10.1002/ajpa.1090>
- Stock, J. T., & Shaw, C. N. (2007). Which measures of diaphyseal robusticity are robust? A comparison of external methods of quantifying the strength of long bone diaphyses to cross-sectional geometric properties. *American Journal of Physical Anthropology*, 134(3), 412–423. <https://doi.org/10.1002/ajpa.20686>
- Susanne, C. (1985). Living conditions and secular trend. *Journal of Human Evolution*, 14(4), 357–370. [https://doi.org/10.1016/S0047-2484\(85\)80042-7](https://doi.org/10.1016/S0047-2484(85)80042-7)
- Tallman, S. D., & Blanton, A. I. (2020). Distal Humerus Morphological Variation and Sex Estimation in Modern Thai Individuals. *Journal of Forensic Sciences*, 65(2), 361–371. <https://doi.org/10.1111/1556-4029.14218>
- Thomas, R. M. (2020). Sex determination using DNA and its impact on biological anthropology. In *Sex Estimation of the Human Skeleton* (pp. 343–350). Elsevier. <https://doi.org/10.1016/B978-0-12-815767-1.00021-3>

- Ubelaker, D. H., & DeGaglia, C. M. (2017). Population variation in skeletal sexual dimorphism. *Forensic Science International*, 278, 407.e1-407.e7.
<https://doi.org/10.1016/j.forsciint.2017.06.012>
- Urbanová, P., Ross, A. H., Jurda, M., & Nogueira, M.-I. (2014). Testing the reliability of software tools in sex and ancestry estimation in a multi-ancestral Brazilian sample. *Legal Medicine*, 16(5), 264–273. <https://doi.org/10.1016/j.legalmed.2014.06.002>
- Valaoras, V. G. (1970). Biometric studies of army conscripts in Greece. Mean stature and ABO blood-group distribution. *Human Biology*, 42(2), 184–201.
- von Cramon-Taubadel, N., Frazier, B. C., & Lahr, M. M. (2007). The problem of assessing landmark error in geometric morphometrics: Theory, methods, and modifications. *American Journal of Physical Anthropology*, 134(1), 24–35. <https://doi.org/10.1002/ajpa.20616>
- Walker, P. L. (2008). Sexing skulls using discriminant function analysis of visually assessed traits. *American Journal of Physical Anthropology*, 136(1), 39–50.
<https://doi.org/10.1002/ajpa.20776>
- Walrath, D. E., Turner, P., & Bruzek, J. (2004). Reliability test of the visual assessment of cranial traits for sex determination. *American Journal of Physical Anthropology*, 125(2), 132–137.
<https://doi.org/10.1002/ajpa.10373>
- Waltenberger, L., Rebay-Salisbury, K., & Mitteroecker, P. (2021). Three-dimensional surface scanning methods in osteology: A topographical and geometric morphometric comparison. *American Journal of Physical Anthropology*, ajpa.24204. <https://doi.org/10.1002/ajpa.24204>
- Wenger, D., Miyanji, F., Mahar, A., & Oka, R. (2007). The Mechanical Properties of the Ligamentum Teres: A Pilot Study to Assess Its Potential for Improving Stability in Children's Hip Surgery. *Journal of Pediatric Orthopaedics*, 27(4), 408–410.
<https://doi.org/10.1097/01.bpb.0000271332.66019.15>
- Wescott, D. J., & Drew, J. L. (2015). Effect of obesity on the reliability of age-at-death indicators of the pelvis: EFFECTS OF OBESITY ON AGE-AT-DEATH. *American Journal of Physical Anthropology*, 156(4), 595–605. <https://doi.org/10.1002/ajpa.22674>
- Wilczak, C. A., Mariotti, V., Pany-Kucera, D., Villotte, S., & Henderson, C. Y. (2017). Training and interobserver reliability in qualitative scoring of skeletal samples. *Journal of Archaeological Science: Reports*, 11, 69–79. <https://doi.org/10.1016/j.jasrep.2016.11.033>

- Williams, F. L., & Richtsmeier, J. T. (2003). Comparison of mandibular landmarks from computed tomography and 3D digitizer data. *Clinical Anatomy*, 16(6), 494–500.
<https://doi.org/10.1002/ca.10095>
- Wright, R. (2012). *Guide to using the CRANID programs Cr6bInd: For linear and nearest neighbours discriminant analysis*.
- Xanthopoulou, P., Valakos, E. D., Youlatos, D., & Nikita, E. (2018). Assessing the accuracy of cranial and pelvic ageing methods on human skeletal remains from a modern Greek assemblage. *Forensic Science International*, 286, 266.e1-266.e8.
<https://doi.org/10.1016/j.forsciint.2018.03.005>
- Yang, W., Liu, X., Wang, K., Hu, J., Geng, G., & Feng, J. (2019). Sex Determination of Three-Dimensional Skull Based on Improved Backpropagation Neural Network. *Computational and Mathematical Methods in Medicine*, 2019, 1–8. <https://doi.org/10.1155/2019/9163547>
- Yang, W., Zhou, M., Zhang, P., Geng, G., Liu, X., & Zhang, H. (2020). Skull Sex Estimation Based on Wavelet Transform and Fourier Transform. *BioMed Research International*, 2020, 1–10.
<https://doi.org/10.1155/2020/8608209>
- Zaki, M. E., Azab, A. A., Yousef, W., Wassal, E. Y., & El-Bassyouni, H. T. (2015). Cross-sectional analysis of long bones in a sample of ancient Egyptians. *The Egyptian Journal of Radiology and Nuclear Medicine*, 46(3), 675–681. <https://doi.org/10.1016/j.ejrn.2015.03.008>
- Χοβαλοπούλου, Μ.-Ε. (2014). *Ανάπτυξη μεθοδολογίας φυλετικού προσδιορισμού με εφαρμογή τριδιάστατης γεωμετρικής μορφομετρικής ανάλυσης (3DGM) σε κρανία σύγχρονου ελληνικού πληθυσμού [Διδακτορική Διατριβή]*. Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών.

Βιογραφικό Σημείωμα

Name: **Μπερτσάτος Ανδρέας, Bertatos Andreas**

Affiliation: Department of Animal and Human Physiology, Faculty of Biology, National and Kapodistrian University of Athens, Panepistimiopolis, GR 157 01 Athens, Greece.

e-mail: abertsatos@biol.uoa.gr, andreas.bertatos@gmail.com

OrCID ID: <http://orcid.org/0000-0001-8734-5236>

GitHub Repository: <https://github.com/pr0m1th3as>

Education

- 2002: Graduated with First Class (Hons) in Electronic Engineering (BEng Electronic Systems Design) from the University of Luton, UK.
- since 2016: Currently enrolled as PhD candidate in Physical Anthropology at the Department of Animal and Human Physiology, Faculty of Biology, University of Athens. Subject: "Three-dimensional geometric morphometrics of the skeleton of the modern Greek population with photogrammetric techniques".

Research Collaborations & Field Experience

- 2015: Collaboration with Research Associate Dr. Maria-Eleni Chovalopoulou at the Wiener Laboratory of Archaeological Sciences of the American School of Classical Studies at Athens (Greece) in the project "Application of the 3DGM method on Greek populations from different era".
- 2016: Collaboration with Professor Sotiris Manolis and Dr. Maria-Eleni Chovalopoulou in the anthropological analysis of the skeletal remains from Kastrouli excavation site in Desfina conducted in the Archaeological Museum of Delphi (Fokis, Greece).
- 2017-2019: Collaboration with Associate Professor Efthymia Nikita at the Science and Technology in Archaeology and Culture Research Center of the Cyprus Institute (Nicosia, Cyprus).
- 2019: Collaboration with Professor Jaroslav Bruzek at the Department of Anthropology and Human Genetics of Charles University (Prague, Czech Republic) under the provision of an Erasmus+ Internship grant in the project "A comparative study between Greek and Czech population in the expression of sexual dimorphism of the skull with special regards on the sex assessment using 3D models from computer tomography, photogrammetry and surface scanner."
- 2019-2020: Collaboration with Research Associate Dr. Maria-Eleni Chovalopoulou at the Science and Technology in Archaeology and Culture Research Center of the Cyprus Institute (Nicosia, Cyprus) in the MSCA-IF project "Human Jigsaw".

Teaching, Mentoring & Supervising Experience

- 2015 - 2020: Mentoring undergraduate final year projects related to Physical Anthropology in collaboration with Prof. Efstratios Valakos at the Department of Animal and Human Physiology of the University of Athens.
- 2016 - 2018: Supervising semester mini project "Assessing MNI from archaeological commingled burials" of the module of Physical Anthropology of the Faculty of Biology at the University of Athens.
- 2017 - 2018: Teaching assistance in Physical Anthropology Laboratory work at the Department of Animal and Human Physiology of the University of Athens.
- 2017 - 2019: Visiting lecturer in the postgraduate module Physical Anthropology Methods of the Master's Program "History of Medicine and Biological Anthropology: Health, Disease and Natural Selection" at the Aristotel University of Thessaloniki.

Publications (peer-reviewed)

1. **Bertsatos A**, Chovalopoulou M-E (2016) Technical note: A GNU octave function for smith's mean measure of divergence. *Bioarchaeology of the Near East* 10:69-73. <http://www.anthropology.uw.edu.pl/10/bne-10-04.pdf>
2. Chovalopoulou M-E, **Bertsatos A**, Manolis SK (2017) Landmark based sex discrimination on the crania of archaeological Greek populations. A comparative study based on the cranial sexual dimorphism of a modern Greek population. *Mediterranean Archaeology and Archaeometry* 17(1):37-46. <https://doi.org/10.5281/zenodo.258084>
3. Chovalopoulou M-E, **Bertsatos A** (2017) Estimating Sex of Modern Greeks Based on the Foramen Magnum Region. *Journal of Anthropology*. <https://doi.org/10.1155/2017/9129801>
4. Chovalopoulou M-E, **Bertsatos A**, Manolis SK (2017) Identification of skeletal remains from a Mycenaean burial in Kastrouli-Desfina, Greece. *Mediterranean Archaeology and Arhaeometry* 17(1):265-269. <https://doi.org/10.5281/zenodo.556353>
5. Chovalopoulou M-E, **Bertsatos A**, Papageorgopoulou C (2017) Age-related changes in the craniofacial region in a modern Greek population sample of known age and sex. *International Journal of Legal Medicine* 131(4):1103–1111. <https://doi.org/10.1007/s00414-016-1470-9>
6. Chovalopoulou M-E, Papageorgopoulou C, **Bertsatos A** (2017) Cranium asymmetry in a modern Greek population sample of known age and sex. *International Journal of Legal Medicine* 131(3):803–812. <https://doi.org/10.1007/s00414-016-1430-4>

7. **Bertsatos A**, Chovalopoulou M-E (2018) Secular change in adult stature of modern Greeks. *American Journal of Human Biology* 30(2):e23077. <https://doi.org/10.1002/ajhb.23077>
8. **Bertsatos A**, Chovalopoulou M-E, Giannaki K, Valakos E (2018) Morphological variation of the femoral head fovea capitis. *European Journal of Anatomy* 22(5):397–402. <http://eurjanat.com/web/paper.php?id=180121ab>
9. **Bertsatos A**, Papageorgopoulou C, Valakos E, Chovalopoulou M-E (2018) Investigating the sex-related geometric variation of the human cranium. *International Journal of Legal Medicine* 132(5):1505–1514. <https://doi.org/10.1007/s00414-018-1790-z>
10. Chovalopoulou M-E, **Bertsatos A**, Zoi I, Manolis SK, Valakos E (2018) Reliability testing of metric methods for sex determination in ancient skeletal remains in Greece. *Mediterranean Archaeology and Archaeometry* 18(1):37–47. <https://doi.org/10.5281/zenodo.1069518>
11. **Bertsatos A**, Athanasopoulou K, Chovalopoulou M-E (2019) Estimating sex using discriminant analysis of mandibular measurements from a modern greek sample. *Egyptian Journal of Forensic Sciences* 9(25):1–12. <https://doi.org/10.1186/s41935-019-0133-7>
12. **Bertsatos A**, Chovalopoulou M-E (2019) A novel method for analyzing long bone diaphyseal cross-sectional geometry. A GNU Octave CSG Toolkit. *Forensic Science International* 297:65–71. <https://doi.org/10.1016/j.forsciint.2019.01.041>
13. **Bertsatos A**, Chovalopoulou M-E (2019) Validation study of osteometric techniques for sorting commingled human skeletal remains in archaeological samples. *International Journal of Osteoarchaeology* 29(2):253–259. <https://doi.org/10.1002/oa.2733>
14. **Bertsatos A**, Christaki A, Chovalopoulou M-E (2019) Testing the reliability of 3D-ID software in sex and ancestry estimation with a modern Greek sample. *Forensic Science International* 297:132–137. <https://doi.org/10.1016/j.forsciint.2019.02.004>
15. Nikita E, Xanthopoulou P, **Bertsatos A**, Chovalopoulou M-E, Hafez I (2019) A three-dimensional digital microscopic investigation of enthesal changes as skeletal activity markers. *American Journal of Physical Anthropology* 169(4):704–713. <https://doi.org/10.1002/ajpa.23850>

16. **Bertsatos A**, Gkaniatsou E, Papageorgopoulou C, Chovalopoulou M-E (2020) “What and how should we share?” An inter-method inter-observer comparison of measurement error with landmark-based craniometric datasets. *Anthropologischer Anzeiger* 77(2):109–120. <https://doi.org/10.1127/anthranz/2019/1047>
17. **Bertsatos A**, Chovalopoulou M-E (2020) Advances in Osteometric Sorting: Utilizing Diaphyseal CSG Properties for Lower Limb Skeletal Pair-Matching. *Journal of Forensic Sciences* 65(5):1400–1405. <https://doi.org/10.1111/1556-4029.14480>
18. **Bertsatos A**, Chovalopoulou M-E, Brůžek J, Bejdová Š (2020) Advanced procedures for skull sex estimation using sexually dimorphic morphometric features. *International Journal of Legal Medicine* 134(5):1927–1937. <https://doi.org/10.1007/s00414-020-02334-9>
19. Garoufi N, **Bertsatos A**, Chovalopoulou M-E, Villa C (2020) Forensic sex estimation using the vertebrae: an evaluation on two European populations. *International Journal of Legal Medicine* 134(6):2307-2318. <https://doi.org/10.1007/s00414-020-02430-w>
20. **Bertsatos A**, Garoufi N, Chovalopoulou M-E (2020) Advancements in sex estimation using the diaphyseal cross-sectional geometric properties of the lower and upper limbs. *International Journal of Legal Medicine* (online 08 October 2020). <https://doi.org/10.1007/s00414-020-02437-3>
21. **Bertsatos A**, Chovalopoulou M-E, Boskovits, N-M, Garoufi N, Nikita E (2020) The impact of activity on pelvic age-at-death estimation. *International Journal of Osteoarchaeology* (online 09 December 2020). <https://doi.org/10.1002/oa.2941>
22. **Bertsatos A**, Chovalopoulou M-E (2021) Integration patterns of the asymmetric components of the skull: A covariational statistical approach on a modern Greek sample. *HOMO – Journal of Comparative Human Biology* (online 06 April 2021). <https://doi.org/10.1127/homo/2021/1369>

Chapters in Books (peer-reviewed)

1. Chovalopoulou M-E, **Bertsatos A** (2018) Exploring the shape variation of the human cranium. A geometric morphometrics study on a modern Greek population sample. In Rissech C, Lloveras L, Nadal J, Fullola JM (Eds), *Geometric Morphometrics. Trends in Biology, Paleobiology and Archaeology*. Barcelona: Universitat de Barcelona. pp.25-39.

Other Publications/Works

- **Bertsatos A** (2017). A set of GNU Octave functions for biodistance analysis based on non-metric traits. Source code available at <https://github.com/pr0m1th3as/non-metric-trait-analysis>
- **Bertsatos A** (2018). A set of GNU Octave functions for manipulating triangular meshes in Wavefront obj format. Source code available at <https://github.com/pr0m1th3as/wavefront-obj-mesh-package>
- **Bertsatos A** (2018). long-bone-diaphyseal-CSG-Toolkit: A GNU Octave CSG Toolkit (Version v1.0.1). Zenodo. <https://doi.org/10.5281/zenodo.1467045>. Source code available at <https://github.com/pr0m1th3as/long-bone-diaphyseal-CSG-Toolkit>
- **Bertsatos A**, Chovalopoulou, M-E. (2018) Validation study of correct operation for the "long-bone-diaphyseal-CSG-Toolkit". Zenodo. <https://doi.org/10.5281/zenodo.1466135>
- **Bertsatos A**, Chovalopoulou, M-E. (2018) Testing dataset for the "long-bone-diaphyseal-CSG-Toolkit" [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.1466962>
- **Bertsatos A** (2019). skullanalyzer: a concrete way of extracting cranial morphometric features (Version v1.0). Zenodo. <https://doi.org/10.5281/zenodo.3594565>. Source code available at <https://github.com/pr0m1th3as/skullanalyzer>
- **Bertsatos, A.** (2019) User Manual & Algorithm Description Document "skullanalyzer" v1.0 (Version v1.0). Zenodo. <https://doi.org/10.5281/zenodo.3519248>
- **Bertsatos A** (2019) Testing dataset for the "skullanalyzer" program (Version v1.0) [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.3519103>
- **Bertsatos A** (2020). long-bone-diaphyseal-CSG-Toolkit: A GNU Octave CSG Toolkit (Version v1.1.0). Zenodo. <https://doi.org/10.5281/zenodo.3813770>. Source code available at <https://github.com/pr0m1th3as/long-bone-diaphyseal-CSG-Toolkit>
- **Bertsatos A** (2020). A GNU Octave function for skeletal sex estimation. Source code available at <https://github.com/pr0m1th3as/skeletal-sex-estimation>

Book Translations

- Scientific supervision of the Greek translation of the textbook: "Jurmain, R., Kilgore, L., Trevathan, W., Ciochon, R.L. (Eds) 2013-14. Introduction to Physical Anthropology", Parisianou Press.

Reviewer in the following Journals (recorded and verified by Publons: K-2156-2019)

- European Journal of Anatomy (Sociedad Anatomica Espanola)
- International Journal of Osteoarchaeology (Wiley)

Ανατύπωση πρωτότυπων δημοσιεύσεων



Investigating the sex-related geometric variation of the human cranium

Andreas Bertatos¹ · Christina Papageorgopoulou² · Efstratios Valakos¹ · Maria-Eleni Chovalopoulou¹

Received: 21 August 2017 / Accepted: 22 January 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract

Accurate sexing methods are of great importance in forensic anthropology since sex assessment is among the principal tasks when examining human skeletal remains. The present study explores a novel approach in assessing the most accurate metric traits of the human cranium for sex estimation based on 80 ectocranial landmarks from 176 modern individuals of known age and sex from the Athens Collection. The purpose of the study is to identify those distance and angle measurements that can be most effectively used in sex assessment. Three-dimensional landmark coordinates were digitized with a Microscribe 3DX and analyzed in GNU Octave. An iterative linear discriminant analysis of all possible combinations of landmarks was performed for each unique set of the 3160 distances and 246,480 angles. Cross-validated correct classification as well as multivariate DFA on top performing variables reported 13 craniometric distances with over 85% classification accuracy, 7 angles over 78%, as well as certain multivariate combinations yielding over 95%. Linear regression of these variables with the centroid size was used to assess their relation to the size of the cranium. In contrast to the use of generalized procrustes analysis (GPA) and principal component analysis (PCA), which constitute the common analytical work flow for such data, our method, although computational intensive, produced easily applicable discriminant functions of high accuracy, while at the same time explored the maximum of cranial variability.

Keywords Geometric morphometrics · Cranial metric traits · Discriminant functions · Greek population

Introduction

Sex assessment is among the principal tasks when examining human skeletal remains. Since human identification as well as various skeletal analyses heavily rely on sex, accurate sexing methods are of great importance in forensic anthropology and bioarchaeology. Over the years, a number of qualitative and quantitative methods have been proposed based on cranial and post-cranial traits [e.g., [1–5]] with varying rates of success and the highest correct sex classification concerning the pelvis and the cranium [4, 5]. Qualitative methods assess specific

sexually dimorphic anatomical traits. In general, qualitative methods are easier to apply but heavily rely on the experience of the observer; hence, they tend to be more subjective and prone to inter-observer error [5, 6]. On the other hand, quantitative methods rely on measurements, which may be more time consuming and often require specialized equipment [7] but are less prone to observer bias and often produce more accurate results [5, 6, 8].

In recent years, the quantitative methods advanced to more complicated techniques using geometric morphometrics, which allow more advanced analysis of the shape variation than traditional measuring methods [5, 9, 10]. Although the application of geometric morphometrics requires the use of sophisticated equipment such as laser scanners and 3D digitizers, its more elaborate shape analysis of skeletal elements has revealed further information on the variation of the human skeleton such as asymmetry [11], age-related changes [12], and secular changes in morphology [13]. Additionally, geometric morphometrics have produced methods for sex estimation with improved correct classification results over traditional quantitative [4, 5, 8]. In contrast to the qualitative methods, geometric morphometrics not only produce far more robust

✉ Maria-Eleni Chovalopoulou
marilenachov@biol.uoa.gr

¹ Department of Animal and Human Physiology, Faculty of Biology, National and Kapodistrian University of Athens, Panepistimiopolis, 157 81 Athens, GR, Greece

² Laboratory of Physical Anthropology, Department of History and Ethnology, Democritus University of Thrace, 1 P. Tsaldari Street, 69100 Komotini, Greece

and reproducible results due to the minimization of inter-observer error, but also allow more accurate measurement of the magnitude of sexual dimorphism [8], which is known to vary in its expression in different populations [14].

The use of ectocranial landmarks in geometric morphometrics for analyzing the shape variation of the human cranium with respect to sex has been previously studied [9, 13, 15–18]. The use of generalized procrustes analysis (GPA) and principal component analysis (PCA), which constitute the common analytical work flow for such data, produce results that are difficult to apply in forensic cases unless all landmarks are present and the reference data are also available to the forensic anthropologist or the bioarchaeologist, who need to examine an isolated sample. This has resulted to some criticism against the usability and effectiveness of geometric morphometrics in sex assessment [7, 19]. On the other hand, the traditional quantitative methods, which are straightforward to implement (simple measurements), often rely on cranial traits known from visual observation to be sexually dimorphic [20, 21]. As a result, significant sex-related variation, which is captured with GPA and PCA, is left unexplored and unused for sex discriminant functions [5]. The present study aims to fill this analytical gap by iteratively exploring the sexually dimorphic properties of almost a quarter of million possible

combinations of euclidean distances and angles based on ectocranial landmarks in a modern Greek skeletal collection.

Materials and methods

The sample consisted of 176 adult crania of Greek individuals of known sex: 94 males and 82 females. These crania constitute part of the modern skeletal reference collection (known as the Athens Collection), which is housed in the Department of Animal and Human Physiology at the National and Kapodistrian University of Athens. The collection consists of 225 skeletons. Information on the sex, age at death, cause of death as well as place and year of birth for each individual in the collection is derived from death records [22]. All individuals are Greek nationals with recorded age-at-death ranging from 19 to 99 years old and their respective year of birth spanning from 1879 to 1965. Individuals without any evidence of pathological deformation and absence of craniotomy were selected.

Data acquisition Eighty landmarks (12 midline and 34 bilateral) on the outer surface of the skull, chosen on the basis of adequately illustrating the geometry of the cranium, were used (Table 1). Most of landmarks definitions were obtained after

Table 1 List of digitized landmarks

1	Nasion	24	Asterion left and right
2	Glabella	25	Entomion left
3	Bregma	26	Supramastoid crest–squamous suture intersection left and right
4	Lambda	27	Crotaphion left and right
5	Opisthocranion	28	Coronale left and right
6	Opisthion	29	Sphenion left and right
7	Basion	30	Frontotemporale left and right
8	Nasospinale	31	Landmark × left and right
9	Hormion	32	Supra auricular left and right
10	Staphylion	33	Zygotemporale superior left and right
11	Staurion	34	Jugale left and right
12	Foramen incisivum	35	Frontomalare temporale left and right
13	Foraminolaterale left and right	36	Porion left and right
14	Occipitocondylion mediale left and right	37	Zygotemporale inferior left and right
15	Occipitocondylion posterior left and right	38	Zygomaxillare left and right
16	Occipitocondylion laterale left and right	39	Frontomalare orbitale left and right
17	Occipitocondylion anterior left and right	40	Supraconchion left and right
18	Caroticum mediale left and right	41	Maxillofrontale left and right
19	Spinale left and right	42	Subconchion left and right
20	Ovale mediale left and right	43	Ectoconchion left and right
21	Postalverion left and right	44	Infraorbitale left and right
22	Infratemporale left and right	45	Apertion left and right
23	Mastoidale left and right	46	Maxillonasofrontale left and right

Bigoni et al. [9], while definition of *supramastoid crest–squamous suture intersection* was obtained after Franklin et al. [20]. Regarding the *porion*, *glabella*, and *opisthocranium* landmarks, definitions were obtained from Whites' book [23]. All landmarks were type I, type II, and type III [24] and can be unambiguously located on the cranial surface [9]. Three-dimensional landmark coordinates were digitized (by MC and AB) with a Microscribe 3DX (Immersion Cor, San Jose, California), whose rated accuracy is ± 0.23 mm. Inter- and intra-observer errors regarding landmark digitization on the crania of the Athens Collection have been previously reported [12] ranging from 0.3 to 1.8 mm. Moreover, for the purposes of the present study, the most sexually dimorphic traits (13 distances and 7 angles) have been measured (by AB) with a mechanical caliper on 20 randomly selected crania (10 males, 10 females) to assess the usability of the produced discriminant functions when measuring with simple equipment.

Data analysis All available coordinates of the 80 landmarks were analyzed in GNU Octave [25]. Purpose specific programming functions, which are available upon request, iteratively calculated all possible combinations of distances and angles from the battery of landmarks included in the present study. For each unique set of the 3160 distances and 246,480 angles, a linear discriminant analysis was performed to calculate the correct classification percentage of the original group for each single variable (distance or angle). Thirteen distances and seven angles with the highest correct classification score were further analyzed in SPSS (IBM SPSS version 23.0, Armonk, NY). Discriminant function analysis (DFA) for leave one out cross-validated correct classification was performed for all 20 variables. Multivariate DFA was further performed on all angles and distances as separately as well as joined. Both enter and step-wise methods were used in multivariate analysis. Additionally, linear regression of each variable on the centroid size was performed to assess the relation of each variable to the size of the cranium. The centroid size was calculated from 63 landmarks that were available in 167 individuals, who were included in all 20 angle and distance variables. Finally, the technical error of measurement (TEM) as well as the scaled TEM has been calculated [26] for each of the most sexually dimorphic traits between the measurements calculated by landmark coordinates and those taken with the caliper to assess the accuracy of the later method and determine sampling variability between the two measurement techniques.

Results

The iterative approach adopted in the present study for investigating the sex-related geometric variation of the human

cranium yielded numerous sexually dimorphic distance and angle measurements that yield reasonably high correct classification. Searching through the univariate results showed 60 distances with correct classification higher than 82% and 61 angles with correct classification higher than 75%. However, we have decided to focus on the most sexually dimorphic distance and angle variables, which were further used in multivariate analysis. Thirteen distances with correct classification higher than 85% and seven angles with correct classification higher than 78% were chosen for this matter. The descriptive statistics of these measurements along with their corresponding landmarks are presented in Table 2.

The aforementioned correct classification percentages were used as a threshold for selecting the most sexually dimorphic measurements based on the original group classification, which was evaluated in the preliminary iterative analysis in GNU Octave. The DFA in SPSS of the selected 20 angles and distances altogether verified the original group correct classification results but also yielded the cross-validated correct classification along with the discriminant functions presented in Table 3. In addition to the total correct classification for each variable, the corresponding percentages for males and females are also provided, since there are certain measurements that provide remarkably better classification in favor of one sex over the other. It should be noted that the sectioning point for all the discriminant functions in Table 3 is zero with positive values classifying for male.

Whereas the highest cross-validated correct total classification for D1 and A1 was 87.4 and 80.6%, respectively, the multivariate analysis yielded even higher percentages. The correct classification results of the multivariate analysis using enter and stepwise DFA on the 20 variables are summarized in Table 4. Although grouping angles and distances in multivariate DFA (enter method) yielded higher correct classification percentages, the stepwise method revealed that the combination of only four measurements, namely angles A3 and A7 and distances D1 and D6, is the most accurate sex classifier with cross-validated correct classification 93.5%. Additionally, the same classifier provides almost the same accuracy for both males and females (93.3 and 93.8%, respectively). It should be noted that although the multivariate discriminant functions still produce different correct classification scores between males and females, these differences are reduced compared to univariate results. Again, the sectioning point for all the discriminant functions in Table 4 is zero with positive values classifying for male.

The results of the regression analysis between sex discriminating variables and the centroid size are presented in Table 5. As expected, the distance variables explain a much higher percentage of the variance of the size of the cranium (mostly between 70 and 80%) rather than the angle variables (ranging from 17 to 27%). The only exceptions are variables D1 and D12, which correspond to the same measurement from the

Table 2 Descriptives of highest correct classification distances and angles

Dist. ^a	Landmark 1	Landmark 2	Sample size ^b	Mean ^b	95% confidence interval		Standard deviation ^b	Standard error ^b	Normality test ^b
					Lower bound ^b	Upper bound ^b			
D1	Ovale mediale right	Mastoidale right	93/81	53.3/47.5	52.7/46.9	53.8/48.1	2.85/2.75	0.30/0.31	0.200/0.200
D2	Mastoidale left	Frontomalare temporale right	92/81	142.5/134.3	141.5/133.3	143.3/135.3	4.53/4.43	0.47/0.49	0.184/0.061
D3	Sphenion left	Mastoidale right	91/81	141.3/131.5	140.2/130.6	142.4/132.5	5.23/4.27	0.55/0.47	0.200/0.200
D4	Mastoidale left	Coronale right	92/82	159.2/149.0	158.1/147.9	160.4/150.1	5.46/4.98	0.57/0.55	0.200/0.057
D5	Mastoidale left	Jugale right	92/81	132.5/125.0	131.6/124.0	133.4/125.9	4.28/4.27	0.45/0.47	0.200/0.200
D6	Mastoidale left	Sphenion right	90/82	141.6/132.5	140.6/131.4	142.7/133.5	4.96/4.63	0.52/0.51	0.200/0.061
D7	Supraconchion left	Mastoidale right	92/80	139.0/129.5	137.9/128.5	140.0/130.5	4.97/4.56	0.52/0.51	0.200/0.200
D8	Frontotemporale left	Mastoidale right	93/81	148.1/138.5	147.0/137.5	149.2/139.5	5.33/4.53	0.55/0.50	0.200/0.200
D9	Ectoconchion left	Mastoidale right	92/80	135.5/127.2	134.5/126.3	136.4/128.2	4.60/4.36	0.48/0.49	0.200/0.200
D10	Zygotemporale superior left	Mastoidale right	91/80	132.4/123.6	131.4/122.7	133.3/124.5	4.69/4.24	0.49/0.47	0.200/0.200
D11	Jugale left	Mastoidale right	91/80	133.1/124.6	132.1/123.7	134.1/125.6	4.56/4.34	0.48/0.48	0.200/0.200
D12	Ovale mediale left	Mastoidale left	93/82	53.6/48.5	53.0/47.9	54.1/49.0	2.62/2.58	0.27/0.29	0.200/0.200
D13	Glabella	Mastoidale Right	93/81	134.3/124.8	133.3/123.7	135.3/125.8	4.94/4.54	0.51/0.50	0.200/0.200
angl. ^a									
A1	Landmark 1 Glabella	Landmark 2 Supra-Auricular left	94/81	5.3/6.6	5.1/6.4	5.5/6.9	0.96/1.16	0.10/0.13	0.200/0.200
A2	Mastoidale Right	Opisthocranion	91/80	25.0/22.6	24.7/22.2	25.4/23.0	1.68/1.88	0.18/0.21	0.200/0.200
A3	Glabella	Asterion left	93/81	4.5/5.5	4.3/5.3	4.6/5.7	0.73/0.88	0.08/0.10	0.072/0.200
A4	Glabella	Maxillonasofrontale right	94/81	67.9/59.2	66.5/58.0	69.3/60.3	6.74/5.24	0.70/0.58	0.200/0.200
A5	Glabella	Maxillofrontale right	92/81	71.2/64.1	70.1/63.1	72.4/65.2	5.51/4.60	0.57/0.51	0.200/0.189
A6	Ovale mediale left	Mastoidale right	92/81	66.4/69.8	65.9/69.2	67.0/70.3	2.63/2.48	0.27/0.28	0.200/0.200
A7	Glabella	Supramastoid crest-squamous suture intersection left	93/81	5.2/6.3	5.0/6.1	5.4/6.5	0.85/1.02	0.09/0.11	0.068/0.200

^a Distances measured in mm and angles in degrees^b Males/females

Table 3 Discriminant functions and correct classification results of distances and angles

Dist. and Angl.	Sample size	Constant	Coefficient	Percentage of correct classification %		
				Total ^a	Males ^a	Females ^a
D1	174	-18.050	0.357	87.4	87.1	87.7
D2	173	-30.920	0.223	86.7/86.1	88.0	85.2/84.0
D3	172	-28.456	0.208	86.6	85.7	87.7
D4	174	-29.482	0.191	86.2	83.7	89.0
D5	173	-30.157	0.234	86.1/85.5	89.1/88.0	82.7
D6	172	-28.576	0.208	86.0	87.8	84.1
D7	172	-28.123	0.209	86.0	85.9	86.3
D8	174	-28.870	0.201	85.6	86.0	85.2
D9	172	-29.340	0.223	85.5/84.3	85.9	85.0/82.5
D10	171	-28.594	0.223	85.4	84.6	86.3
D11	171	-28.969	0.224	85.4	85.7	85.0
D12	175	-19.690	0.385	85.1	83.9	86.6
D13	174	-27.315	0.210	85.1	84.9	85.2
A1	175	-5.604	0.948	80.6	87.2	72.8
A2	171	-13.462	0.563	79.5/78.9	83.5/82.4	75.0
A3	174	-6.180	1.248	79.3	84.9	72.8
A4	175	-10.484	0.164	78.9	78.7	79.0
A5	173	-13.313	0.196	78.6	77.2	80.2
A6	173	-26.567	0.391	78.6	82.6	74.1
A7	174	-6.102	1.068	78.2	86.0	69.1

^a Single values are given when original group classification equals cross-validated classification, otherwise both values are shown in the form: original group/cross-validated

right and left side, respectively, between landmarks *Ovale mediale-Mastoidale* and the explained variance is about 47%. All regression models presented in Table 5 are statistically significant.

TEM and scaled TEM for distance and angle variables are shown in Table 6. The scaled TEM for manually measured distances varies from 0.8 to 2.3% with respect to the calculated distances based on landmark coordinates. Additionally, TEM is systematically lower than the standard deviation (see Table 2) of each variable for both males and females from the entire skeletal sample and most important is much lower than the difference between upper and lower confidence interval bounds between sexes. The manual measurement of angles revealed high values of scaled TEM for angles A1, A3, and A7 ranging from 10.4 up to 19.1%, whereas the scaled TEM for all other angles does not exceed 4%. Although TEM for angles A1, A3, and A7 is low (1.1, 0.6, and 0.6, respectively) suggesting high precision in measurement, it is of the same magnitude as the observed standard deviation which may result in low accuracy when measuring these particular variables.

In light of the TEM results, a stepwise method DFA was conducted including all angle and distance variables from the entire skeletal sample except for high-level scaled TEM angles. The results (shown in Table 4)

revealed even higher classification scores with cross-validated correct classification reaching 95.2% for the total sample, whereas males are correctly classified at 94.3% and females at 96.3%. This classification is based on a three-variable discriminant function including distances D1 and D6 as well as angle A5.

Discussion

The aim of the present study has been two-fold. We aimed to explore the geometry of the cranium in an iterative manner to identify the most sexually dimorphic traits that can be used for accurate sex identification, while at the same time minimizing the complexity of the application of modern geometric morphometric methods. Indeed, although our analysis was based on three-dimensional coordinates of landmarks, which requires specialized equipment (laser scanner or 3D digitizer in our case), all traits reported in our results can be easily and accurately measured with the use of simple equipment (caliper). For distance traits, all is required a straightforward distance measurement between two landmarks, whereas the traits regarding angles can be easily calculated from the distances associated with the respective landmarks that define the angle of interest.

Table 4 Discriminant functions and correct classification results of multivariate analysis

Variables included ^a	Sample size	Constant	Coefficients	Percentage of correct classification %		
				Total ^b	Males ^b	Females ^b
All angles A1 to A7	169	-0.087	-0.260, +0.303, -2.065, +0.029, +0.039, -0.136, +1.649	86.4/85.2	88.8/86.5	83.8
A2 + A3 + A4 + A6 + A7	169	0.563	0.310, -2.159, +0.056, -0.130, +1.402	86.5/85.3	88.9/86.7	83.8
All distances D1 to D13	167	-30.233	0.107, -0.017, +0.052, +0.069, -0.055, +0.056, +0.008, -0.148, -0.029, +0.193, -0.058, +0.118 0.073	92.2/88.6	90.8/86.2	93.8/91.3
D1 + D4 + D6	167	-30.745	0.214, +0.061, +0.077	90.6	87.8	93.8
all angles and distances A1-A7 and D1-D13	166	-23.655	0.020, +0.093, -2.217, +0.004, +0.027, -0.020, +1.401, +0.011, -0.020, +0.046, +0.013, -0.072, +0.060, +0.028, -0.194, -0.054, +0.270, -0.064, +0.199, +0.102	96.4/91.6	96.5/90.7	96.3/92.5
A3 + A7 + D1 + D6	166	-23.874	-1.951, +1.200, +0.241, +0.106	94.1/93.5	94.4/93.3	93.8
A5 + D1 + D6 ^c	166	-31.370	0.084, +0.214, +0.109	95.8/95.2	95.5/94.3	96.3

^a All variables were included using enter and stepwise methods of the discriminant function analysis

^b Single values are given when original group classification equals cross-validated classification; otherwise, both values are shown in the form: original group/cross-validated

^c Stepwise method on all angle and distance variables except for A1, A3, and A7

This can be achieved using the formula $B = \arccos\left(\frac{a^2+c^2-b^2}{2ac}\right)$, which is derived from the law of sines [27], or simply use an online calculator such as <http://www.calculator.net/triangle-calculator.html>. For illustration purposes, considering the angle **B** formed by three landmarks, namely A, B, C, distance **a** corresponds to the distance between landmarks B and C, distance **b** corresponds to landmarks A and C, and distance **c** between landmarks A and B. Although the discriminant functions reported in the present study can be easily implemented with a single vernier caliper, the use of geometric morphometrics and digitization of 3D coordinates of landmarks constitutes the backbone of our research objective, since almost a quarter of a million distinct combinations of distances and angles could not have been evaluated manually. The TEM results reported in the present study suggest that manual measurements are quite accurate and precise for implementing the proposed sexual discriminant functions without the need for specialized and expensive equipment (3D microscribe). With the exception of angles A1, A3, and A7, the scaled TEM lies well within acceptable levels [28].

With respect to exploring accurate sex identifiers, our results have identified sexually dimorphic traits that have not been previously reported, such as the distance between *Ovale mediale*–*Mastoidale* landmarks. Moreover, all other distance traits exhibiting correct classification higher than 85% concern distances between landmarks on opposite hemispheres

of the cranium but not symmetric (e.g., bilateral landmarks). Our results show that the most sexually dimorphic distances correspond to non-symmetric traits which are not usually evaluated by other researchers [20, 21, 29]. The same pattern of favoring non-symmetrical traits over symmetric measurements is also present for the traits concerning angles, which although they exhibit lower correct sex classification rates than distance traits, they still produce better results than symmetrical angles often used on other research [30].

The application of iterative analysis of all possible distance and angle combinations based on the battery of landmarks used in the present study not only revealed different sex estimation traits with higher correct classification rates than those commonly used for producing discriminant functions for various populations [20, 21, 29–34], but also produced better classification results than previous work on the same population sample using GPA–PCA workflow [15–17]. More specifically, Franklin et al. [20] reported 85% correct classification for bizygomatic breadth as the most accurate single variable and 90% for multivariate classification using three variables (namely, glabello-occipital length, bizygomatic breadth, and mastoid height) in a study based on 3D volume rendered multi-detector computed tomography scans of 400 adult individuals. Mahakkanukrauh et al. [21] working on a sample of 200 Thai skulls reported maximum accuracy of 90.6% when combining six measurements (maximum cranial length, bizygomatic breadth, biauricular breadth, nasal height, biorbital breadth, and right mastoid length). Marinescu et al.

Table 5 Regression analysis between sex discriminating variables and centroid size

Dist. and Angl.	R square	F	sig.	Dist. and Angl.	R square	F	sig.
D1	0.468	145.282	<.001	D11	0.740	474.398	<.001
D2	0.776	570.744	<.001	D12	0.471	146.907	<.001
D3	0.723	430.008	<.001	D13	0.764	537.645	<.001
D4	0.749	492.157	<.001	A1	0.202	41.737	<.001
D5	0.748	488.675	<.001	A2	0.220	46.633	<.001
D6	0.706	396.801	<.001	A3	0.168	33.221	<.001
D7	0.765	540.661	<.001	A4	0.271	61.453	<.001
D8	0.805	681.596	<.001	A5	0.219	46.161	<.001
D9	0.745	481.415	<.001	A6	0.163	32.039	<.001
D10	0.738	469.595	<.001	A7	0.176	35.179	<.001

[29] produced multivariate discriminant functions achieving 88% correct classification for a Romanian population sample based on four cranial measurements (cranial length, cranial height, facial breadth, and nose height). Similar classification scores have been reported by Ogawa et al. [33] from a Japanese sample. Kranioti et al. [34] working on a modern Cretan population sample, which may be regarded as the most closely related to the Athens Collection since they both comprise from contemporary Greek individuals, reported bizygomatic breadth as the most discriminatory variable with 82% correct classification whereas a stepwise method involving five measurements (bizygomatic breadth, cranial length, nasion–prosthion and mastoid heights, and nasal breadth) raised the classification accuracy to 88.2%.

Regarding the sample of the present study, the research of Chovalopoulou et al. reported cross-validated correct classification of 74.8% for the palate region (five landmarks) and 90.4% for the cranial base (25 landmarks) [15], original group correct classification of 79.4% for the midsagittal curve of the neurocranium (32, 2 landmarks and 30 semi-landmarks) and 89.2% for the cranial vault (31 landmarks) [16] as well as

86.7% for the upper face (31 landmarks) and 83% for the orbits (10 landmarks) [17]. In the present study, although the maximum cross-validated correct classification of single distance variable (2 landmarks) reaches 87.4%, the multivariate analysis produced a discriminant function with only three traits (calculated from only seven landmarks), namely two distances and one angle, with a total cross-validated correct classification of 95.2%.

The direct comparison of our results with previous work on the same population sample not only shows better classification results with easier applicability on forensic cases for sex identification, but also illustrates the advantages of the methodological approach of the present study, where an exhaustive iterative approach (although computationally intensive) was selected over the GPA–PCA analysis often used in landmark-based analysis [9, 10, 15–17]. Of course, using discriminant analysis for landmark-based measurements is a long established method for sex identification [e.g., [35–37]], which is still adopted in recent research [20, 21, 29–34]. Nevertheless, the use of “standard” craniometric measurements or their variants poses limitations on the possible

Table 6 Technical error measurement between digitized landmarks and manual measurements

Dist. and Angl.	TEM ^a	Scaled TEM ^b	Dist. and Angl.	TEM ^a	Scaled TEM ^b
D1	0.90	1.8	D11	1.12	0.9
D2	1.15	0.8	D12	0.69	1.4
D3	2.49	1.8	D13	1.84	1.4
D4	3.51	2.3	A1	1.1	19.1
D5	1.25	1.0	A2	0.8	3.5
D6	2.18	1.6	A3	0.6	13.2
D7	2.75	2.1	A4	2.5	4.0
D8	1.32	0.9	A5	2.2	3.2
D9	2.32	1.8	A6	1.0	1.4
D10	1.28	1.0	A7	0.6	10.4

^a Distances measured in mm and angles in degrees

^b Values are shown in percentage %

findings for each population being examined, which the present study tries to overcome by “ignoring” the prior knowledge of cranial sexual dimorphism expression. The benefits of non-standard interlandmark distances (ILDs) over standard ILDs as well as GM methods have also been reported by Spradley and Jantz [38]. In their work on ancestry estimation of American Black and White and Hispanic males and females, they have demonstrated that nonstandard ILDs provide the best classifications for the groups used in their analyses.

Although our correct classification results are mostly higher than those reported for other populations regarding craniometric traits [20, 21, 29–34], direct comparison of our results with those from different populations does not necessarily reflect higher expression of sexual dimorphism in the Greek population, since different measurements are concerned. However, the highest correct classification discriminant functions reported in the present paper do not necessarily outperform other craniometric discriminant functions in other populations, although they definitely outperform them in the Greek sample, since the extended list of landmarks in the present study includes virtually all landmarks used in other studies and all possible combinations have been assessed. Nevertheless, it is valuable to compare our results with the recent findings of Oikonomopoulou et al. [7], who examined the accuracy of existing binary logistic equations for sex prediction based on cranial traits in the same population sample with our study. Oikonomopoulou et al. used the cranial traits of the mental eminence, supra-orbital margin, glabella, nuchal area, and mastoid process as described by Walker [2] and further included the vertical femoral head diameter (FHD) for further improving the classification results of their newly proposed functions based on the Athens collection. According to their findings, Walker’s equations produced variable degree of success with correct classification ranging from 76 to 99% for males, but only from 22 to 61% for females. Additionally, applying the discriminant functions from Soficaru et al. [39] produced even poorer results (males: 66–98%; females: 26–77%) [7]. However, the high sex bias reported by Oikonomopoulou et al. [7] shows rather poor overall sex estimation for these discriminant functions. On the other hand, the newly proposed equations by Oikonomopoulou et al. [7] based on the Greek assemblage yielded correct classification accuracy ranging from 87.4 to 88.4% for males and 91.4 to 92.6% for females, which are almost as good as the discriminant functions reported in the present study. The inclusion of FHD in their functions further increased the achieved accuracy in the range of 92.6–94.7%, but the gain in classification accuracy is very small for the requirement of the presence of the femoral bone belonging to the same individual as the cranium.

Usually, size produces the most prominent variation between males and females, whether it concerns the qualitative assessment of the mastoid process or the quantitative assessment of the maximum cranial length. Previous

research by Chovalopoulou et al. [15–17] has shown that cranial sex discrimination based on form variation, which includes both shape and size variation, always produces better results than shape variation. Additionally, Oikonomopoulou et al. have shown that using a size related variable as a proxy in sex estimation further improves the classification accuracy of discriminant functions [7]. Nevertheless, using size as a proxy renders the discriminant functions even more population specific, since size-based variables are more influenced by secular trend and are usually population specific [40]. Evaluating the correlation of our distance and angle variables with the centroid size of the crania used in the present study aims to illustrate how much of each sex discriminating variable is explained from the variation in size. Of course, this should not be regarded as a population specificity metric of our discriminant functions. Nevertheless, it provides some insight into how the expression of sexual dimorphism is influenced by the observed size variation. Further research on different population samples is required to evaluate the performance of the proposed discriminant functions for sex assessment and to establish how this performance is interrelated with the observed size variation.

Conclusions

Iteratively exploring nonstandard ILDs as well as interlandmark angles revealed a number of sex-discriminating single traits with reasonably high correct classification, which can be valuable when examining partial cranial remains. It also produced one of the most highly accurate sex-discriminating functions reported in literature involving only two distances and one angle, which also has been demonstrated that can be measured accurately with a simple caliper relaxing the need for use of specialized equipment. The results of our study further emphasizes the need for revising the standard data collection protocol as previously suggested by Spradley and Jantz. Although the use of equipment such as microscribe and landmark based GM methods has become increasingly popular in forensic anthropology, the use of discriminant functions based on simple measurements remains very important since it relaxes the time constraints and need for elaborate setups by the forensic examiner.

Acknowledgements The authors would like to thank the two anonymous reviewers for their comments, which not only improved the quality of the present paper but also helped us further explore certain aspects of our work resulting in improved classification results. The present study and the use of the Athens Collection have been approved by the Department of Animal and Human Physiology. Preliminary results of this study have been communicated (Oral Presentation) at the 3rd Iberian Symposium on Geometric Morphometrics in Girona, Spain. The present study did not receive any funding.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

- Phenice TW (1969) A newly developed visual method of sexing the os pubis. *Am J Phys Anthropol* 30(2):297–301. <https://doi.org/10.1002/ajpa.1330300214>
- Walker PL (2008) Sexing skulls using discriminant function analysis of visually assessed traits. *Am J Phys Anthropol* 136(1):39–50. <https://doi.org/10.1002/ajpa.20776>
- Spradley MK, Jantz RL (2011) Sex estimation in forensic anthropology: skull versus postcranial elements. *J Forensic Sci* 56(2): 289–296. <https://doi.org/10.1111/j.1556-4029.2010.01635.x>
- Harrison DL (2014) An evaluation of the methods used in the estimation of sex. Dissertation, University College London
- Krishan K, Chatterjee PM, Kanchan T, Kaur S, Baryah N, Singh R (2016) A review of sex estimation techniques during examination of skeletal remains in forensic anthropology casework. *Forensic Sci Int* 261:165.e1–165.e8. <https://doi.org/10.1016/j.forsciint.2016.02.007>
- Wilczak CA, Mariotti V, Pany-Kucera D, Villotte S, Henderson CY (2017) Training and interobserver reliability in qualitative scoring of skeletal samples. *J Archaeol Sci Rep* 11:69–79. <https://doi.org/10.1016/j.jasrep.2016.11.033>
- Oikonomopoulou E, Valakos E, Nikita E (2017) Population-specificity of sexual dimorphism in cranial and pelvic traits: evaluation of existing and proposal of new functions for sex assessment in a Greek assemblage. *Int J Legal Med* 131(6):1731–1738. <https://doi.org/10.1007/s00414-017-1655-x>
- Gómez-Valdés JA, Quinto-Sánchez M, Menéndez Garmendia A, Velemínska J, Sánchez-Mejorada G, Bruzek J (2012) Comparison of methods to determine sex by evaluating the greater sciatic notch: visual, angular and geometric morphometrics. *Forensic Sci Int* 221(1–3):156.e1–156.e7. <https://doi.org/10.1016/j.forsciint.2012.04.027>
- Bigoni L, Velemínská J, Brůžek J (2010) Three-dimensional geometric morphometric analysis of cranio-facial sexual dimorphism in a central European sample of known sex. *HOMO - Journal of Comparative Human Biology* 61(1):16–32. <https://doi.org/10.1016/j.jchb.2009.09.004>
- Gonzalez PN, Bernal V, Perez SI (2009) Geometric morphometric approach to sex estimation of human pelvis. *Forensic Sci Int* 189(1–3):68–74. <https://doi.org/10.1016/j.forsciint.2009.04.012>
- Chovalopoulou M-E, Papageorgopoulou C, Bertsatos A (2016) Cranium asymmetry in a modern Greek population sample of known age and sex. *Int J Legal Med* 131(3):803–812. <https://doi.org/10.1007/s00414-016-1430-4>
- Chovalopoulou M-E, Bertsatos A, Papageorgopoulou C (2016) Age-related changes in the craniofacial region in a modern Greek population sample of known age and sex. *Int J Legal Med* 131(4): 1103–1111. <https://doi.org/10.1007/s00414-016-1470-9>
- Manthey L, Jantz RL, Bohnert M, Jellinghaus K (2016) Secular change of sexually dimorphic cranial variables in euro-Americans and Germans. *Int J Legal Med* 131(4):1113–1118. <https://doi.org/10.1007/s00414-016-1469-2>
- Walker PL (2005) Greater sciatic notch morphology: sex, age, and population differences. *Am J Phys Anthropol* 127(4):385–391. <https://doi.org/10.1002/ajpa.10422>
- Chovalopoulou M-E, Valakos ED, Manolis SK (2013) Sex determination by three-dimensional geometric morphometrics of the palate and cranial base. *Anthropol Anz* 70(4):407–425. <https://doi.org/10.1127/0003-5548/2013/0363>
- Chovalopoulou M-E, Valakos ED, Manolis SK (2016) Sex determination by three-dimensional geometric morphometrics of the vault and midsagittal curve of the neurocranium in a modern Greek population sample. *HOMO - Journal of Comparative Human Biology* 67(3):173–187. <https://doi.org/10.1016/j.jchb.2015.09.007>
- Chovalopoulou M-E, Valakos ED, Manolis SK (2016) Sex determination by three-dimensional geometric morphometrics of craniofacial form. *Anthropol Anz* 73(3):195–206. <https://doi.org/10.1127/anthranz/2016/0470>
- Chovalopoulou M-E, Bertsatos A, Manolis SK (2017) Landmark based sex discrimination on the crania of archaeological Greek populations. A comparative study based on the cranial sexual dimorphism of a modern Greek population. *Journal of Mediterranean archaeology and Archaeometry* 17(1):37–49. <https://doi.org/10.5281/zenodo.258084>
- Langley NR, Dudzik B, Cloutier A (2017) A decision tree for non-metric sex assessment from the skull. *J Forensic Sci* 63(1):31–37. <https://doi.org/10.1111/1556-4029.13534>
- Franklin D, Cardini A, Flavel A, Kuliukas A (2013) Estimation of sex from cranial measurements in a western Australian population. *Forensic Sci Int* 229(1–3):158.e1–158.e8. <https://doi.org/10.1016/j.forsciint.2013.03.005>
- Mahakkanukrauh P, Sinthubua A, Prasitwattanaseree S, Ruengdit S, Singsuwan P, Praneatpolgrang S, Duangto P (2015) Craniometric study for sex determination in a Thai population. *Anatomy & Cell Biology* 48(4):275–283. <https://doi.org/10.5115/acb.2015.48.4.275>
- Eliopoulos C, Lagia A, Manolis SK (2007) A modern, documented human skeletal collection from Greece. *HOMO - Journal of Comparative Human Biology* 58(3):221–228. <https://doi.org/10.1016/j.jchb.2006.10.003>
- White TD, Black MT, Folkens PA (2012) *Human Osteology*. Elsevier Academic Press, Amsterdam
- Bookstein FL (1991) *Morphometric tools for landmark data: geometry and biology*. Cambridge University Press, New York
- Eaton JW, Bateman D, Hauberg S, Wehbring R (2015) GNU octave version 4.0.0 manual: a high-level interactive language for numerical computations. <http://www.gnu.org/software/octave/doc/interpreter/> accessed 20 august 2017
- Lewis SJ (1999) Quantifying measurement error. In: Anderson S (ed) *Current and recent research in osteoarchaeology 2: proceedings of the 4th, 5th and 6th meetings of the Osteoarchaeological research group*. Oxbow Books, Oxford, pp 54–55
- Coxeter HSM, Greitzer SL (1967) *Geometry revisited*. The mathematical Association of America, Washington, DC
- Ulijaszek S, Kerr DA (1999) Anthropometric measurement error and the assessment of nutritional status. *Br J Nutr* 82(03):165–177. <https://doi.org/10.1017/S0007114599001348>
- Marinescu M, Panaitescu V, Rosu M, Maru N, Punga A (2014) Sexual dimorphism of crania in a Romanian population: discriminant function analysis approach for sex estimation. *Romanian J Legal Med* 22(1). <https://doi.org/10.4323/rjlm.2014.21>
- Jain D, Jasuja O, Nath S (2013) Sex determination of human crania using mastoid triangle and Opisthion–Bimastoid triangle. *J Forensic Legal Med* 20(4):255–259. <https://doi.org/10.1016/j.jflm.2012.09.020>
- Fortes de Oliveira O, Lima Ribeiro Tinoco R, Daruge Junior E, Sayuri A, Terada SD, Henrique R, da Silva A, Paranhos LR (2012) Sexual dimorphism in Brazilian human skulls: discriminant function analysis. *Journal of Forensic Odontostomatology* 30(2): 26–33
- Ibrahim A, Alias A, Nor FM, Swarhib M, Abu Bakar SN, Das S (2017) Study of sexual dimorphism of Malaysian crania: an

- important step in identification of the skeletal remains. *Anatomy & Cell Biology* 50(2):86–92. <https://doi.org/10.5115/acb.2017.50.2.86>
33. Ogawa Y, Imaizumi K, Miyasaka S, Yoshino M (2013) Discriminant functions for sex estimation of modern Japanese skulls. *J Forensic Legal Med* 20(4):234–238. <https://doi.org/10.1016/j.jflm.2012.09.023>
 34. Kranioti EF, Iscan MY, Michalodimitrakis M (2008) Craniometric analysis of the modern Cretan population. *Forensic Sci Int* 180(2–3):110.e1–110.e5. <https://doi.org/10.1016/j.forsciint.2008.06.018>
 35. Hanihara K (1959) Sex diagnosis of Japanese skulls and scapulae. *The Journal of Anthropological Society of Nippon* 67(4):191–197. <https://doi.org/10.1537/ase1911.67.191>
 36. Hanihara K, Kimura K, Minamidate T (1964) The sexing of Japanese skeleton by means of discriminant function. *Nihon Hoigaku Zasshi* 18(2):107–114
 37. Tanaka T, Hanihara K, Koizumi K (1979) Sex determination of the modern Japanese skull by means of discriminant function. *Sapporo medical. Journal* 48:582–593
 38. Spradley MK, Jantz RL (2016) Ancestry estimation in forensic anthropology: geometric morphometric versus standard and non-standard Interlandmark distances. *J Forensic Sci* 61(4):892–897. <https://doi.org/10.1111/1556-4029.13081>
 39. Soficar A, Constantinescu M, Culea M, Ionică C (2014) Evaluation of discriminant functions for sexing skulls from visually assessed traits applied in the Rainer osteological collection (Bucharest, Romania). *Homo* 65(6):464–475. <https://doi.org/10.1016/j.jchb.2014.08.004>
 40. Kotěrová A, Velemínská J, Dupej J, Brzobohatá H, Pilný A, Brůžek J (2016) Disregarding population specificity: its influence on the sex assessment methods from the tibia. *Int J Legal Med* 131(1): 251–261. <https://doi.org/10.1007/s00414-016-1413-5>

Morphological variation of the femoral head fovea capitis

Andreas Bertsatos, Maria-Eleni Chovalopoulou, Klitemnistra Giannaki, Efstratios Valakos

Department of Animal and Human Physiology, Faculty of Biology, School of Sciences, National and Kapodistrian University of Athens, Panepistimiopolis, GR 157 81 Athens, Greece

SUMMARY

The *fovea capitis femoris* receives the distal attachment of the *ligamentum teres femoris*. Although recent research has shown that the latter may have a number of functions including mechanical stability to the hip joint, there is little published information on the morphological variation of the *fovea capitis femoris*. The present study investigates the morphological variation of the *fovea capitis femoris* with respect to sex and age. Morphometric properties were recorded from both left and right femurs of 212 individuals from the Athens skeletal collection. The *fovea capitis femoris* was photographed *en face* with a reference scale and a polyline outlining its boundary edges was extracted. Two shape variables and three size variables of the *fovea capitis femoris* were calculated and used in the morphological analysis. Two variables, one size and one shape variable, exhibited bilateral asymmetry. The sexual dimorphism of *fovea capitis femoris* is attributed to size variables, while at the same time there are age-related changes in its shape. The *fovea capitis* area and the *fovea capitis* maximum diameter have significant higher values in males, while the perimeter of *fovea capitis* tends to have a more irregular shape in older individuals. However, *fovea capitis femoris* cannot be used for age estimation or sex determination of a human skeleton.

Key words: Fovea capitis – Morphometric properties – Size and shape variation – Sex and age related variation – Age estimation – Sex determination

Corresponding author: Andreas Bertsatos. Dept. of Animal and Human Physiology, National and Kapodistrian University of Athens, Panepistimiopolis, GR 157 81 Athens, Greece.
E-mail: abertsatos@biol.uoa.gr

INTRODUCTION

The main function of the *ligamentum teres femoris* is to provide a functional blood supply to the femoral head (Brewster, 1991; Gray and Villar, 1997) in the fetus. Although it remains unclear whether this role is retained through into adulthood (Chatha and Arora, 2005; Byrd, 2013; Chang and Huang, 2013), recent studies have suggested that the *ligamentum teres* may contribute to stability of the hip joint (Bardakos and Villar, 2009; Cerezal et al., 2010). The *ligamentum teres* is reported to be tensioned in combined flexion–adduction–external rotation of the hip joint (Gray and Villar, 1997; Bardakos and Villar, 2009; Chung, 1976). It extends from the edges of the acetabular notch, to the fovea capitis on the head of the femur (Cerezal et al., 2010; Crelin, 1976).

The *fovea capitis femoris* (FCF) is located slightly posterior and inferior to the center of the articular surface of the femoral head (Michaels and Matles, 1970; Wenger et al., 2007). In contrast to other ligament attachment sites, the FCF is represented as focal depression and is typically slightly ovoid, with an oblique, superior-to-posteroinferior orientation (Bardakos and Villar, 2009). The receptacle zone of the FCF, the smooth lower half, is said to help accommodate the fibers of the *ligamentum teres* when it is tensed (Bardakos and Villar, 2009; Rao et al., 2001), which is thought to contribute to its shape (Kapandji, 2007). Although in adults the *ligamentum teres* is occasionally absent on one or both sides (Tan and Wong, 1990), the FCF is a consistent entity (Tan and Wong, 1990; Cerezal et al., 2012). Even in cases of congenital absence of the *ligamentum teres*, the FCF is present although hypoplastic (Cerezal et al., 2010).

The FCF is used as a marker in various studies

Submitted: 15 April, 2018. Accepted: 19 June, 2018.

and its morphology has been evaluated for several anthropological enquiries. When measuring variations of the acetabular orientation through image analysis, the FCF is a standard landmark (Beltran et al., 2012). Additionally, the FCF serves as a marker for evaluating dysplastic hips (Siebenrock et al., 2013). From an anthropological perspective, the FCF's morphology, apart from being used for differentiation of long bone fragments from each other even in juveniles (Scheuer et al., 2004; Bass, 2005), recently has also been evaluated for sex determination of an individual (Murton et al., 2015). However, there has been very little research reported on the anatomy of the FCF (Perumal et al., 2017; Acar et al., 2017).

The present study aims to contribute to the existing literature by evaluating whether the *fovea capitis femoris* can be used for age-at-death estimation or sex assessment of a human skeleton. More specifically, the present study evaluates the variation in size and shape of the FCF with respect to age and sex using dry femora.

MATERIALS AND METHODS

For the purpose of this research 212 individuals (120 males and 92 females) were studied from the Athens Collection. The Athens Collection is housed in the Department of Animal and Human Physio-

logy of the University of Athens. All individuals were acquired from cemeteries in the Athens area, are Greek nationals, with the vast majority born in Greece, and their respective years of birth span from 1879 to 1965. Information on the name and age at death of each individual in the collection is derived from death records (Eliopoulos et al., 2007). Adult individuals without any evidence of pathology or morphological deformities in femoral head were selected.

Data acquisition

Morphometric properties were recorded from both left and right femora. A Mitutoyo 150 mm dial caliper (graduation 0.02 mm) was used for the measurement of the femoral head's maximum diameter (FHMD) and a Canon EOS 100D camera with Canon EF 40 mm f / 2.8 STM fixed focal length lens for image acquisition of the *fovea capitis femoris* (FCF). All FCF images were shot from a stationary position with a vertical camera view perpendicular to the reference scale at a fixed distance of 35 cm. All femoral samples were manually positioned with the FCF surface facing upwards and at the same plane with the reference scale as shown in Fig. 1. The 2D coordinates of a polyline outlining the boundary edges of the FCF were extracted using AutoCAD software. In order to evaluate the reproducibility of the digitization method, 10 femoral samples were positioned and photographed twice and polylines were extracted from both sets by the same observer. Additionally, polyline extraction by two different observers was further performed on 10 FCF images. Nevertheless, the sample of extracted polyline 2D coordinates used in the analysis of FCF's morphological variation was digitized by the same observer.

In total, 5 variables (2 shape variables and 3 size variables) were used in order to investigate the variation of the FCF. More specifically, the 3 size variables used in our study are the *fovea capitis* area (FCA) and the maximum diameter (FCMD) of the FCF, as well as the ratio of the *fovea capitis* area to the femoral head surface (FCA/FHS). The FCA/FHS ratio is expressed in percentage and the femoral head surface (FHS) was approximated as 74% of the sphere (Ruff, 1988) calculated from the FHMD as $4\pi r^2$. Shape variables comprise the following two ratios:

{ArPerIndex $\in R:0 < x \leq 1$ }: the ratio of the area to the squared perimeter of the *fovea capitis* multiplied by 4π so that ArPerIndex reaches maximum when *fovea capitis* is perfectly circular.

{SMA $\in R:0 < x \leq 1$ }: the ratio of the minimum (I_{MIN}) to the maximum (I_{MAX}) principal component of second moments of area of the FCF, which maximizes when *fovea capitis* area is evenly distributed across principal axes and as close as possible around its centroid.

The FCA, its perimeter, the FCMD, as well as the I_{MAX} and I_{MIN} were calculated with GNU Octave from the coordinates of the recorded polylines (Fig. 1). The three ratios were also calculated with GNU Octave. Regarding intra- and inter-observer error,

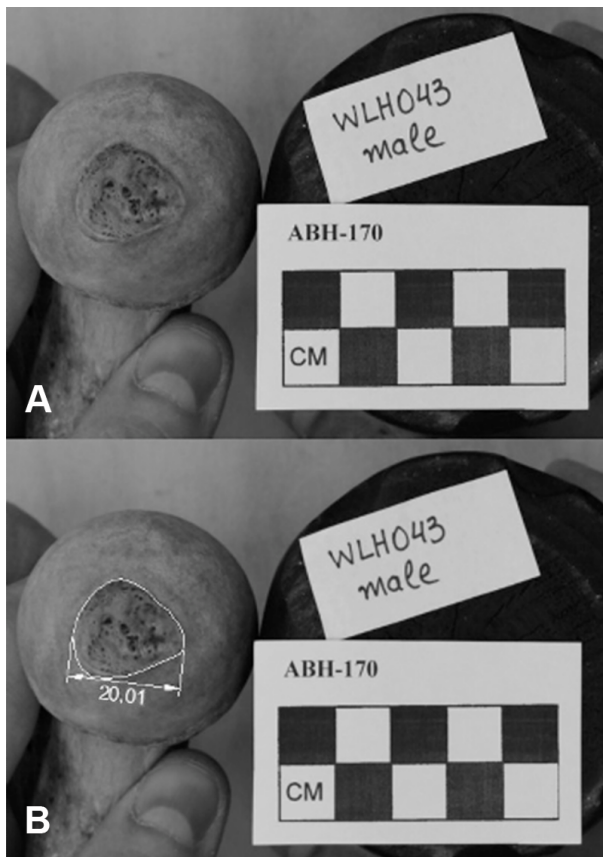


Fig 1. A: The fovea capitis femoris (FCF) photographed in face with a reference scale. **B:** The fovea capitis maximum diameter and the polyline outlining the boundary edges of the FCF.

the mean absolute difference (MAD) between repeated digitized polylines was calculated for the FCA and FCMD variables.

Statistical analysis

Prior to any statistical analysis, a Kolmogorov-Smirnov test was conducted in order to determine whether parametric or non-parametric test will be used. Initially, paired sample t-test or Wilcoxon test (non-parametric), depending on the results of the normality test, was performed in order to test for bilateral asymmetry and whether the left and right sides should be analyzed separately. Independent sample t-test or Mann-Whitney test (non-parametric) was used in order to evaluate the presence of sexual dimorphism in the FCF. Additionally, the sample was divided in 3 age group categories: 1) 20-39 years old, 2) 40-59 years old and 3) ≥ 60 years old and tested for age related changes through either ANOVA or Kruskal-Wallis H test (non-parametric). The specific age ranges were chosen for similar size of each subgroup based on the availability of our sample, while they are still representative of young-aged, middle-aged, and old-aged groups of individuals. In cases where statistically significant difference was detected, Dunn's Test, with Bonferroni adjusted p values, was used to pinpoint which specific groups differ significantly from the others. Finally, the appropriateness of the *fovea capitis femoris* as a sex and age estimator was evaluated through Discriminant Function Analysis (DFA). SPSS software (PASW Statistics 23.0) was used to perform all statistical analyses.

RESULTS

Regarding the reproducibility of the digitization method in the present study, the MAD concerning

intra-observer error was found approximately 0.05 mm for FCMD and 2 mm² for FCA, whereas for inter-observer error the MAD values were 0.08 mm and 2.5 mm², respectively. Considering the overall observed variability of the FCMD and FCA variables in our sample, as shown in table 1, the intra- and inter-observer errors may be regarded as negligible.

All 5 FCF variables, for male and female groups separately, were tested for bilateral asymmetry. According to the Kolmogorov-Smirnov test, the ArPerIndex variable of the left side both for males and females failed the normality test (p-values: <0.001 and 0.019 for males and females respectively). The SMA variable of the left side for females also failed the normality test (p-values: 0.034). Consequently, the Wilcoxon test was performed for the ArPerIndex variable for both sexes, as well as the SMA variable for the female group, while in all other cases the paired sample t-test was performed. The FCA/FHS and SMA variables exhibited bilateral asymmetry in females. Henceforth, the left and right sides were analyzed separately for these variables, whereas the average values of the left and right side of the FCA, FCMD and ArperIndex variables were used for the subsequent analyses.

Table 1 presents the descriptive statistics for all 5 FCF variables analyzed with respect to sex and age. In most of the size variables males exhibit higher values than females, the exception concerns the ratio of the *fovea capitis* area to the femoral head surface. In case of the shape variables the reverse applies. However, regarding the FCA/FHS variable as well as the shape variables, the differences between sexes are small, when considering the confidence intervals for their mean values. These observations are confirmed by the evaluation of the presence of sexual dimorphism.

Table 1. Summary statistics for the femoral head fovea capitis variables. FCA in mm²; FCMD in mm.

		N	Mean	Std. Deviation	95% Confidence Interval for Mean		Kolmogorov-Smirnov
					Lower Bound	Upper Bound	Sig.
FCA	Males	120	219.939	64.738	207.410	232.468	.200
	Females	92	177.565	48.299	166.952	188.177	.181
FCMD	Males	120	19.959	3.034	19.372	20.546	.200
	Females	92	17.874	2.737	17.273	18.475	.200
ArPerIndex	Males	120	0.873	0.043	0.864	0.881	.013
	Females	92	0.880	0.039	0.871	0.889	.200
SMA (Left side)	Males	120	0.588	0.133	0.562	0.614	.200
	Females	92	0.608	0.155	0.574	0.642	.034
SMA (Right side)	Males	120	0.564	0.142	0.537	0.592	.200
	Females	92	0.571	0.134	0.541	0.601	.200
FCA/FHS (Left side)	Males	120	4.336	1.241	4.096	4.576	.200
	Females	92	4.498	1.168	4.242	4.755	.092
FCA/FHS (Right side)	Males	120	4.207	1.268	3.962	4.453	.200
	Females	92	4.287	1.255	4.011	4.563	.200

FCA: fovea capitis area; FCMD: fovea capitis maximum diameter; ArPerIndex: ratio of the area to the squared perimeter of the fovea capitis multiplied by 4π ; SMA: ratio of the I_{MIN} to the I_{MAX} principal components of second moments of area of the FCF; FCA/FHS: ratio of the fovea capitis area to the femoral head surface expressed in percentage.

Table 2 presents the independent sample t-test results regarding the presence of sexual dimorphism in the *fovea capitis femoris* variables. In males the ArPerIndex variable and in females the SMA variable of the right side, failed the normality test (K-S sig. Males: 0.013; Females: 0.034). Consequently, a Mann-Whitney test was performed for these variables.

According to the results, only the size variables FCA and FCMD are sex dimorphic. Applying discriminant function analysis on the FCA and FCMD variables resulted in significant models for both variables (Wilk's lambda: 0.888 and 0.884 respectively; $p < 0.001$ for both). The cross-validated classification accuracy ranges from 62.7% to 65.6% for the FCA and FCMD variables respectively (Table 3).

Regarding the FCF's size and shape changes related to aging process, the FCA and the FCMD were evaluated separately for males and females since they exhibit sexual dimorphism, while the SMA, the FCA/FHS and the ArPerIndex were evaluated for the total sample. Regarding the results of the Kruskal-Wallis H test or the ANOVA, depending on the results of the normality test (Table 4), only the ArPerIndex exhibits statistically significant difference ($p = 0.028$) between age group categories. Finally, Dunn's pairwise tests adjusted by Bonferroni correction revealed statistically significant difference ($p = 0.025$) only between the age groups of «20-39 years old» and «≥60 years old» in the ArPerIndex variable.

DISCUSSION

Very few studies in the literature have focused on

Table 2. Results of Independent sample t-test (A) and Mann-Whitney test (B) regarding the presence of sexual dimorphism in the fovea capitis femoris variables.

A)			
	t	df	Sig.
FCA	5.335	208	<.001
FCMD	5.151	209	<.001
SMA (Right side)	-0.788	196	.431
FCA/FHS (Left side)	-0.969	198	.334
FCA/FHS (Right side)	-0.773	196	.440
Significant difference ($p \leq 0.05$)			
B)			
	Mann-Whitney U		Sig.
ArPerIndex	4778		.114
SMA (Left side)	4696		.589
Significant difference ($p \leq 0.05$)			

FCA: fovea capitis area; FCMD: fovea capitis maximum diameter; ArPerIndex: ratio of the area to the squared perimeter of the fovea capitis multiplied by 4π ; SMA: ratio of the I_{MIN} to the I_{MAX} principal components of second moments of area of the FCF; FCA/FHS: ratio of the fovea capitis area to the femoral head surface expressed in percentage.

the anatomy of the FCF, even though the femur is widely used in anthropology (Bass, 2005). The present study evaluated the sex- and age-related morphological variation of the FCF using a number of variables that accurately represent its size and shape properties and aimed to assess whether the *fovea capitis femoris* can be used for sex and/or age-at-death estimation of a human skeleton. Perumal et al. (2017) analyzed the morphology and morphometry of the FCF on 125 dry adult femurs (61 right and 64 left) from the Anatomy Museum of the University of Otago in New Zealand. They photographed the FCF *en face* with a reference scale and they recorded the maximum and the minimum length (LL: longitudinal length and TL: transverse length respectively), as well as the cross-sectional area (CSA) of the FCF using the ImageJ software. According to Perumal et al.'s research, the LL and CSA mean values of the FCF were 1.8 cm and 1.8 cm² respectively. These findings are similar to the results of our research regarding the FCMD and FCA variables respectively in the female group. Unfortunately, Perumal et al. (2017) were not able to classify the bones based on sex, due to lack of demographic data. Additionally, Perumal et al. (2017) found that 17% of the femoral head was occupied by the FCF with the use of the

Cross – sectional area of the FCF

Cross – sectional area of the head of the femur

ratio. In our study, as determined by the FCA/FHS ratio, only 4.21% and 4.29% of the femoral head is occupied by the FCF in males and females respectively. This disagreement with Perumal et al. (2017) is due to the different approaches used for femoral head surface estimation. Nevertheless, we consider the estimate used in the present study more accurate based on Ruff, who reported it to be approximately 74% of the sphere calculated by the FMHD (Ruff, 1988) as opposed to the cross-sectional area utilized by Perumal et al. (2017).

Sampatchalit et al. in 2009 used 11 cadaveric hip joints (8 males and 3 females) in order to examine the degenerative changes in the *ligamentum teres* of the hip with magnetic resonance arthrography, anatomical inspection and histological evaluation. According to their results, the maximum width of the *fovea capitis* ranges from 12 to 15 mm in males and 13 to 16 mm in females. In our sample, the mean values of the FCMD are higher for both sexes. This disagreement with our study is more likely due to the small sample size used by Sampatchalit et al. (2009) in their research rather than differences between populations.

Murton et al. (2015) used the maximum *fovea capitis* height (MFCH: the distance between the most superior and the most inferior points on the FCF), as well as the maximum *fovea capitis* breadth (MFCB: the distance between the most medial and the most lateral points on the FCF) and calculated the *fovea capitis* index using the formula: $(MFCH \times 100)/MFCB$, in order to establish a new method for sex determination. Measurements

Table 3. Discriminant functions and correct classification results of FCA and FCMD variables.

	N	Constant	Coefficient	Percentage of correct classification %		
				Total ^a	Males ^a	Females ^a
FCA	212	-3.271	0.016	62.7	69.2	54.3
FCMD	212	-6.375	0.333	65.6 / 65.1	75.8	52.2 / 51.1

^aSingle values are given when original group classification equals cross-validated classification, otherwise both values are shown in the form: original group / cross-validated

FCA: fovea capitis area; FCMD: fovea capitis maximum diameter

were taken with a caliper. According to Murton et al.'s results (2015), although the MFCH and the MFCB didn't exhibit sexual dimorphism, when both variables were combined into an index of shape there were significant differences between sexes for both left and right sides. In our study the reverse applies. More specifically, according to our results, shape variables of the FCF are not sexually dimorphic, while the FCMD exhibits sexual dimorphism.

Acar et al. (2017) used x-ray views to record the FCF variant configurations and age-related changes according to the *fovea capitis* index:

$$FCI = \frac{\text{Fovea Capitis Diameter} \times 100}{\text{Femoral Head Diameter}}$$

The *fovea capitis* diameter was measured from its upper margin to its lower margin, while the femoral head diameter from the diameter of a complete circle around the femoral head.

According to Acar et al.'s results, the FCI does not exhibit sexual dimorphism. This observation is consistent with our results, since we did not find statistically significant differences between sexes regarding the FCA/FHS variable, which is equivalent to the Acar's FCI. However, Acar's FCI exhibited bilateral asymmetry in males, whereas in our research statistically significant bilateral asymmetry of the FCA/FHS variable was observed in females. This may be due to the morphological variation between populations.

Regarding the FCF's age-related changes, Acar et al. (2017) found statistically significant differences in FCI between the age group categories «20-39 years old» and «≥60 years old» as well as between the age group categories «40-59 years old» and «≥60 years old». More specifically, according to their results, the *fovea capitis* diameter increases with aging process. In our research, the mean value of the FCMD is almost the same in all age group categories in males, and while in females the FCMD mean value increases with aging process, the differences are not statistically significant. This disagreement with Acar et al.'s study (2017) may be due to the different techniques used to obtain the *fovea capitis* diameter. On the other hand, our research revealed that there are age-related changes in FCF's shape. According to our research, the perimeter of the *fovea capitis* tends to have a more irregular shape in older individuals (ArPerIndex decreases). This could be the result of bone proliferation related to the aging pro-

cess (Boskey and Coleman, 2010), or presumably a combination of aging and activity patterns that may result to the observed morphological variation.

The present study utilizes certain size and shape variables that aim to describe the morphological variability of the FCF in a most representative way possible, while still maintaining an easily reproducible and accurate method. However, it should be emphasized that not all researchers have used the same methods or variables for studying the *fovea*

Table 4. Results of ANOVA (A) and Kruskal Wallis test (B) regarding age related changes of the fovea capitis femoris variables.

A)			
	F	df	Sig.
FCMD (Male group)	0.601	2 / 117	.550
FCMD (Female group)	0.662	2 / 89	.518
SMA (Right side)	0.041	2 / 196	.960
FCA/FHS (Left side)	0.801	2 / 198	.450
Significant difference (p≤0.05)			
Degrees of freedom are shown in the form: Between Groups /			
B)			
	Chi-Square	df	Sig. ^a
FCA (Male group)	0.019	2	.991
FCA (Female group)	0.009	2	.996
SMA (Left side)	5.664	2	.059
FCA/FHS (Right side)	0.300	2	.861
	7.137	2	.028
	Dunn's pair-	Median	Adj. Sig. ^b
ArPerIndex	20-39 / 40-59	0.895	.511
	40-59 / ≥60	0.885	.593
	≥60 / 20-39	0.867	.025

^aSignificant difference (p≤0.05)

^bSignificant difference adjusted using the Bonferroni correction (p≤0.05)

FCA: fovea capitis area; FCMD: fovea capitis maximum diameter; ArPerIndex: ratio of the area to the squared perimeter of the fovea capitis multiplied by 4π; SMA: ratio of the I_{MIN} to the I_{MAX} principal components of second moments of area of the FCF; FCA/FHS: ratio of the fovea capitis area to the femoral head surface expressed in percentage.

capitis, hence the aforementioned comparisons with previous studies is on a more general basis.

CONCLUSIONS

Our findings indicate that the size variables FCA and FCMD are sexually dimorphic, while at the same time the shape variable ArPerIndex exhibits age-related changes. However, as shown by the results of the discriminant function analysis and Dunn's pairwise tests, the magnitude of changes related both to sexual dimorphism and aging process is small, and therefore the *fovea capitis femoris* cannot be used for age-at-death estimation or sex assessment of an unidentified case. Nevertheless, further research including confounding factors such as mechanical stress and activity patterns may shed light to the morphological variation of the FCF, which, to our current understanding, for the most part may be considered idiosyncratic.

REFERENCES

- ACAR N, KARAAARSLAN A, KARAKASLI A, ERDURAN M (2017) Femoral head fovea capitis variant configurations and age related changes- a radiological study. *Iranian J Radiol*, InPress. doi:10.5812/iranjradiol.41130.
- BARDAKOS NV, VILLAR RN (2009) The ligamentum teres of the adult hip. *J Bone Joint Surg*, 91-B(1): 8-15.
- BASS WM (2005) *Human osteology: a laboratory and field manual*. Missouri Archaeological Society, Columbia, p 223.
- BELTRAN LS, MAYO JD, ROSENBERG ZS, TUESTA MD, MARTIN O, NETO LP, BENCARDINO JT (2012) Fovea alta on MR images: is it a marker of hip dysplasia in young adults? *Am J Roentgenol*, 199(4): 879-883.
- BOSKEY AL, COLEMAN R (2010) Aging and bone. *J Dental Res*, 89(12): 1333-1348.
- BREWSTER SF (1991) The development of the ligament of the head of the femur. *Clin Anat*, 4(4): 245-255.
- BYRD JW (2013) *Operative hip arthroscopy*. Springer, New York, p 108.
- CEREZAL L, ARNAIZ J, CANGA A, PIEDRA T, ALTÓNAGA JR, MUNAFO R, PÉREZ-CARRO L (2012) Emerging topics on the hip: Ligamentum teres and hip microinstability. *Eur J Radiol*, 81(12): 3745-3754.
- CEREZAL L, KASSARJIAN A, CANGA A, DOBADO MC, MONTERO JA, LLOPIS E, ROLÓN A, PÉREZ-CARRO L (2010) Anatomy, biomechanics, imaging, and management of ligamentum teres injuries. *Radiographics*, 30: 1637-1651.
- CHANG CY, HUANG AJ (2013) MR imaging of normal hip anatomy. *Magn Reson Imaging Clin N Am*, 21(1): 1-19.
- CHATHA DS, ARORA R (2005) MR imaging of the normal hip. *Magn Reson Imaging Clin N Am*, 13(4): 605-615.
- CHUNG S (1976) The arterial supply of the developing proximal end of the human femur. *J Bone Joint Surg*, 58(7): 961-970.
- CRELIN ES (1976) An experimental study of hip stability in human newborn cadavers. *Yale J Biol Med*, 49: 109-121.
- ELIOPOULOS C, LAGIA A, MANOLIS S (2007) A modern, documented human skeletal collection from Greece. *HOMO - J Comp Human Biol*, 58(3): 221-228.
- GRAY AJ, VILLAR RN (1997) The ligamentum teres of the hip: an arthroscopic classification of its pathology. *Arthroscopy*, 13(5): 575-578.
- KAPANDJI IA (2007) *The physiology of the joints*. Churchill Livingstone/Elsevier, Edinburgh, p 24.
- MICHAELS G, MATLES AL (1970) 21 The role of the ligamentum teres in congenital dislocation of the hip. *Clin Orthop Rel Res*, 71(1): 199-201.
- MURTON N, BORRINI M, ELIOPOULOS C (2015) Sexual dimorphism of the fovea capitis femoris in a medieval population from Gloucester, England. *Global J Anthropol Res*, 2(2): 9-14.
- PERUMAL V, WOODLEY SJ, NICHOLSON HD (2017) The morphology and morphometry of the fovea capitis femoris. *Surg Radiol Anat*, 39(7): 791-798.
- RAO J, ZHOU Y, VILLAR R (2001) Injury to the ligamentum teres. *Clin Sports Med*, 20(4): 791-800.
- RUFF CB (1988) Hindlimb articular surface allometry in hominoidea and Macaca, with comparisons to diaphyseal scaling. *J Human Evol*, 17(7): 687-714.
- SAMPATCHALIT S, BARBOSA D, GENTILI A, HAGHIGHI P, TRUDELL D, RESNICK D (2009) Degenerative changes in the ligamentum teres of the hip. *J Comput Assist Tomogr*, 33(6): 927-933.
- SCHEUER L, BLACK S, CHRISTIE A (2004) *The juvenile skeleton*. Elsevier Academic Press, Amsterdam, p 355.
- SIEBENROCK KA, STEPPACHER SD, ALBERS CE, HAEFELI PC, TANNAST M (2013) Diagnosis and management of developmental dysplasia of the hip from triradiate closure through young adulthood. *J Bone Joint Surg*, 95(8): 749-755.
- TAN CK, WONG WC (1990) Absence of the ligament of head of femur in the human hip joint. *Singapore Med J*, 31: 360-363.
- WENGER D, MIYANJI F, MAHAR A, OKA R (2007) The mechanical properties of the ligamentum teres. *J Pediatr Orthop*, 27(4): 408-410.

Secular change in adult stature of modern Greeks

Andreas Bertsatos  | Maria-Eleni Chovalopoulou 

Department of Animal and Human Physiology, Faculty of Biology, School of Sciences, National and Kapodistrian University of Athens, Panepistimiopolis, Athens, Greece

Correspondence

Maria-Eleni Chovalopoulou, Department of Animal and Human Physiology, Faculty of Biology, School of Sciences, National and Kapodistrian University of Athens, Panepistimiopolis, GR 157 81 Athens, Greece.
Email: marilenachov@biol.uoa.gr

Abstract

Objectives: In Greece, during the late 19th and early 20th century, the rural population adopted a more or less urban lifestyle. Furthermore, the first half of the 20th century finds Greece involved in five major wars, including a civil war, and consequent financial deprivation. This study investigates how the socioeconomic changes in Greece, during this period of time, have affected the stature of its population.

Methods: The Athens collection constitutes our sample with 189 adult individuals (104 males and 85 females). Stature was estimated with regression equations and secular change was evaluated by linear regression of stature with respect to the year of birth. Further analysis of our population sample was based on three time periods to explore the correlation between secular change and historical events.

Results: From 1879 to 1965, stature increased for both males and females. The subsequent analysis among different periods revealed that the male group exhibited a small although non-significant decline in stature during the years 1912 to 1950, which coincides with the long inter-war period. However, females appear less affected by the consequent deprivation.

Conclusions: Our findings are consistent with earlier studies based on a much smaller time span. The negative effect of the economic and nutritional deprivation on stature as a result of warfare is apparent in our sample, at least for males. Furthermore, the positive effect of economic growth on stature is prominent for the entire population in the postwar period.

1 | INTRODUCTION

Documentation of variation in standard of living across time has become a major aim of anthropologists, social historians, and economists, to explain spatial and temporal variation in human health and population well-being. Stature is considered a good index of socioeconomic status (SES) or the biocultural environment of human populations (Komlos, 2009; Schweich & Knüsel, 2003; Steckel, 2008). Although stature is mainly affected by genetic factors (Lango et al., 2010; Wood et al., 2014), it is also influenced by sanitary conditions, nutritional quality, the amount of energy expenditure, and overall health (Keep & Bogin, 1999; Schell, 1986; Susanne, 1985).

Stature data can be obtained either from measurements of living people, or can be calculated by regression equations

derived from the measurements of the lengths of long bones from skeletal remains. Several regression equations for stature estimation have been developed (Ruff et al., 2012; Sjøvold, 1990; Trotter & Gleser, 1952). The stature estimates from different formulas differ (Waldron, 1998) because the various reference samples from which the equations were derived differ in relative proportions of their limbs. Therefore, researchers have to ensure consistency of their results when examining changes in stature over time. The incomparability of stature estimates over time can be avoided either by comparing absolute long bone lengths, or by applying the same formula.

Over the last century, countries have shown various patterns of secular change in stature depending on their socioeconomic and biocultural environment. For instance, in England, the United States, Sweden, France, and Australia,

average stature began to increase with industrialization in the 19th century, while in Germany the average stature remained relatively unchanged until the beginning of the 20th century (Steckel, 2001).

Throughout the 19th century, Greece was mostly an agricultural society. Its predominantly agricultural economy was capable of feeding its rapidly rising population, while sustaining a net trade surplus in agricultural and livestock products (Petmezas, 2006). Despite importing a large part of its provisions in subsistence goods (such as wheat and other cereals), these land-intensive food imports were largely outnumbered in value by the labor-intensive Greek agricultural products (Petmezas, 2006). The late 19th and early 20th century characterizes the transition of the rural Greek population in adapting a more or less urban lifestyle following the policy shift of the Greek state toward public investment in urban infrastructure. Nevertheless, nutritional standards for the majority of the population remained almost constant with the rural population, being very frugal in its eating habits (Sapounaki-Dracaki, 1998). Additionally, the first half of the 20th century finds Greece involved in four major wars, namely the Balkan Wars (1912–1913), World War I (1916–1918), the Greco-Turkish War (1919–1922), and World War II (1940–1944), as well as the Greek Civil War (1946–1950) that followed soon after WWII. Furthermore, the socioeconomic consequences of the Great Depression of the 1930s, which affected agricultural production in Greece, also played an important role in the dramatic change of Greek society from a mainly agricultural society at the beginning of the 20th century to a mainly urban one in the second half of the century.

The long inter-war period (1912–1950) greatly hindered the agricultural development of the 19th century, whereas the consequent economic crisis along with the Great Depression of the 1930s resulted in the transatlantic out-migration movement during the early 20th century. However, it should be noted that during the first half of the 20th century (1913–1947) Greece doubled its territory by annexing Macedonia, Epirus, Crete, Samos, and the Northern Aegean islands in 1913, Thrace in 1919, Eastern Thrace in 1923, as well as Dodecanese Islands in 1947. Greece's population also tripled during that period from 2.5 million in 1897 to 7.5 million in 1947 (Petmezas, 2006).

Secular change in the stature of the Greek population during the first half of the 20th century has been studied based on the biometrics of army conscripts (Valaoras, 1970). However, this only accounts for male individuals born in Greece during the period of 1927–1945. The present study uses stature estimates based on the measurements of skeletal remains in order to investigate the secular change of the Greek population during a wider period of time. With a time span from the late 19th until the middle of the 20th century, our goals are to capture the entire transformation period of

Greek society and to investigate how the socioeconomic changes in Greece affected the stature of its population.

2 | MATERIALS AND METHODS

For the purpose of this research, 189 adult individuals (104 males and 85 females) were studied from the Athens Collection. Only individuals without any skeletal deformities that could severely bias the estimated stature were studied. The Athens collection is housed in the Department of Animal and Human Physiology (Faculty of Biology, National & Kapodistrian University of Athens). All individuals were acquired from cemeteries in the Athens area and their respective years of birth span from 1879 to 1965. According to information from death certificates all individuals are of known sex and age (Eliopoulos, Lagia, & Manolis, 2007). All individuals are Greek nationals, with the vast majority born in Greece, some were born abroad but of Greek descent, and there were six cases for whom the birth place is unknown and only Greek ethnicity is provided in their death certificates. More details about the geographical distribution of our sample is provided in Table 1. The mean age at death of our sample is 59.3 years old. Although the registered occupation on the death certificate does not necessarily reflect the actual SES for each individual, it indicates that our sample covers a wide range of SES backgrounds.

Stature estimation was based on femur bones due to preservation limitations of the skeletal material. The estimated stature was cross-validated against anatomical height on 16 individuals among the most well-preserved skeletons of our sample. For the anatomical determination of living stature Raxter, Auerbach, & Ruff, (2006) first equation was used, which also includes the age of the individual. The maximum length of the femur bone was measured according to Martin's method (Buikstra & Ubelaker, 1994; Martin, 1928). Femur bones free of deformities were selected and all measurements were performed using an osteometric board. All measurements were taken on the right side, and only if this was not possible due to damage was the left side used.

For the Greek population, no regression equation for stature estimation has yet been developed. Therefore, we selected the most appropriate one from among those previously developed in other populations. The racial background of the reference population from which the equations were derived was considered. The equations of Sjøvold for Caucasians (1990), of Trotter and Gleser for Whites (1952), and those of Ruff et al. (2012) for all populations regarding the femur were tested. Stature values were calculated based on the femoral maximum length with the corresponding regression equations and to select the most appropriate equation, the absolute mean differences and average differences

TABLE 1 Geographic distribution regarding the birthplace of the Greek population sample

Geographic region of birthplace	Sample size	Percentage of the total sample
Attica	49	26.2%
Central Greece	15	8%
Northern Greece	24	12.8%
Aegean Islands	21	11.2%
Ionian Islands	6	3.2%
Peloponnese	21	11.2%
Crete	10	5.4%
Asia Minor	20	10.7%
Other countries	15	8%
Unknown	6	3.2%

between calculated stature and living height of 16 individuals (8 males and 8 females) were compared.

Secular change was evaluated by linear regression of the stature with respect to the year of birth. Autocorrelation, which can affect parametric significance tests, was examined using the Durbin-Watson test. Normal least-squares regression was used, since “year of birth” was used as the independent variable. Additionally, the sample was divided into three time periods, namely 1879–1911, 1912–1950, and 1951–1965, based on the year of birth of each individual. The aforementioned time periods correspond to the different political and economic situation in Greece based on historical events that influenced the Greek economy. Indeed, the first period corresponds to agricultural stagnation at the turn of the century, whereas the second period corresponds to the long war interval when Greece was involved in a series of wars. The third period signifies the postwar rapid economic growth and modernization of the Greek society. The three groups were analyzed against each other for variance by means of one-way ANOVA tests in order to assess whether the dramatic shifts in the economy have affected the stature

of the Greek population. The selected time division along with the limitations imposed by the overall small sample of the skeletal collection resulted in large differences in the sample sizes of the groups. Hence, Hochberg’s post hoc test was used in ANOVA for testing between group differences.

The aforementioned analyses were performed separately for males and females, since stature exhibits sexual dimorphism in human populations (also valid for our skeletal sample), hence secular change may also follow different patterns between sexes. All statistical analyses regarding linear regression and ANOVA were performed with SPSS statistical software (IBM SPSS version 23.0, Armonk, NY).

3 | RESULTS

The absolute mean differences as well as the average differences between living height and stature values obtained from different stature estimation equations are given in Table 2. The equations of Ruff et al. (2012) exhibited the lowest values. Most importantly, the estimated stature from Ruff’s

TABLE 2 Absolute mean differences and average differences (in cm) between living height and stature values obtained from different stature estimation equations

Sex	N	Ruff et al. equations		Trotter & Gleser equations		Sjøvold equations	
		MAD	AD	MAD	AD	MAD	AD
Females	8	1.8	−0.1	6.3	−6.3	3.7	−3.7
Males	8	1.7	0.6	4.2	−4.2	2.3	−2.3
Pooled sexes	16	1.75	0.25	5.25	−5.25	3.0	−3.0

Abbreviations: MAD, mean absolute difference; AD, average difference.

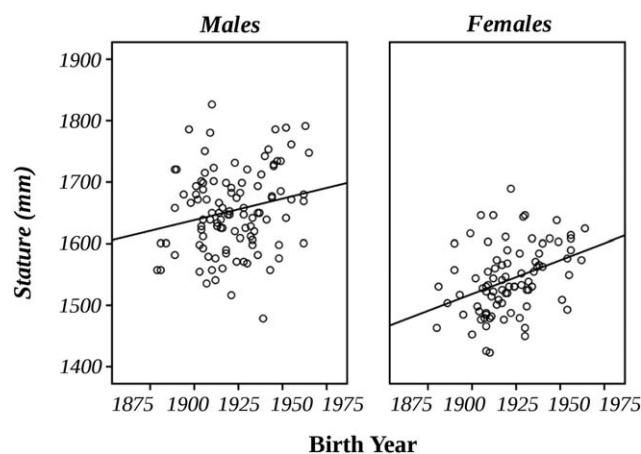


FIGURE 1 Least squares regression lines for male (left) and female (right) stature in a Greek population

regression equations exhibit symmetry over the mean living height, whereas the Trotter and Gleser (1952) as well as Sjøvold (1990) estimators are biased toward overestimating stature. Therefore, we chose Ruff's et al. equations as the most appropriate for our sample.

As shown in Figure 1, during 1879 to 1965, stature increased for both males and females with the latter exhibiting a higher increment. Regression analysis on stature data showed that both male and female stature increased significantly over time (Males regression: $R^2 = 0.044$, $F = 4.658$, $P = .033$; Females regression: $R^2 = 0.139$, $F = 13.370$, $P < .001$; Table 3). The slope of the regression line for males was 0.696 and the corresponding for females was 1.098. However, the regression coefficient for males was not significantly different from females (T value = 0.879, $P = .380$).

The smaller amount of variance explained by the male regression model is in agreement with the results of the F statistic presented in Table 4. The ANOVA results showed statistically significant differences in mean stature between time periods for both males (ANOVA: $F = 3.506$, $P = .034$) and females (ANOVA: $F = 6.872$, $P = .002$). When stature was compared among groups of different time periods using

TABLE 3 Least squares regression results of secular change in stature for both sexes

Regression results	Males	Females
R^2	.044	.139
Durbin-Watson	2.000	1.895
F-Statistic	4.658	13.370
Prob (F-statistic)	.033	.000
Regression coefficient	.696	1.098
Regression slopes comparison	T -value	Prob
	.879	.380

Hochberg's post hoc test, males exhibited a statistically significant difference only between the second and third period, whereas females had a statistically significant increase in mean stature between consecutive time periods. More specifically, as shown in Table 5, during 1879–1911 the stature of females was 151.4 cm, increased to 155.4 cm during 1912–1950, and further increased to 157.1 cm during 1951–1965 (Figure 2). On the contrary, males exhibited a minor decrease in mean stature between the first two periods from 165.7 to 164.4 cm, although it increased significantly in the third period reaching 170.6 cm.

4 | DISCUSSION

This study is based on the assumption that no substantial genetic change affecting stature took place during the time period of interest. This assumption is largely based on the historical evidence concerning migration patterns in the Greek peninsula. Since the formation of the Greek state, two major waves of emigration took place (Kasimis & Kassimi, 2004). The first mass emigration was triggered by the rapid fall in the price of currants and the consequent economic crisis of 1893, while the second wave took place after World War II. The only in-migration during the first half of the 20th century concerns the repatriation of native Asia Minor Greeks, who either fled from Asia Minor during the Greco-Turkish war or were forcibly repatriated soon after, during the population exchange based on the convention signed at Lausanne, Switzerland, on 30 January, 1923. However, they may not be considered a foreign population, since they are of Greek descent and both culturally and biologically may be considered Greeks. Despite a positive migration balance in the 1970s mainly due to return migration, it is not until the 1980s that foreign immigrants started arriving in Greece. Thus, it is safe to assume that the Greek population largely remained genetically isolated during the period between 1879 and 1965, which our skeletal sample accounts for. Consequently, we assume that any secular trend observed in stature during this period should be a result of socioeconomic factors.

Another assumption of this study is that there are no changes in the relationship between femoral length and stature of the Greek population during the examined period, since our findings based on the maximum femoral length in reality reflect secular change in bone length. However, such an assumption cannot be tested with the available sample. Nevertheless, we decided to abide by this assumption so we can make direct comparisons with previous studies based on living heights of the Greeks. A number of authors during the 20th century have published biometric records of stature based on measurements from living people from various parts of Greece. A comprehensive list of their findings is

TABLE 4 One-Way ANOVA results: Comparison of male and female stature between time-periods

Between groups	Sum of squares	df	Mean square	F	Sig.	Levene's test sig.
ANOVA results						
Males	30688.58	2	15344.29	3.506	.034	.426
Females	37220.76	2	18610.38	6.872	.002	.881
				<u>95% Confidence interval</u>		
Comparisons	Mean difference (mm)		Std. error	Sig.	Lower bound	Upper bound
Male's Comparison for each time period						
1879–1911	1912–1950	13.02	13.99	.729	–20.94	46.97
	1951–1965	–49.12	24.66	.139	–108.95	10.71
1912–1950	1879–1911	–13.02	13.99	.729	–46.97	20.93
	1951–1965	–62.13 ^a	23.68	.030	–119.58	–4.69
1951–1965	1879–1911	49.12	24.66	.139	–10.71	108.95
	1912–1950	62.13 ^a	23.68	.030	4.69	119.58
Female's Comparison for each time period						
1879–1911	1912–1950	–40.02 ^a	12.29	.005	–69.95	–10.09
	1951–1965	–57.25 ^a	19.86	.015	–105.61	–8.89
1912–1950	1879–1911	40.02 ^a	12.29	.005	10.09	69.95
	1951–1965	–17.22	18.94	.742	–63.35	28.89
1951–1965	1879–1911	57.25 ^a	19.86	.015	8.89	105.61
	1912–1950	17.22	18.94	.742	–28.89	63.35

^aThe mean difference is significant at the 0.05 level.

provided in Table 6 covering the same time period that the present study is focused upon. Despite the fact that our study is based on skeletal remains and regression equations for estimating stature, a direct comparison is useful for assessing the accuracy of the regression estimates as well as the representability of our sample. Furthermore, the skeletal material at hand is the only available sample representing the Greek population for such a long time interval. And despite its size-related limitations, we consider it significant for extracting some useful and consistent information on the secular change

in stature against which previous studies with larger samples but very limited time spans can be compared. Additionally, the well-distributed geography of our skeletal sample with regard to birthplace, which covers the entire Greek territory, further emphasizes its reference value.

Regarding the first period (1879–1911), there are observed differences compared with our sample both in male mean stature, ranging from 1.5 to 3.3 cm, and female mean stature, from 3.5 to 5.6 cm (Exarchopoulos, 1931; Gedeon, 1931; Labadarios, 1921; values taken from Fthenakis, 1970).

TABLE 5 Descriptive statistics of stature for both sexes

Sex	Stature (cm)				
	N	Mean	Std. error	Std. deviation	Kolmogorov-Smirnov
Males					
1879–1911	36	165.7	1.2	7.17	0.200
1912–1950	59	164.4	0.8	6.23	0.200
1951–1965	9	170.6	2.3	6.83	0.200
Total	104	165.3	0.7	6.78	0.200
Females					
1879–1911	29	151.4	1.0	5.41	0.200
1912–1950	47	155.4	0.8	5.17	0.200
1951–1965	9	157.1	1.5	4.60	0.200
Total	85	154.2	0.6	5.56	0.200

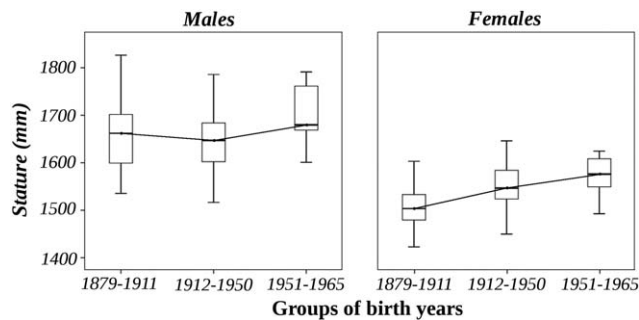


FIGURE 2 Male (left) and female (right) stature compared between groups of birth years: 1879–1911, 1912–1950, and 1951–1965

However, both samples, which Exarchopoulos and Labadaros measured, originate from high school students of Athens, implying an upper class SES, since at that time high school graduates were found mostly among the privileged parts of the Greek population. However, Gedeon (1931) findings concerning 17 year old males of the same period, and also from the Athens area, show a 2.3 cm height difference for the samples with high SES, whereas the reported mean height for the deprived population is almost identical with the stature estimated from our skeletal sample. Roberts (1954) and Hasluck and Morant (1929) have also reported height measurements for male Cretan and Greek Macedonian general populations of the late 19th century, respectively. Although both populations appear taller than our sample mean for this period, it is expected as Valaoras (1970) reports in his findings that the Cretans and the Greek Macedonians are the tallest subpopulations of the Greek ethnicity pool. Valaoras (1970) also reports significant variation among different parts of Greece and concludes that the larger part of this variation results from environmental factors, which is not surprising considering the vast differences in the morphology of the Greek landscape.

Concerning the second period of interest (1912–1950), the only available data on Greek stature are reported by Valaoras and Papaioannou (1948) obtained from high school students of Athens during the period 1925–1930 and once again by Valaoras (1970), who assembled a massive sample (nearly half a million males) based on the records of the Greek Armed Forces and reported on yearly stature secular change of Greek army conscripts born between 1927 and 1945. Once again, their findings show a difference in measured height compared to our sample. Regarding the Athenian high school students, it is remarkable that although the female sample shows analogous difference (3.3 cm) with respect to the observed differences in the previous period, the mean height for male high school students is 7.3 cm higher than our sample. What is more important, is that although our sample exhibits a negative slope of secular change for the stature of males between the first and second period, comparing the available height data from the male

school students of Athens indicates the opposite. Considering the aforementioned socioeconomic background of the high school students, one possible explanation is that the increase of height, which also reflects advancements in health, nutrition, and SES, is only reflected in a small part of the Greek population and more specifically in upper-class urban citizens such as the high school student samples of the capital city of Greece. This social class-oriented explanation is further reinforced by the fact that the height of the general young male population (army conscripts) during almost the same years (1927–1930) as reported by Valaoras (1970) averages at 166.76 cm (sample size: 89,269 males), which negatively offsets the high school students by 5 cm.

The important finding of our study concerning the second period is that Greek males exhibit a decline in stature, whereas females follow a continuous increase. This is in agreement with previously reported results as shown in Table 6. With the sole exception of the poor 17 year old Athenians between 1905 and 1910 (Gedeon, 1931), all other male samples from the first period exhibit higher mean stature than the mean stature of army conscripts between 1927 and 1945. On the contrary, the female mean stature in the first period is lower than that reported in the second period by Valaoras and Papaioannou (1948). Unless we assume measurement error or misprint in the reported results regarding the male high school students of the second period, one could argue that at the same time span that general male population stature appears to decline as a result of continuous warfare and economic deprivation during the period 1912–1950, the socioeconomic elite in the urban areas benefits from technological advances in health care and nutrition. However, the fact that the female Athenian high school students do not show such radical increases in stature challenges this argument. It should also be emphasized that the army conscripts sample, apart from being massive and well-distributed over the entire Greek territory, constitutes the most socioeconomically unbiased data sample available recording the height of the Greeks. Hence, we consider it a far more solid reference for mean stature over the second period than the Athenian high school students.

The third period between 1951 and 1965 is well-documented due to the works of Fthenakis (1970) as well as Manolis, Neroutsos, Zafeiratos, and Pentzou-Daponte (1995). The only drawback is that all samples include higher education students, who may be biased in their SES background to some extent. Nevertheless, it should be noted that from a geographical perspective these samples are most likely broadly distributed all over Greece. During the 50s and 60s, the universities of Athens and Thessaloniki were the two most prominent academic institutions in Greece, attracting students from the entire country. Manolis et al. (1995), who studied the secular changes in Greek students born between 1950 and 1970, reported that the average

TABLE 6 A comprehensive list of previous findings on average stature in Greece

Author and year of publication	Population surveyed	Time period based on year of birth	Sample (size and sex)	Mean stature (cm)	Difference from skeletal reference collection (cm) ^a
Roberts (1954)	Cretans, general population	1870–1880	1981 Males	168.5	2.8
Hasluck & Morant (1929)	Greek Macedonians, general population	1880–1900	200 Males	167.3	1.6
Labadarios (1921) ^b	High school students, Athens	1900–1905	Males	169.0	3.3
			Females	158.0	5.6
Exarchopoulos (1931) ^b	High school students, Athens	1900–1910	Males	167.2	1.5
			Females	154.0	2.6
Gedeon S (1931) ^b	Poor 17 year old, Athens	1905–1910	Males	165.5	–0.2
			Females	154.9	3.5
Gedeon S (1931) ^b	Rich 17 year old, Athens	1905–1910	Males	168.1	2.3
			Females	155.9	4.5
Valaoras & Papaioannou (1948)	High school students, Athens	1925–1930	Males	171.7	7.3
			Females	158.7	3.3
Valaoras (1970)	Army conscripts	1927–1945	468,056 Males	167.2	2.8
Fthenakis (1970)	Higher education students, Thessaloniki	Early 1950s	Males	174.7	4.1
			Females	158.6–161.1	1.5–4.0
Manolis et al. (1995)	Higher education students, Athens	1950–1955	357 males	173.7	3.1
			232 females	161.4	4.3
Manolis et al. (1995)	Higher education students, Athens	1955–1960	110 males	175.9	5.3
			198 females	162.7	5.6
Manolis et al. (1995)	Higher education students, Athens	1960–1965	115 males	177.0	6.4
			115 females	162.7	5.6

^aDifferences are calculated with respect to the corresponding period of this study. The differences relate to contemporary but not identical timespan.

^bValues taken from Fthenakis (1970).

height of male students increased continuously during that period. Their findings are partially presented in Table 6 only for the time period that concerns this study and both male and female groups exhibit higher figures compared to our sample. More specifically, there are 9 male and 9 female individuals in our sample born between 1950 and 1965 with a mean stature of 170.6 and 157.1 cm, respectively. Comparing these figures with the weighted average over the same period from Manolis et al. shows a difference in height of

4.2 cm for male group and 5 cm for female group. Similar differences in the mean stature appear between our sample and the sample reported by Fthenakis (1970), who documented stature records from higher education students in Thessaloniki born in the early 1950s. Both university samples follow the same pattern of secular change in stature with respect to previous periods, also present in our skeletal sample, which shows that the stature of the Greek population increases in the post-war period both for males and females.

Additionally, the reported differences with our skeletal sample (see Table 6) during the third period are mostly higher than the observed differences in the previous two periods, which could be another attribute of different SES background between samples. However, the very small sample size of our skeletal collection during the third period cannot account for such assumption. More specifically, the 9 males and 9 females corresponding to the third period, although adequate for validating the assumptions of the statistical analysis presented in our study, they exhibit a much higher standard error (see Table 5) than the previous two periods examined. Consequently, the larger differences between our sample and the reported mean statures of the third period may well be attributed to the higher standard error of our sample over the same period.

Despite the fact that the size of our skeletal sample does not allow a proper statistical correlation with the available published data on height, a careful observation reveals some consistency in the observed differences between the estimated stature from the skeletal collection and the reported measured heights of living people. Some explanation has already been attributed to differences in socioeconomic background or regional dispersal between the population samples. However, other factors including systematic errors can contribute to the observed difference with the most prominent being sampling error (Särndal et al., 1992). When considering the size, standard error, and standard deviation values of our sample along with the differences between mean heights of different authors from the same period (Exarchopoulos, 1931; Labadarios, 1921; see Fthenakis, 1970), the observed differences may well be attributed to sampling error. Furthermore, it should be pointed out that stature estimation from regression formulas based on femoral length induces further bias compared with documented stature.

The statistical analysis of our skeletal sample has shown a statistically significant increase in secular change of stature of the Greek population over the period between 1879 and 1965. Compared to other populations, the overall positive secular change in stature of the Greek population during the studied time period remains moderate, and it is not until the 1960s and onward that a rapid increase in stature is observed (Floud, Fogel, Harris, & Hong, 2011), which is consistent with the rapid industrialization and economic growth in Greece since that decade. Regarding the Greek population, our findings on secular trend are consistent with previous findings, which also report a positive secular change in the height of male and female Greeks (Fthenakis, 1970; Manolis, Neroutsos, Zafeiratos, & Pentzou-Daponte, 1995). However, there are also some discrepancies with previously reported findings with the most prominent concerning the decline of male stature during the second period. Our sample shows that males suffer a decrease in mean stature between 1912 and 1950, which coincides well with the war period in

Greece. Nutritional deprivation during childhood as a result of war may well have had an impact on height (Akresh, Bhalotra, Leone, & Osili, 2012) thus explaining our findings. However, this is not evident in our data regarding female stature, which continues to increase during war times. One possible explanation of this phenomenon is that males and females follow a distinct growth pattern based on different environmental stress. It is not uncommon for male children to be more often directly involved in warfare than female children (Wessells, 1997), although no such evidence is available concerning the Greek wars examined in the present study. Nevertheless, even if not directly involved, the absence of adult males due to their involvement in warfare would definitely put a heavier burden on the male children for supporting their families left behind, which can be explained by the cultural and moral standards of the Greek society during those times. Analogous opposing secular changes between males and females have been reported (Himes & Mueller, 1977).

Another sample-related implication regarding the secular change of the second period may be attributed to the geographic distribution of our skeletal sample. Following Valaoras' (1970) findings on significant height variation of Greek subpopulations due to geographically related differences, we examined our data for such sources of sampling error that could significantly bias the observed secular changes of the second period. The Greek subpopulation that mostly deviates from the overall Greek sample (see Table 1) is that from Asia Minor. The descriptive statistics regarding the Greeks from Asia Minor compared to the rest of the Greek population, designated as "Mainland" is presented in Table 7. Despite the large standard errors due to the small sample size from Asia Minor, the males follow an opposite secular change between the first and second periods, whereas females show an even more radical positive secular change than the rest population during the same time. However, a follow-up examination of the "Mainland" samples revealed the same secular patterns as the overall Greek population (presented in Table 4) with somewhat different statistical values. It should be noted, although, that whether the 2 males and 3 females from Asia Minor constituting the subsamples of the second period reflect an average of Asia Minor Greeks during that period or whether they are just "sampling errors" of our skeletal collection could be no more than an educated guess. Nevertheless, despite the observed differences, whether a result of geographical variation or sampling error, the distinctively opposing patterns of secular change between Greek males and females during the second period still hold. An observation that suggests sexual dimorphism is present in the secular change of the Greek population during that time.

Finally, it should be noted that although the army conscripts in Valaoras' work exhibit a mean stature lower than those reported by other researchers with regard to the first

TABLE 7 Differences in mean stature of Asia Minor Greek individuals as compared with the rest of the Greek population

Sex	Birthplace	Stature (cm)			
		N	Mean	Std. Deviation	Std. Error
Males					
1879–1911	Mainland	30	166.2	7.4	1.3
	Asia Minor	6	162.9	5.9	2.4
1912–1950	Mainland	57	164.3	6.2	0.8
	Asia Minor	2	165.4	10.9	7.8
1951–1965	Mainland	9	170.6	6.8	2.3
	Asia Minor	–	–	–	–
Females					
1879–1911	Mainland	20	151.6	4.6	1
	Asia Minor	9	150.8	7.2	2.4
1912–1950	Mainland	44	155.0	4.9	0.7
	Asia Minor	3	160.0	7.8	4.5
1951–1965	Mainland	9	157.0	4.6	1.5
	Asia Minor	–	–	–	–

period, Valaoras' results show that the decline in mean male stature does not begin until 1935. However, the time span covered by the conscripts' record (1927–1945) is a subdivision of the second period examined in this study. A closer look at the male individuals of our sample coinciding with Valaoras' examined time span yielded a mean stature of 164.7 cm ($N = 27$, $SD = 6.5$, $SE = 1.2$), whereas the mean stature of the previous years since the beginning of the second period (1912–1926) yielded a mean stature of 163.1 cm ($N = 27$, $SD = 6.5$, $SE = 1.2$). Although these values do not contradict the tendency in stature reported by Valaoras, they are still lower than the mean stature of our skeletal sample corresponding to the first period examined (see Table 5). Thus, we can argue that male individuals suffered a decline in stature after 1912 as compared with earlier years. This study, which covers a significantly wider time span than any other published work, aimed to identify how the socioeconomic conditions in Greece have affected the stature of its population. In this aspect, we can argue that the negative effect of economic and nutritional deprivation on stature as a result of warfare is apparent in our sample, at least for males. Furthermore, the positive effect of economic growth on stature is also prominent for the entire population.

5 | CONCLUSIONS

Although this study is constrained by the limited sample size and abides by certain assumptions, there are a number of substantial findings. The Greek population shows a moderate positive secular change in stature during the period 1879–1965, which is more prominent in females. The economic and nutritional deprivation of the Greek population as a

result of the long inter-war period in the first half of the 20th century has left its mark on male stature during that period. There is evidence of sexual dimorphism with regard to the secular patterns of the Greek population that may be attributed to cultural and moral standards of the Greek society. The Athens skeletal reference collection may be regarded reasonably representative of the Greek population to a certain extent. Finally, secular change in the Greek population does not really get underway until the 1960s, which postdates the time interval of the present study.

AUTHOR CONTRIBUTIONS

All authors read and approved the final version of the manuscript.

Study design: Andreas Bertsatos, Maria-Eleni Chovalopoulou

Directed implementation: Andreas Bertsatos, Maria-Eleni Chovalopoulou

Data collection: Andreas Bertsatos, Maria-Eleni Chovalopoulou

Analyzed the data: Andreas Bertsatos

Drafted the manuscript: Andreas Bertsatos

Edited the manuscript for intellectual content: Maria-Eleni Chovalopoulou

Critical comments on the manuscript: Maria-Eleni Chovalopoulou

ORCID

Andreas Bertsatos  <http://orcid.org/0000-0001-8734-5236>
 Maria-Eleni Chovalopoulou  <http://orcid.org/0000-0001-8187-8031>

REFERENCES

- Akresh, R., Bhalotra, S., Leone, M., & Osili, U. O. (2012). War and Stature: Growing Up during the Nigerian Civil War. *American Economic Review*, 102(3), 273–277.
- Buikstra, J. E., & Ubelaker, D. H. (1994). *Standards for data collection from human skeletal remains. (Report Number 44)*. Fayetteville: Arkansas Archeological Survey Publications.
- Eliopoulos, C., Lagia, A., & Manolis, S. (2007). A modern, documented human skeletal collection from Greece. *HOMO-Journal of Comparative Human Biology*, 58(3), 221–228.
- Floud, R., Fogel, R. W., Harris, B., & Hong, S. C. (2011). *The changing body: Health, nutrition, and human development in the Western world since 1700*. Cambridge: Cambridge University Press.
- Fthenakis, W. (1970). Größe und Gewicht griechischer Kinder und Jugendlicher unter Berücksichtigung der Akzeleration. *Anthropologischer Anzeiger*, 32(1/2), 24–45.
- Gedeon, S. (1931) *Pedometric studies in Greece*. Athens: Dimitrakos. (In Greek, English summary).
- Hasluck, M. M., & Morant, G. M. (1929). Measurements Of Macedonian Men. *Biometrika*, 21(1–4), 322–336.
- Himes, J. H., & Mueller, W. H. (1977). Aging and secular change in adult stature in rural Colombia. *American Journal of Physical Anthropology*, 46(2), 275–279.
- Kasimis, C., & Kassimi, C. (2004). Greece: A History of Migration. Retrieved from <http://www.migrationpolicy.org/article/greece-history-migration>. Accessed on June 1, 2004.
- Keep, R., & Bogin, B. (1999). Eight thousand years of economic and political history in Latin America revealed by anthropometry. *Annals of Human Biology*, 26(4), 333–351.
- Komlos, J. (2009). Anthropometric history: An overview of a quarter century of research. *Anthropologischer Anzeiger*, 67(4), 341–356.
- Lango, A. H., Estrada, K., Lettre, G., Berndt, S. I., Weedon, M. N., Rivadeneira, F., . . . Hirschhorn, J. N. (2010). Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature*, 467(7317), 832–838.
- Manolis, S., Neroutsos, A., Zafeiratos, C., & Pentzou-Daponte, A. (1995). Secular changes in body formation of greek students. *Human Evolution*, 10(3), 199–204.
- Martin, R. (1928). *Lehrbuch der anthropologie*. Jena: Fischer.
- Petmezas, S. D. (2006). Agriculture and economic growth in Greece. *IEHC 2006 Helsinki Session*, 60.
- Raxter, M. H., Auerbach, B. M., & Ruff, C. B. (2006). Revision of the Fully technique for estimating statures. *American Journal of Physical Anthropology*, 130(3), 374–384.
- Roberts, D. F. (1954). The Cretans. A Geographical Analysis of Some Aspects of Their Physical Anthropology. *The Journal of the Royal Anthropological Institute of Great Britain and Ireland*, 84(1/2), 145.
- Ruff, C. B., Holt, B. M., Niskanen, M., Sladák, V., Berner, M., Garofalo, E., . . . Tompkins, D. (2012). Stature and body mass estimation from skeletal remains in the European Holocene. *American Journal of Physical Anthropology*, 148(4), 601–617.
- Sapounaki-Dracaki, L. (1998). Heights and Nutritional Status in Greece. In J. Komlos & J. Baten (Eds.), *The biological standard of living in comparative perspective* (Contributions to the conference held in Munich January 18–22, 1997, for the XIIIth congress of the International Economic History Association, pp. 408–412). Stuttgart: Franz Steiner.
- Särndal, C. E., Swensson, B., & Wretman, J. (2003). *Model assisted survey sampling*. New York: Springer.
- Schell, L. (1986). Community Health Assessment Through Physical Anthropology: Auxological Epidemiology. *Human Organization*, 45(4), 321–327.
- Schweich, M., & Knüsel, C. (2003). Bio-cultural effects in medieval populations. *Economics & Human Biology*, 1(3), 367–377.
- Sjøvold, T. (1990). Estimation of stature from long bones utilizing the line of organic correlation. *Human Evolution*, 5(5), 431–447.
- Steckel, R. H. (2001). *Health and nutrition in the preindustrial era: Insights from a millennium of average heights in Northern europe*. Cambridge, MA: NBER.
- Steckel, R. H. (2008). Biological Measures of the Standard of Living. *Journal of Economic Perspectives*, 22(1), 129–152.
- Susanne, C. (1985). Living conditions and secular trend. *Journal of Human Evolution*, 14(4), 357–370.
- Trotter, M., & Gleser, G. C. (1952). Estimation of stature from long bones of American Whites and Negroes. *American Journal of Physical Anthropology*, 10(4), 463–514.
- Valaoras, V. (1970). Biometric studies of army conscripts in Greece. Mean stature and ABO blood-group distribution. *Human Biology*, 42(2), 184–201.
- Valaoras, V., & Papaioannou, S. (1944). La taille et le poids moyens des écoliers grecs, pendant la guerre. *Proceedings of the Academy of Athens 19*, 293–299. (In Greek, French summary).
- Waldron, T. (1998). A note on the estimation of height from long-bone measurements. *International Journal of Osteoarchaeology*, 8(1), 75–77.
- Wessells, M. (1997). Child Soldiers. *Bulletin of the Atomic Scientists*, 53(6), 32–39.
- Wood, A. R., Esko, T., Yang, J., Vedantam, S., Pers, T. H., Gustafsson, S., . . . Frayling, T. M. (2014). Defining the role of common variation in the genomic and biological architecture of adult human height. *Nature Genetics*, 46(11), 1173–1186.

How to cite this article: Bertsatos A, Chovalopoulou M-E. Secular change in adult stature of modern Greeks. *Am J Hum Biol*. 2017;e23077. <https://doi.org/10.1002/ajhb.23077>



“What and how should we share?” An inter-method inter-observer comparison of measurement error with landmark-based craniometric datasets

Andreas Bertsatos^{1,*}, Elissavet Gkaniatsou², Christina Papageorgopoulou², and Maria-Eleni Chovalopoulou¹

¹ Department of Animal and Human Physiology, Faculty of Biology, School of Sciences, National and Kapodistrian University of Athens, Panepistimiopolis, GR 157 01 Athens, Greece

² Department of History and Ethnology, Democritus University of Thrace, 1 P. Tsaldari Street, 69100 Komotini, Greece

* Corresponding author: abertsatos@biol.uoa.gr

With 2 figures and 7 tables

Abstract: The present study evaluates the precision and accuracy of photogrammetric 3D modeling of human crania in landmark acquisition and explores the limitations of combining datasets acquired by different observers and different measurement methods. Our working sample comprises 50 adult human crania, which were modeled with 3D photogrammetry. 3D coordinates of 56 landmarks were collected from the 3D models with Meshlab software and an existing corresponding dataset digitized with Microscribe-3DX has been utilized. Measurement error for landmark configurations and Inter Landmarks Distances (ILDs) for each type of landmarks has been assessed through least root mean squared deviation and mean absolute error respectively. Inter-observer error has been assessed on a sub-sample of 20 crania, which was also used for caliper measured ILDs. Between-methods Technical Error Measurement (TEM) based on ILDs has been calculated for evaluating the interchangeability for different datasets. Photogrammetric 3D models and Microscribe-3DX share identical rated accuracy regarding craniometric applications and both methods show increased accuracy in locating type I landmarks as opposed to types II and III. However, photogrammetric 3D models perform better in terms of inter-observer error suggesting higher reliability of measurements. Furthermore, ILDs are less prone to measurement error than landmark configurations. Finally, ILDs exhibit similar relative TEM of about 1.5% between Microscribe, caliper and 3D model based measurement methods. Combining datasets of landmark coordinates acquired from photogrammetric 3D models does not compromise the statistical integrity in terms of measurement error, which also applies to pooling ILD datasets from multiple methods. Nevertheless, compiling 3D datasets from multiple methods for 3DGM analysis should be done cautiously.

Keywords: photogrammetric accuracy of skeletal modelling; cranial landmarks; inter-method error; inter-observer error; combining datasets

1 Introduction

Analysis of metric traits may well be regarded the cornerstone of osteological research. Traditionally, caliper-derived distance and angle measurements have been used for size and shape analysis (Ross & Williams 2008), whereas the advent of landmark based three-dimensional geometric morphometrics (3DGM) over the past decades has opened new horizons in analyzing complex shape data in physical anthropology utilizing digital specimens (Shearer et al. 2017). Both approaches offer pros and cons over one another and are utilized in modern osteological research. Regardless of their individual differences, which still triggers some discussion over their strengths and weaknesses (Spradley & Jantz 2016;

Bertsatos et al. 2018), they both share an intrinsic property which is measurement error.

Despite measurement error being always present when assessing metric traits, reports containing measured parameters rarely included any indication of the error or the reliability of those measurements (Lewis 1999). To our benefit, this habit has been changing over the past decade with recent work increasingly including inter- and intra-observer error estimates for evaluating the reliability of their measurements. Furthermore, a number of studies have explicitly focused on measurement error as a function of different observers or particular quantities being measured, whether they involve standard osteometric measurements (Langley et al. 2018) or 3D landmark-derived datasets in geometric morphometrics

(Barbeito-Andrés et al. 2012; Fruciano 2016; Shearer et al. 2017). Such studies are of major importance, since they not only systematically detect sources of measurement error and measure its magnitude on specific osteometric parameters, but also they allow conclusions to be drawn about the accuracy and repeatability of relevant work, where measurement error may not be reported or quantified.

The observer and his/her interaction with the measuring equipment as well as the sample constitute the primary sources of measurement error in any osteometric study. Most often well-designed studies report on inter- and intra-observer errors as a means of evaluating the induced variation related to error in their data collection, the repeatability/reproducibility of their results and by extension the validity of their results (von Cramon-Taubadel et al. 2007). The popularity of geometric morphometrics in osteological research along with the ever-increasing availability of diverse equipment for digitizing osteological specimens (Robinson & Terhune 2017) as well as the increasingly frequent sharing of geometric morphometric datasets among researchers (Fruciano et al. 2017) has led to another “potentially harmful” source of systematic measurement error, which results from the use of different measuring methods or equipment in pooled samples. As Fruciano et al. (2017) conclude, such measurement error can be a source of substantial variation with unexpected effect on parameter estimates when combining different morphometric datasets. Additionally, these concerns not only apply when pooling datasets from diverse sources, but also valid when testing methods. For example, testing sex discriminant functions produced with one measurement method from dataset A on another dataset B measured with a different method introduces analogous confounding measurement error as pooling both datasets together.

The aforementioned example merely illustrates a common practice of indirect “data sharing” and its implications can be prominent especially in forensic anthropology cases or bioarchaeology, where equipment or other resources may vary between labs or in field work. In order to measure the potential effect of such practices, a number of studies have focused on the precision and accuracy of various equipment especially for producing and measuring digital specimens (Ross & Williams 2008; Barbeito-Andrés et al. 2012; Toneva et al. 2016; Toneva et al. 2017). Other researchers have extended their focus on inter-method error that arises from the use of different equipment and its implications in pooled samples (Sládek et al. 2012; Fruciano et al. 2017; Robinson & Terhune 2017; Shearer et al. 2017). Although comparative studies in current literature cover a diverse range of measuring equipment from calipers and 3D digitizers to laser or structured light scanners as well as X-ray computed tomography (CT) systems, the precision and accuracy of 3D photogrammetry for digitizing skeletal samples has not been thoroughly evaluated yet to provide a measurable reference error between different types of landmarks, types of mea-

surements and methods. Previous work by Katz & Friess (2014), who compared 4 cranial samples reconstructed with Agisoft Photoscan Pro v0.9.1 and a Breuckmann smartscan stereo 3D surface scanner, has shown that 3D photogrammetry is suitable for use in many types of morphometric studies, since it produced similar models and exhibited similar repeatability with the Breuckmann scanner, although they suggested that mixing samples might be inappropriate due to the detectable differences between the two 3D reconstruction methods. Nevertheless, their study, which was based on a fairly limited sample, didn’t investigate the precision of landmarks as a function of type (Katz & Friess 2014), which is also important when combining datasets (Fruciano et al. 2017; Robinson & Terhune 2017).

Photogrammetry is a method for producing 3D digital models of real life objects from a set of pictures taken with a digital camera. Despite being very popular in other fields such as land surveying, cultural heritage or archaeology, photogrammetric modeling is relatively new in osteological research, where the vast majority of digital datasets derive from 3D digitizers, laser scanners or X-ray CT systems. The present study evaluates the precision and accuracy of photogrammetric 3D modeling of human crania in landmark acquisition. This is accomplished by extracting linear and angular measurements from the digital models and comparing them with corresponding measurements taken with caliper and 3D Microscribe digitizer. To further evaluate accuracy and precision in landmark acquisition, 3D coordinates of different types of landmarks are extracted and compared between observers and against its corresponding dataset obtained with 3D Microscribe. Since digital specimens are increasingly being used for both landmark-based geometric morphometrics but also for extracting “traditional” measurements in terms of inter-landmark distances (ILDs) and/or inter-landmark angles (ILAs), precision and accuracy of model-based landmark acquisition was evaluated for superimposed landmark configurations as well as possible combinations of ILDs. It is important to note that the present research evaluates the measurement error among types of landmarks or measurements both as a means of assessing the precision and accuracy of photogrammetry as an osteometric tool, but also in comparison to other measurement tools. Following Fruciano’s and colleagues question if “sharing is caring?” (Fruciano et al. 2017) and acknowledging that sharing either datasets or methods is inevitable, the present study aims to answer “what and how should we share?”.

2 Material and methods

The cranial samples used in the present study are part of the modern reference skeletal collection (Athens collection) housed in the Department of Animal and Human Physiology at the National and Kapodistrian University of

Athens. 3D cranial models from 50 adult individuals have been created with photogrammetry. The sample comprises intact crania from both males and females with their respective age-at-death ranging from 35 to 86 years old. Sample selection was based on equal representation of sexes and adult age groups but also on the availability of previously recorded data.

2.1 Photogrammetric modeling

Each sample was statically positioned in two opposite orientations and was photographed from 65 to 70 different views in each orientation with a DSLR Canon 100D camera and a fixed focal length Canon EF 40 mm f/2.8 STM lens. Each set of photographs was processed as a separate chunk with Photoscan v.1.4 software (Agisoft LLC, Russia). After initial camera alignment, repeated alignment optimization was applied after removing sparse cloud points with reconstruction uncertainty higher than 8.5 and reprojection error higher than 8 pixels. High accuracy dense point clouds were generated for each chunk and point based chunk alignment was used to merge the paired chunks for each sample into a single model. After removing the points that were not part of the crania, each model's dense cloud, which counted from 11 to 14 million points, was used for mesh generation. The generated mesh for each model was limited to 500,000 faces and texture was generated with a single texture atlas set to 8192×8192 pixels. Each cranial model was exported in Wavefront object format and Meshlab software was used for scaling and extracting landmark 3D coordinates.

2.2 Data acquisition

Scaling was performed by measuring three arbitrary distance measurements on the 3D models and their respective real world measurements on the crania with a Mitutoyo 150 mm dial caliper with a graduation of 0.02 mm and a rated accuracy of ± 0.03 mm. The scaling measurements were taken between distant pseudo-landmarks that could be accurately spotted on the textured mesh of the 3D models as well as on the real crania. Subsequently, 3D coordinates were extracted for 56 landmarks (12 midline and 22 bilateral), which are listed in Table 1 along with their corresponding type according to the categorical classification of landmarks after Bookstein (1991). Landmark 3D coordinates from the crania of the Athens collection have been previously acquired with a Microscribe-3DX digitizer (Immersion Corp. San Jose, California), which has a rated accuracy of ± 0.23 mm, and the related intra- and inter-observer errors regarding landmark digitization have been reported (Chovalopoulou et al. 2017) ranging from 0.3 to 1.8 mm. The subset of Microscribe coordinates corresponding to the sample and landmarks of the present study has been reutilized for comparison. However, it should be noted that both datasets (present: on 3D models, previous: with Microscribe) utilized for inter-method comparison were acquired by the same observer. From the present sample, the landmark coordinates from 20 cranial models have been extracted twice, each time by a different observer, to account for the inter-observer error on the acquisition of 3D landmark coordinates from the models with Meshlab software. The same 20 samples have been manually mea-

Table 1. List of digitized landmarks.

Abbrev.	Landmark name	Type	Abbrev.	Landmark name	Type
ba	Basion	II	fmo	Frontomalare orbitale Left & Right	III
b	Bregma	I	fmt	Frontomalare temporale Left & Right	III
inc	Foramen incisivum	II	ft	Frontotemporale Left & Right	III
g	Glabella	III	io	Infraorbitale Left & Right	II
ho	Hormion	I	it	Infratemporale Left & Right	III
l	Lambda	I	ju	Jugale Left & Right	II
n	Nasion	I	x	Landmark x Left & Right	I
ns	Nasospinale	III	ms	Mastoidale Left & Right	III
o	Opisthion	II	mf	Maxillofrontale Left & Right	II
op	Opisthocranion	III	mnf	Maxillonasofrontale Left & Right	I
sta	Staphylion	III	ovm	Ovale mediale Left & Right	II
sr	Staurion	I	sphn	Sphenion Left & Right	I
ast	Asterion Left & Right	I	sau	Supra Auricular Left & Right	III
cam	Caroticum mediale Left & Right	II	scs	Supramastoid crest – squamous suture intersection Left & Right	III
co	Coronale Left & Right	III	zm	Zygomaxillare Left & Right	III
k	Crotaphion Left & Right	I	zti	Zygotemporale inferior Left & Right	III
ec	Ectoconchion Left & Right	II	zts	Zygotemporale superior Left & Right	III

Table 2. List of measurements used for TEM.

id ^a	type	Landmarks (abbrev)
D1	distance	ovm R – ms R
D2	distance	ms L – fmt R
D3	distance	sphn L – ms R
D4	distance	ms L – co R
D5	distance	ms L – ju R
D6	distance	ms L – sphn R
A2	angle	ms R – op – zts R
A4	angle	g – mnf R – b
A5	angle	g – mf R – b
A6	angle	ovm L – ms R – co R

^a The measurement names have been adopted from Bertatos et al. (2018) for easier cross-referencing.

sured with a dial caliper over a set of ILDs and ILAs, which are listed in Table 2, for inter-method comparison regarding linear measurements.

2.3 Data analysis

Accuracy of the produced photogrammetric models was evaluated by their corresponding scaling precision, which was based on the mean absolute percentage error of the paired measurements used for scaling. For each sample the scaling factor was calculated as the mean ratio of the three paired measurements used for scaling. Subsequently, scaling precision was calculated as the percentage of the absolute errors between each digital measurement, multiplied by the respective sample's scaling factor, and its corresponding manual measurement divided by the latter. Descriptive statistics for evaluating scaling precision were applied on the total sample accounting for 150 paired measurements. Measurement precision from the subsample of 20 photogrammetric 3D models was evaluated as inter-observer error for model-based landmark acquisition, which was calculated as the least root mean squared deviation (LRMSD) between paired sample landmark configurations after superimposition by the Kabsch algorithm (Kabsch 1976). For each type of landmark, a separate paired configuration was superimposed. Along with the LRMSD for each type of landmarks, the mean deviation for each individual landmark was also calculated. Additionally, descriptive statistics were calculated on the pool of absolute differences between all combinations of ILDs for each paired configuration.

In order to evaluate the accuracy of model-based landmark acquisition the same analytical workflow was applied on the paired sample landmark configurations between model-derived landmark coordinates from our entire sample and the available data set from the same 50 crania that have been previously digitized with Microscribe-3DX. Finally, the accuracy of photogrammetric modeling was also evaluated by calculating the technical error measurement (TEM)

as well as the scaled TEM (Lewis 1999) over the set of 10 ILAs and ILDs on the subsample of 20 crania. These measurements, which are considered the most sexually dimorphic cranial metric traits for the Greek population, have been previously used for evaluating TEM between Microscribe coordinates and caliper measurements (Bertatos et al. 2018). In the present study, these measurements have been computed from the model-derived landmark coordinates in order to calculate TEM among model-based measurements, Microscribe-based measurements as well as manual measurements.

All data analysis has been performed with GNU Octave programming language (Eaton et al. 2016). Implementation of the Kabsch algorithm along with some useful functions for exchanging point coordinates with Meshlab are provided as open source under the GPL v3.0 and can be found at <https://github.com/pr0m1th3as/wavefront-obj-mesh-package>. Statistical inference results by means of ANOVAs between measurement methods or types of landmarks have been produced with JASP software (JASP Team 2018).

3 Results

The distribution of the scaling precision values is plotted as a histogram in Fig. 1. Since precision was based on the mean absolute percentage error, the resulted distribution is heavily skewed but unimodal. Hence, the use of robust descriptive statistics for evaluating precision is more appropriate. The median scaling precision of the photogrammetric 3D models calculated as a percentage of the actual distance being measured was 0.145% with an inter-quartile range (IQR) of 0.215% (1st quartile = 0.069%, 3rd quartile = 0.284%). Regarding accuracy of linear and angular measurements, the TEM results between each pair of different measurement methods are listed in Table 3. Both TEM and scaled TEM values are given for each individual measurement as well as their respective mean values for distance and angular measurements. The average relative TEM for distance measurement is consistent among the different methods and is approximately 1.5%. Consistency is also observed among angular measurements with an average relative TEM about 3.1%. The average TEM for distance and angle measurements is approximately 1.7 mm and 1.7 degrees, respectively. An observed difference is that distance TEM is lower between model-based – microscribe measurements or model-based – manual measurements as opposed to microscribe – manual measurements (difference: ~ 0.3 mm), whereas angular TEM is marginally lower for model-based – manual measurements and Microscribe – manual measurements. Nevertheless, there is no statistically significant difference among pairs of measurement methods when considering the scaled TEM (ANOVA results for scaled TEMs: $F(2,18) = 0.126$, $p = 0.883$, $\eta^2 = 0.014$).

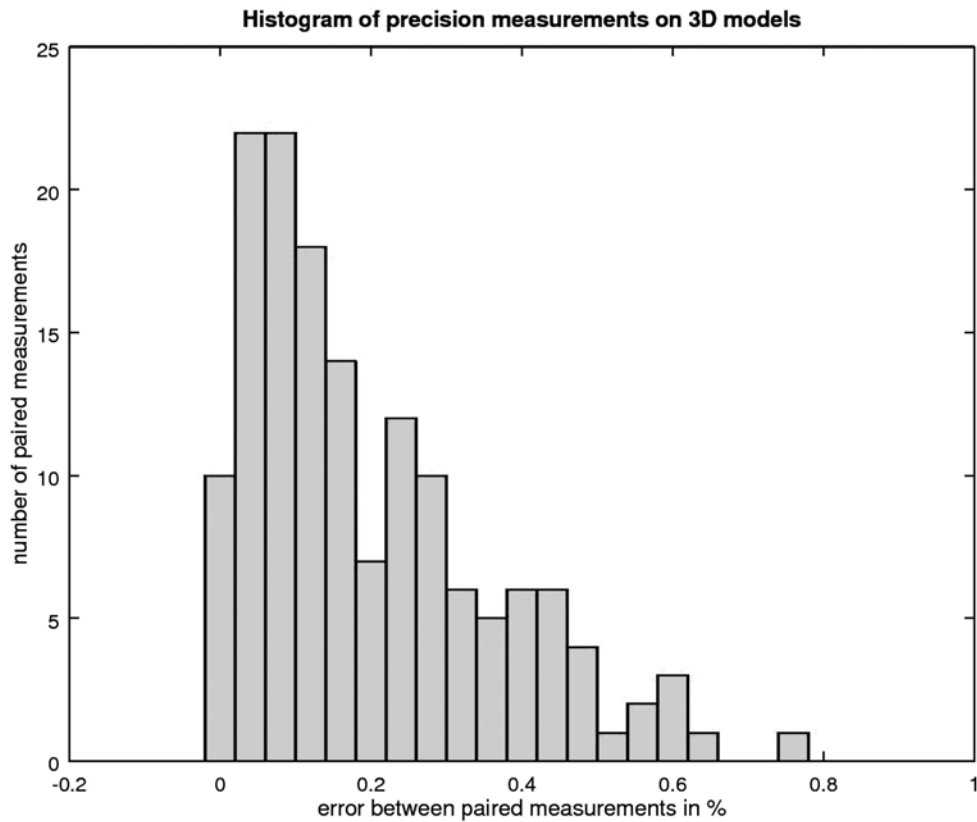


Fig. 1. Rated accuracy of photogrammetric modeling in terms of scaling precision repetitive measurements.

Table 3. TEM results among different measurement techniques.

	D1	D2	D3	D4	D5	D6	A2	A4	A5	A6	Distance mean	Angle mean
TEM ^a : 3D Models – Microscribe	1.26	1.43	2.42	1.98	1.46	1.28	0.93	2.98	2.21	0.91	1.64	1.75
Relative TEM ^b : 3D Models – Microscribe	2.5	1.0	1.8	1.3	1.1	0.9	4.0	4.8	3.2	1.3	1.46	3.33
TEM ^a : 3D Models – manual	1.33	1.01	1.86	3.07	0.88	1.62	0.83	2.82	1.43	1.41	1.62	1.62
Relative TEM ^b : 3D Models – manual	2.7	0.7	1.4	2.0	0.7	1.2	3.6	4.6	2.1	2.0	1.45	3.06
TEM ^a : Microscribe – manual	0.93	1.18	2.51	3.52	1.28	2.18	0.81	2.56	2.11	0.98	1.93	1.61
Relative TEM ^b : Microscribe – manual	1.8	0.9	1.9	2.3	1.0	1.6	3.5	4.0	3.0	1.4	1.57	2.98

^a Distances measured in mm and angles in degrees

^b Values are shown in percentage %

Accuracy and precision for landmark digitization has been assessed independently for each type of landmarks. Accuracy results are summarized in Fig. 2, which illustrates the measurement error between model-based and Microscribe data. The LRMSD and the median ILD absolute

deviation from each paired sample are shown in separate box plots. The standard 1.5*IQR rule is used for the whiskers, whereas outliers are annotated with '+' for points between 1.5 and 3 times the IQR and 'o' for points that lie outside 3 times the IQR. The mean deviation for each individual

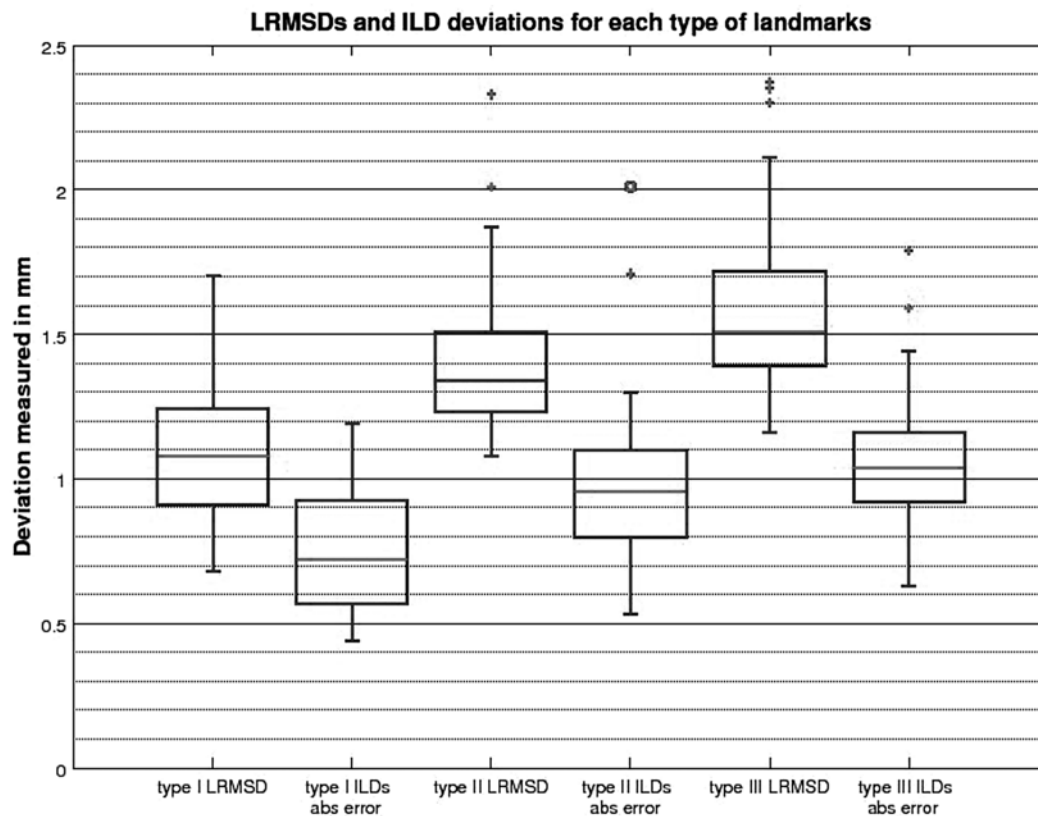


Fig. 2. Measurement errors between 3D model-based and Microscribe derived data.

landmark along with the respective inter-observer error on model-based landmark digitization are provided in Table 4, Table 5 and Table 6 for landmark types I, II and III respectively. Additionally, averaged descriptive statistics of ILD deviations for the entire sample are provided both for model-based with Microscribe paired configurations as well as for inter-observer error of model-based paired configurations. Finally, the sample mean of LRMSD between paired configurations is reported in each table accordingly. Average of medians of ILD deviations as well as average LRMSD values reported for each type of landmark are slightly different from what is presented in Fig. 2, since the box plots in Fig. 2 are showing the median of each corresponding distribution rather than the mean values. From a statistical perspective, the medians of the ILD deviations are correctly being used as measure of expected (typical) error based on the heavily skewed distributions observed in each paired sample (see averaged descriptives on Tables 4, 5, 6). However, the average of these medians is much more appropriate descriptor of the overall estimated landmark acquisition accuracy than the median (of medians) illustrated in the box plots or the averaged mean of ILD deviations, which is also reported in the relevant tables.

Regarding landmark digitization accuracy, the sample mean LRMSD showed that type I landmarks can be digitized most accurately (between methods: 1.1 mm; model-based inter-observer error: 0.47 mm). Type II landmarks produced a sample mean LRMSD of 1.4 mm between methods and type III landmarks exhibited the least accuracy (between methods: 1.6 mm), whereas both types had virtually the same model-based inter-observer error (type II: 0.71 mm; type III: 0.72 mm). When considering the ILD related accuracy as expressed by the average of ILD medians, all types of landmarks exhibited the same ascending order as with LRMSDs although with consistently less error between the paired configurations. More specifically, between methods related measurement error ranged from 0.74 to 1.05 mm, whereas inter-observer error was less than 0.5 mm in all cases with type I landmarks outperforming the rest. Subsequent ANOVA with Tukey's post-hoc or Kruskal-Wallis with Dunn's pairwise tests, which are summarized in Table 7, revealed statistically significant differences among all types of landmarks for LRMSDs between methods, whereas for ILDs only type I landmarks differ from type II and type III. Regarding model-based inter-observer error only type I landmarks differ significantly for both LRMSDs and ILDs.

Table 4. Digitization accuracy for type I landmarks.

Landmark (abbrev)	Landmark deviations after superimposition		ILD deviation between paired configurations		
	3D Models – Microscribe deviations ^a	3D Models inter-observer error ^a	Descriptive statistics	3D Models – Microscribe ^a	3D Models inter-observer error ^a
n	1.36	0.47	average of medians	0.74	0.27
b	1.01	0.47			
l	0.87	0.30	average mean	0.90	0.35
ho	0.93	0.40	average skewness	0.98	1.19
sr	1.41	0.41	average kurtosis	3.71	4.18
ast L	1.04	0.42	Average LRMSD between paired configurations after landmark superimposition		
k L	0.82	0.44			
sphn L	1.53	0.66	Descriptive statistics	3D Models – Microscribe ^a	3D Models inter-observer error ^a
x L	1.19	0.80			
mnf L	0.97	0.33	sample mean	1.10	0.47
ast R	1.17	0.18			
k R	0.98	0.69	Average LRMSD between paired configurations after landmark superimposition		
sphn R	1.12	0.52			
x R	1.01	0.94	Descriptive statistics	3D Models – Microscribe ^a	3D Models inter-observer error ^a
mnf R	0.97	0.42			

^aAll values shown in mm**Table 5.** Digitization accuracy for type II landmarks.

Landmark (abbrev)	Landmark deviations after superimposition		ILD deviation between paired configurations		
	3D Models – Microscribe deviations ^a	3D Models inter-observer error ^a	Descriptive statistics	3D Models – Microscribe ^a	3D Models inter-observer error ^a
O	1.46	0.55	average of medians	0.97	0.49
ba	1.56	0.61			
cam L	1.19	0.59	average mean	1.10	0.59
ovm L	1.42	1.14	average skewness	0.78	1.17
inc	1.13	0.54	average kurtosis	3.26	4.29
ju L	1.31	0.37	Average LRMSD between paired configurations after landmark superimposition		
mf L	1.25	0.75			
ec L	1.41	0.58	Descriptive statistics	3D Models – Microscribe ^a	3D Models inter-observer error ^a
io L	1.38	0.78			
cam R	1.37	0.75	sample mean	1.40	0.71
ovm R	1.30	0.83			
ju R	1.57	0.68	Average LRMSD between paired configurations after landmark superimposition		
mf R	1.40	0.78			
ec R	1.57	0.62	Descriptive statistics	3D Models – Microscribe ^a	3D Models inter-observer error ^a
io R	1.64	1.01			

^a All values shown in mm

Table 6. Digitization accuracy for type III landmarks.

Landmark (abbrev)	Landmark deviations after superimposition		ILD deviation between paired configurations		
	3D Models – Microscribe deviations ^a	3D Models inter-observer error ^a	Descriptive statistics	3D Models – Microscribe ^a	3D Models inter-observer error ^a
g	1.34	0.38	average of medians	1.05	0.41
op	2.55	0.83			
sta	1.31	0.41	average mean	1.25	0.51
it L	1.07	0.87	average skewness	1.11	1.30
ms L	1.79	0.87	average kurtosis	4.41	4.80
scs L	1.60	0.93	Average LRMSD between paired configurations after landmark superimposition		
co L	1.68	0.78			
ft L	1.79	0.68	Descriptive statistics	3D Models – Microscribe ^a	3D Models inter-observer error ^a
sau L	1.53	1.11			
zts L	1.56	0.61	sample mean	1.60	0.72
fnt L	1.67	0.51	Landmark deviations after superimposition		
zti L	1.22	0.91			
zm L	1.16	0.53	Landmark (abbrev)	3D Models – Microscribe deviations ^a	3D Models inter-observer error ^a
fmo L	1.53	0.73			
ns	2.43	0.60	zts R	1.33	0.38
it R	1.80	1.21	fnt R	1.55	0.81
ms R	1.62	0.65	zti R	1.13	0.74
scs R	2.25	0.84	zm R	1.21	0.49
co R	2.13	1.06	fmo R	1.28	0.45
ft R	1.82	0.70			
sau R	2.02	0.52			

^a All values shown in mm

Table 7. Digitization accuracy inferential statistics among types of landmarks.

	Between-methods			Model-based inter-observer		
	χ^2	df	sig. ^a	χ^2	df	sig. ^a
LRMSD	63.319	2	< 0.001	27.761	2	< 0.001
	Dunn's pairwise test		adj. sig. ^b	Dunn's pairwise test		adj. sig. ^b
	Type I / Type II		< 0.001	Type I / Type II		< 0.001
	Type I / Type III		< 0.001	Type I / Type III		< 0.001
	Type II / Type III		0.004	Type II / Type III		1
ILD	F	df ^c	sig. ^a	χ^2	df	sig. ^a
	24.734	2 / 147	< 0.001	38.613	2	< 0.001
	Tuckey's post-hoc test		sig. ^a	Dunn's pairwise test		adj. sig. ^b
	Type I / Type II		< 0.001	Type I / Type II		< 0.001
	Type I / Type III		< 0.001	Type I / Type III		< 0.001
Type II / Type III		0.226	Type II / Type III		0.280	

^a statistical significance ($p \leq 0.05$)

^b statistical significance adjusted using the Bonferroni correction ($p \leq 0.05$)

^c degrees of freedom are shown in the form: Between Groups / Within Groups

4 Discussion

Measurement accuracy refers to the degree of closeness of a measured quantity to its true value, whereas precision relates to the closeness of agreement between replicate measurements (BIPM 2012). Most commonly, accuracy is a description of systematic errors induced by the applied measurement method and is usually established by repeatedly measuring some traceable reference standard. Nevertheless, such reference standard is not available when measuring skeletal material, since all measurement methods are prone to observer induced bias, which not only varies among measurement methods or equipment but also with the actual quantities being measured. The latter play an even more crucial role when measured quantities rely on landmark acquisition from skeletal material. When assessing measurement error, there are two different conceptual approaches: repeated measures and comparisons with a “gold standard” (Fruciano 2016). In this context, the present study aims to evaluate the accuracy of each measurement method as a between-methods comparison rather than a comparison to a certain method that one might consider a “gold standard”.

Skeletal work involving measurements is usually accompanied with intra- and inter-observer error or other repeatability/reproducibility measures, which in most cases are adequate for assessing both the accuracy and precision of the method applied. The present work is not concerned about which method is more accurate or precise for landmark digitization per se, but aims to answer whether the underlying systematic errors for each measurement method are small enough and consistent so one may use these methods, hence their data, interchangeably. This is important in osteometric applications, since it provides an insight whether it is appropriate to apply regression equations or discriminant functions produced with one measurement method to a data set measured with a different method or even if it is acceptable to pool data sets measured with different methods. Most importantly, the present study aims to determine what is the additional variability that is introduced to such data when pooled together.

Most often studies on landmark digitization error rely on repetitive measurements over a very small sample for assessing accuracy and repeatability (von Cramon-Taubadel et al. 2007; Ross & Williams 2008; Tallman et al. 2014; Toneva et al. 2016; Shearer et al. 2017; Toneva et al. 2017). Apart from the intrinsic precision and accuracy of every measuring equipment, landmark digitization error has two fundamental contributors. One being the level of experience of the observer in accurately locating the landmarks and the other being the surface morphological variability of the cranial samples, which directly affects the difficulty for any given observer to accurately locate a landmark on the cranial surface and of course this difficulty varies per individual sample. The latter is a mere empirical observation that became apparent during the digitization process, since the experi-

mental method of the present study has not been designed to assess this source of systematic error. Nevertheless, the sample size was large enough to allow for such variation to be integrated in the collected data and subsequently expressed in our findings.

4.1 Inter-method error

The results of the present study regarding inter-method error show that all three tested methods for acquiring linear measurements introduce similar amounts of measurement error in the collected data. Moreover, the average between-methods relative TEMs for distance measurements fall within the acceptable range of inter-observer error regarding skillful anthropometrists (Perini et al. 2005), which are widely adapted in osteometric studies (Goto & Mascie-Taylor 2007; Langley et al. 2018). This implies that combining data sets of linear measurements collected with different methods could be a viable option in terms of the inter-method induced error. Our results show that error-induced variation is more related to the measured variable rather than the measuring method. For example, the maximum difference in TEM between different measurement methods of linear measurements D1 and D2 are 0.4 mm and 0.42 mm, respectively. However, relative TEM for D1 is considered unacceptable (1.8–2.7%), whereas for D2 it is considered very small (0.7–1%). Furthermore, the relative TEMs associated with angular measurements are even higher, which is expected since they require three distance measurements when measured with a caliper or three digitized landmarks when calculated from landmark coordinates. Nevertheless, even with angular measurements the inter-method error variation is less than the error observed across different measured variables.

Robinson & Terhune (2017) performed similar analysis on linear measurements acquired with a Mitutoyo digital caliper, a Microscribe-3DX digitizer and 3D models produced with Next Engine Laser scanner and microCT scans from 14 adult anthropoid crania. Despite the differences in the utilized samples and type of measurements with respect to the present study as well as the different analytical approach in calculating measurement error, they also found that all aforementioned methods result in fairly minor differences in terms of measurement acquisition error, although they observed that collecting linear data from 3D scanned models rather than directly on specimens using calipers or a Microscribe reduces error. It should be noted that Robinson and Terhune evaluated inter-method related error by means of random effects hierarchical ANOVA tests, in which nested observations cannot be considered as independent. This implies that observed inter-method error differences could be an effect of observed differences in intra- or inter-observer errors per method. Other studies that have performed direct pairwise comparisons of linear measurements between different methods are also congruent with our findings. Muñoz-Muñoz et al. (2016) studied 10 skulls of adult specimens of the western European subspecies of the house mouse. Comparing

linear measurements from photogrammetric 3D models with those taken directly on the skulls of the studied specimens with a Mitutoyo digital caliper reported a strong association between model based ILDs and directly measured ILDs from the skulls. In addition, Fourie et al. (2011) reported no “clinical differences” between the accuracy and reliability of anthropometric measurements taken from cadaver heads using calipers and from models produced with 3D imaging systems (cone beam CT, Laser surface scanning and 3D stereo-photogrammetry). Congruent with the findings of the present study, they also concluded that anthropometric data obtained by any of these methods are capable of being combined or compared statistically (Fourie et al. 2011).

In the present study the measurement accuracy for the dial caliper is ± 0.03 mm and ± 0.23 mm for the Microscribe-3DX as rated by their respective manufacturers. Furthermore, considering the accuracy of the photogrammetric 3D models to be bound by the precision of the repetitive measurements used for scaling, extrapolating the typical scaling precision (median = 0.145%) onto the maximum ILD measured in our sample (~165 mm) provides some statistical measure of rated accuracy for the photogrammetric models (± 0.24 mm), which is practically identical to that of the Microscribe-3DX. This further justifies our conclusion that the observed variation in between-methods TEMs is more related to the measured variables and observer experience in properly identifying landmark positions rather than the measurement method itself. Of course, each method has pros and cons regarding its inherent difficulty in acquiring certain measurements or landmark coordinates (Robinson & Terhune 2017), but pooling datasets measured with different methods should not produce any significant statistical noise as long as intra- and inter-observer errors are in concordance across sub-samples. The results of the present study suggest that evaluating the reliability of each measured variable in terms of observer error is the most important test for controlling measurement error when pooling different datasets.

4.2 Inter-observer error

Evaluation of inter-observer error in the present study is limited to landmark acquisition on the photogrammetric 3D models, since landmark-based observer error with Microscribe has been conducted on the crania of the Athens Collection as part of previous work on cranial age-related changes (Chovalopoulou et al. 2017). However, inter-observer error was evaluated both for ILDs and actual 3D coordinates represented as landmark configurations for each type of landmark separately in order to provide a direct measure how geometric morphometrics and more traditional analyses based on linear measurements may be affected by landmark acquisition error when pooling datasets from different sources.

Our results showed lower inter-observer error for landmark acquisition on photogrammetric 3D models with mean values per type of landmark ranging from 0.47 mm to 0.72 mm

as opposed to previously reported values for Microscribe with mean inter-observer error of 0.87 mm (Chovalopoulou et al. 2017). The reduced error may be partially attributed to the benefits of locating landmarks on photogrammetric models, where the capacity of rotating and focusing on a high quality textured mesh model allows much more precise landmark placement, which is most evident for type I landmarks. However, it should be noted that the reduced error is also influenced by technical differences between the two studies. Chovalopoulou et al. (2017) assessed inter-observer error as the mean standard deviation on 31 mixed-type non-reference landmarks (method proposed by von Cramon-Taubadel et al. 2007) between observers with different experience levels with Microscribe digitizer. In the present study, the inter-observer error is evaluated between equally well-experienced observers by super-imposing configurations with single type of landmarks for each paired sample. Nevertheless, the averaged inter-observer error is in the sub-millimeter range in both cases suggesting that both landmark acquisition methods, namely Microscribe-3DX and photogrammetric 3D models, share the same levels of accuracy and precision with the latter perhaps taking the lead. The benefits of photogrammetric 3D models regarding landmark digitization error have been also appraised by previous work (Muñoz-Muñoz et al. 2016). On the other hand, various researchers have reported reduced measurement precision on 3D models derived from NextEngine or CT scans as opposed to Microscribe-3DX data (Sholts et al. 2011; Robinson & Terhune 2017). CT scan based models have no texture and NextEngine scans usually result in poor textural information and noisy surface models as opposed to the high-quality texture present in photogrammetric 3D models, which improve their visualization and observer’s capacity to identify certain structures such as cranial sutures. This may provide some explanation over these differences with the findings of the present study.

3D models have been systematically used for acquiring landmark coordinates in geometric morphometrics studies, which benefit from the inherent advantages of 3DGM in quantitatively analyzing the shape of biological structures (Bookstein 1991). The ever increasing availability of 3D modeling techniques and model samples over the past few years resulted in researchers utilizing 3D models to derive osteometric data directly from 3D models rather than acquiring such measurements with calipers and other traditional osteometric tools. Hence, measurement error was evaluated separately for landmark coordinates and ILDs in the present study. Despite the latter being directly calculated from the former by simple Euclidean geometry, ILDs possess distinct properties as measured variables (distances vs point coordinates), which is also evident in our results with ILDs systematically showing significantly less inter-observer error than landmark coordinates (see Tables 4, 5, 6) but also less inter-method error as illustrated in Fig. 2. For the most part this distinction relates to the overall geometry of the specimen and is most evident with type III landmarks. For example,

the ILD of the bilateral *coronale* landmarks is less affected by the accuracy of landmark positions, since locally both sides of the cranial surface may be regarded as being approximately parallel and landmark positioning error is bound by the cranial surface. Additionally, the clustering of landmarks in landmark configurations along with the inevitable dispersion of sampling error on superimposed configurations, which is known as the Pinocchio effect (Chapman 1990; von Cramon-Taubadel et al. 2007; Klingenberg 2013), also influences the measured observer error.

4.3 Inter-type of landmark differences in observer error

Many researchers have reported on differential observer error related to the three types of landmarks (e.g. Williams & Richtsmeier 2003; von Cramon-Taubadel et al. 2007; Ross & Williams 2008; Barbeito-Andrés et al. 2012). The consensus is that landmark acquisition is most precise for type I and least precise for type III. An exception has been previously reported by Sholts et al. (2011) regarding NextEngine scanned models for which they reported highest precision for type III landmarks and lowest for type I, although their results regarding 3D digitizer measurements were in line with previous studies. They argued that their results were “readily explained by the extensive geometric tools available in software for analyzing 3D laser scanner models”, although they did not use such tools for locating type III landmarks in the sampling procedures. Furthermore, their results regarding type I landmarks on 3D models appear to be heavily influenced by landmark acquisition error on at least two of their samples and by one of the observers (referring to tables 2 and 3 in Sholts et al. 2011), suggesting that imprecise acquisition of type I landmarks might have been a result of poor quality laser scan models. Criticism about poor quality 3D models produced by laser scanners and especially for NextEngine scans has been previously reported (Polo & Felicísimo 2012; Robinson & Terhune 2017).

The results of the present study show that type I landmarks can be digitized with the highest precision on photogrammetric 3D models as demonstrated by inter-observer error values on paired landmark configurations and ILDs. This is in line with previous findings and is mostly attributable to the high quality texture that is inherent in photogrammetric 3D models. As in previous studies, type III landmarks exhibited lower accuracy than type I, but inter-observer error was statistically indifferent between type II and type III landmarks both for LRMSD of landmark configurations and ILDs. On the other hand, between-methods comparison revealed statistically significant difference between type II and type III landmarks for LRMSD of landmark configurations, suggesting that type III landmarks are more difficult to digitize with a Microscribe. This may relate to the restriction of 3D digitizers requiring specimens to be static throughout landmark acquisition, which often hinders the capacity of the observer to accurately locate extremal points such as

type III landmarks. However, such restriction does not seem to introduce any further measurement error when utilizing ILDs.

Conclusions

Overall, our results on the digitization accuracy between different types of landmarks from photogrammetric 3D models suggest that sharing/combining 3D datasets from different sources is a safe option without compromising the statistical integrity of the pooled dataset in subsequent analyses. Moreover, pooling ILD datasets from multiple methods, such as caliper, Microscribe, or photogrammetric 3D models, is also a statistically viable option. Nevertheless, compiling 3D datasets from multiple methods that are intended for 3DGM analysis should be done cautiously, as previously suggested (Sholts et al. 2011; Fruciano et al. 2017; Robinson & Terhune 2017). Our findings that ILDs are more immune to inter-observer/inter-method error further suggest that applying ILD-based discriminant methods with various measurement techniques or equipment is also statistically safer than utilizing 3DGM for analogous work. Finally, we conclude that properly conducted intra- and inter-observer error in any osteometric work is paramount and researchers should exercise and report, whether combining datasets or evaluating methods on samples with different equipment.

References

- Barbeito-Andrés, J., Anzelmo, M., Ventrice, F., & Sardi, M. L. (2012). Measurement error of 3D cranial landmarks of an ontogenetic sample using Computed Tomography. *Journal of Oral Biology and Craniofacial Research*, 2(2), 77–82. <https://doi.org/10.1016/j.jobcr.2012.05.005> PMID:25737840
- Bertsatos, A., Papageorgopoulou, C., Valakos, E., & Chovalopoulou, M. E. (2018). Investigating the sex-related geometric variation of the human cranium. *International Journal of Legal Medicine*, 132(5), 1505–1514. <https://doi.org/10.1007/s00414-018-1790-z> PMID:29380124
- BIPM. (2012). International vocabulary of metrology – Basic and general concepts and associated terms (VIM) (3rd ed.). *JCGM 200*, 1012.
- Bookstein, F. (1991). *Morphometric Tools for Landmark Data: Geometry and Biology*. Cambridge, UK: Cambridge University Press.
- Chapman, R. E. (1990). Conventional Procrustes approaches. In F. J. Rohlf, & F. L. Bookstein (Eds.), *Proceedings of the Michigan Morphometrics Workshop. Spec. Publ. 2* (pp. 251–267). Ann Arbor: The University of Michigan Museum of Zoology.
- Chovalopoulou, M.-E., Bertsatos, A., & Papageorgopoulou, C. (2017). Age-related changes in the craniofacial region in a modern Greek population sample of known age and sex. *International Journal of Legal Medicine*, 131(4), 1103–1111. <https://doi.org/10.1007/s00414-016-1470-9> PMID:27757579
- Eaton, J. W., Bateman, D., Hausberg, S., & Wehbring, R. (2016). GNU Octave version 4.2.0 manual: a high-level interactive

- language for numerical computations. <http://www.gnu.org/software/octave/doc/interpreter/>
- Fourie, Z., Damstra, J., Gerrits, P. O., & Ren, Y. (2011). Evaluation of anthropometric accuracy and reliability using different three-dimensional scanning systems. *Forensic Science International*, 207(1-3), 127–134. <https://doi.org/10.1016/j.forsciint.2010.09.018> PMID:20951517
- Fruciano, C. (2016). Measurement error in geometric morphometrics. *Development Genes and Evolution*, 226(3), 139–158. <https://doi.org/10.1007/s00427-016-0537-4> PMID:27038025
- Fruciano, C., Celik, M. A., Butler, K., Dooley, T., Weisbecker, V., & Phillips, M. J. (2017). Sharing is caring? Measurement error and the issues arising from combining 3D morphometric datasets. *Ecology and Evolution*, 7(17), 7034–7046. <https://doi.org/10.1002/ece3.3256> PMID:28904781
- Goto, R., & Mascie-Taylor, C. G. (2007). Precision of measurement as a component of human variation. *Journal of Physiological Anthropology*, 26(2), 253–256. <https://doi.org/10.2114/jpa2.26.253> PMID:17435374
- JASP Team. (2018). *JASP (Version 0.8.6)* [Computer software]. <https://jasp-stats.org/>
- Kabsch, W. (1976). A solution for the best rotation to relate two sets of vectors. *Acta Crystallographica. Section A, Crystal Physics, Diffraction, Theoretical and General Crystallography*, 32(5), 922–923. <https://doi.org/10.1107/S0567739476001873>
- Katz, D., & Friess, M. (2014). Technical note: 3D from standard digital photography of human crania—a preliminary assessment. *American Journal of Physical Anthropology*, 154(1), 152–158. <https://doi.org/10.1002/ajpa.22468> PMID:24711122
- Klingenberg, C. P. (2013). Visualizations in geometric morphometrics: how to read and how to make graphs showing shape changes. *Hystrix, the Italian Journal of Mammalogy*, 24(1), 15–24. doi:10.4404/hystrix-24.1-7691
- Langley, N. R., Meadows Jantz, L., McNulty, S., Maijanen, H., Ousley, S. D., & Jantz, R. L. (2018). Error quantification of osteometric data in forensic anthropology. *Forensic Science International*, 287, 183–189. <https://doi.org/10.1016/j.forsciint.2018.04.004> PMID:29698914
- Lewis, S. J. (1999). Quantifying measurement error. In S. Anderson (Ed.), *Current and recent research in osteoarchaeology 2: proceedings of the 4th, 5th and 6th meetings of the Osteoarchaeological research group* (pp. 54–55). Oxford: Oxbow Books.
- Muñoz-Muñoz, F., Quinto-Sánchez, M., & González-José, R. (2016). Photogrammetry: A useful tool for three-dimensional morphometric analysis of small mammals. *Journal of Zoological Systematics and Evolutionary Research*, 54(4), 318–325. <https://doi.org/10.1111/jzs.12137>
- Perini, T. A., de Oliveira, G. L., Ornellas, J. S., & de Oliveira, F. P. (2005). Technical error measurement in anthropometry. *Revista Brasileira de Medicina do Esporte*, 11(1), 86–90. <https://doi.org/10.1590/S1517-86922005000100009>
- Polo, M. E., & Felicísimo, Á. M. (2012). Analysis of uncertainty and repeatability of a low-cost 3D laser scanner. *Sensors (Basel)*, 12(7), 9046–9054. <https://doi.org/10.3390/s120709046> PMID:23012532
- Robinson, C., & Terhune, C. E. (2017). Error in geometric morphometric data collection: Combining data from multiple sources. *American Journal of Physical Anthropology*, 164(1), 62–75. <https://doi.org/10.1002/ajpa.23257> PMID:28573732
- Ross, A. H., & Williams, S. (2008). Testing repeatability and error of coordinate landmark data acquired from crania. *Journal of Forensic Sciences*, 53(4), 782–785. <https://doi.org/10.1111/j.1556-4029.2008.00751.x> PMID:18537868
- Shearer, B. M., Cooke, S. B., Halenar, L. B., Reber, S. L., Plummer, J. E., Delson, E., & Tallman, M. (2017). Evaluating causes of error in landmark-based data collection using scanners. *PLoS One*, 12(11), e0187452. <https://doi.org/10.1371/journal.pone.0187452> PMID:29099867
- Sholts, S. B., Flores, L., Walker, P. L., & Wärrmländer, S. K. (2011). Comparison of coordinate measurement precision of different landmark types on human crania using a 3D laser scanner and a 3D digitiser: Implications for applications of digital morphometrics. *International Journal of Osteoarchaeology*, 21(5), 535–543. <https://doi.org/10.1002/oa.1156>
- Sládek, V., Galeta, P., & Sosna, D. (2012). Measuring human remains in the field: Grid technique, total station, or MicroScribe? *Forensic Science International*, 221(1-3), 16–22. <https://doi.org/10.1016/j.forsciint.2012.03.018> PMID:22560958
- Spradley, M. K., & Jantz, R. L. (2016). Ancestry Estimation in Forensic Anthropology: Geometric Morphometric versus Standard and Nonstandard Interlandmark Distances. *Journal of Forensic Sciences*, 61(4), 892–897. <https://doi.org/10.1111/1556-4029.13081> PMID:27364267
- Tallman, M., Amenta, N., Delson, E., Frost, S. R., Ghosh, D., Klukkert, Z. S., ... Sawyer, G. J. (2014). Evaluation of a new method of fossil retrodeformation by algorithmic symmetrization: Crania of papionins (Primates, Cercopithecidae) as a test case. *PLoS One*, 9(7), e100833. <https://doi.org/10.1371/journal.pone.0100833> PMID:24992483
- Toneva, D., Nikolova, S., & Georgiev, I. (2016). Reliability and Accuracy of Angular Measurements on Laser Scanning Created 3D Models of Dry Skulls. *Journal of Anthropology*, 2016, 1–6. <https://doi.org/10.1155/2016/6218659>
- Toneva, D., Nikolova, S., Georgiev, I., & Tchorbadijeff, A. (2017). Accuracy of Linear Craniometric Measurements Obtained from Laser Scanning Created 3D Models of Dry Skulls. In K. Georgiev, M. Todorov, & I. Georgiev (Eds.), *Advanced Computing in Industrial Mathematics*. Studies in Computational Intelligence, vol. 681 (pp. 215–229). Cham: Springer. doi:10.1007/978-3-319-49544-6_18
- von Cramon-Taubadel, N., Frazier, B. C., & Lahr, M. M. (2007). The problem of assessing landmark error in geometric morphometrics: Theory, methods, and modifications. *American Journal of Physical Anthropology*, 134(1), 24–35. <https://doi.org/10.1002/ajpa.20616> PMID:17503448
- Williams, F. L., & Richtsmeier, J. T. (2003). Comparison of mandibular landmarks from computed tomography and 3D digitizer data. *Clinical Anatomy (New York, N.Y.)*, 16(6), 494–500. <https://doi.org/10.1002/ca.10095> PMID:14566895

Manuscript received: 03 April 2019

Revisions requested: 23 August 2019

Revised version received: 04 October 2019

Accepted: 04 October 2019



Testing the reliability of 3D-ID software in sex and ancestry estimation with a modern Greek sample



Andreas Bertatos, Aikaterini Christaki, Maria-Eleni Chovalopoulou*

National and Kapodistrian University of Athens, School of Sciences, Faculty of Biology, Department of Animal and Human Physiology, Panepistimiopolis, GR 157 01, Athens, Greece

ARTICLE INFO

Article history:

Received 24 October 2018

Received in revised form 31 January 2019

Accepted 5 February 2019

Available online 14 February 2019

Keywords:

3D-ID software

Classification performance

Sex estimation

Ancestry estimation

Greek population sample

ABSTRACT

A primary concern in forensic anthropology, when reconstructing the biological profile of an unidentified individual is ancestry and sex estimation. The development of multivariate statistical methods and the assembly of large reference sample databases gave rise to the development of specialized computer software for sex and ancestry estimation. Among various such software, the 3D-ID is the only freely-available program that can handle missing values in the input dataset. The present study evaluates the reliability of 3D-ID in correctly classifying ancestry and sex of 158 test subjects from the Athens Collection, a documented Greek population sample. 3D-ID's classification performance was evaluated both separately and collectively for sex and ancestry. According to our results, the accuracy regarding sex estimation ranged from 74.05% to 86.7% for cases with unknown ancestry and reached 89.87% when testing within the Southeastern European reference group, whereas ancestry estimation accuracy reached 70.9% for correctly classifying the Greek individuals to European population groups.

We conclude that 3D-ID software exhibits moderate reliability in ancestry estimation and adequate reliability in sex estimation. The Greek population seems to deviate from the 3D-ID software's reference samples and therefore caution should be taken in interpreting 3D-ID's results of unknown subjects, for which the software's reference sample database may not be representative. 3D-ID's guidelines for using 19-landmark configuration improves the accuracy of ancestry estimation and form variables should be preferred for sexing samples.

© 2019 Elsevier B.V. All rights reserved.

1. Introduction

Forensic anthropology may be defined as the application of skeletal biology to issues of the law [1]. When human remains are found, the primary objective in unidentified cases is the reconstruction of the biological profile of the individual represented. Estimation of sex, ancestry, age as well as stature, is fundamental in the reconstruction of a biological profile. However, sex and ancestry have to be initially determined, since stature and age are population and sex specific [2,3].

Although the os coxae have accurate criteria in predicting sex [4,5], they are fragile and therefore their conservation is poor [6]. Hence sex estimation also relies on the skull, which is often available, since it is more resistant to taphonomic phenomena [7]. Furthermore, DNA analysis, although established as sex estimation method from skeletal material, it cannot be applied in all cases,

either because it may not be recoverable or due to contamination bias [8–10]. In addition to cranial-based sexing, it has been demonstrated that the human skull, especially the midface, is widely used for estimating ancestry compared to the other skeletal elements [11].

The development of multivariate statistical methods in combination with the increasing availability of large reference sample skeletal records allowed the quantitative analysis for ancestry and sex estimation among different population groups. This in turn gave rise to the development of specialized computer software for sex and ancestry estimation of unidentified cranial remains, such as 3D-ID [12], FORDISC 3 [13] and CRAN-ID [14]. Although quantitative methods exhibit greater objectivity in comparison to qualitative ones [15], the performance of these software programs in terms of correct classification results is still under question, with various researchers reporting poor to mediocre performance in cases where the ancestry group of the test subjects is not included in the software's reference database [9,10,16–18]. As a consequence of ancestry misclassification, sex can also be misclassified as well since various populations have different sexually dimorphic levels [19].

* Corresponding author.

E-mail addresses: abertatos@biol.uoa.gr (A. Bertatos), aichristaki@gmail.com (A. Christaki), marilenachov@biol.uoa.gr (M.-E. Chovalopoulou).

The present study focuses on the 3D-ID software. In contrast to its alternatives (FORDISC and CRAN-ID) which rely on traditional craniometric variables, 3D-ID is the only implementation that utilizes 3-dimensional geometric morphometrics in order to estimate sex and ancestry of unknown individuals based on their cranial morphology. Additionally, 3D-ID is a cross-platform package and it is the only freely-available program that can handle missing values in the input dataset. Both features provide significant merits to its userbase. The user provides the program with the three dimensional landmark coordinates of the unknown cranium and the program constructs classification functions based on the available reference groups derived from a database of over 2000 individuals. Our aim is to assess the reliability of the latest version of the 3D-ID (version 2018) software in correctly classifying ancestry and sex using landmark configurations from a documented Greek population sample as the test subjects. The 3D-ID software does not include any Greek population samples in its reference database, which hinders its capability in correct ancestral classification. Nevertheless, it contains a southeastern European classification group, which geographically, at least, represents the population under test. In forensic practice the unidentified sample at hand will not necessarily be represented in the software's reference database and since 3D-ID has not been tested before against Greek population samples, this rises the opportunity to test its performance on a worst-case scenario. That is testing 3D-ID's performance on samples that are not directly represented in its reference database.

2. Materials and Methods

The present study uses 158 crania (80 males and 78 females without any evidence of pathological deformation) of adult individuals from the Athens Collection, which is housed in the Department of Animal and Human Physiology of the National & Kapodistrian University of Athens. Information on the sex, age at death, occupation, cause of death and place of birth of each individual in the collection is derived from death records [20]. Year of birth for all individuals ranges between 1879 and 1965 while their age-at-death ranges from 19 to 99 years old.

Although 3D-ID's manual [12] defines thirty-four landmarks that can be analyzed with the software, the authors recommend in their website (www.3d-id.org/download-3d-id) that certain landmarks should be excluded from the analysis due to high inter-observer error. Hence, landmark configurations of 3D coordinates of up to twenty-five landmarks (Table 1), based on each sample's availability, were used in the present study. However, all thirty-four landmark coordinates were extracted from high quality textured 3D models with Meshlab software [21]. The 3D models have been produced by 3D photogrammetry software (Photoscan Pro v1.4, Agisoft LLC) from the respective crania of the Athens Collection with a calibrated accuracy of 0.2 mm. Landmark acquisition was performed by MEC and visual re-inspection of placed landmarks by AB. Furthermore, the data integrity of the

extracted coordinates have been tested in GNU Octave [22] with a custom routine to ensure that there are no misplaced landmarks in the dataset that would bias the classification results of 3D-ID. The ".3did" files containing the landmark configuration for each specimen have also been produced with GNU Octave script. The complete landmark dataset [23] along with the associated Meshlab (.pp) and 3D-ID (.3did) files and the relevant GNU Octave source code is available to download from zenodo repository.

3D-ID software classifies each unknown specimen to a specific ancestry and sex group collectively. However, it is possible to separate these two classifications, if the user indicates the sex or the ancestry group of the unknown specimen. In the present study, the correct classification performance of 3D-ID software was evaluated both when sex and ancestry are predicted collectively as well as when they are predicted separately. Additionally, classification performance was also tested with respect to shape as well as form variables, which emerge from the Procrustes analysis of the landmark configurations including the centroid size. The aforementioned classification performance tests were evaluated through two distinct groups based on the number of landmarks used for each unknown sample's configuration. The first group comprised of up to 25-landmark configurations, as suggested by the author's of 3D-ID software, for collectively predicting sex and ancestry (full reference sample), but also for evaluating ancestry classification of individuals with known sex (sex determined reference sample). The second group utilized a reduced set of up to 19-landmark configurations (Table 1) to perform the aforementioned tests respectively. The reduced landmark configurations were applied in order to increase the reference sample size of the European populations according to 3D-ID's guidelines. Furthermore, the 19-landmark configurations were used to evaluate the software's performance when considering only European reference samples for collectively predicting sex and ancestry, but also for predicting sex by specifying Southeastern European ancestral group for each specimen. Finally, 3D-ID's performance for predicting sex in both full reference and European reference samples as well as with full and reduced landmark configurations was measured independently from each test subject's classified ancestral group. Performance of 3D-ID software was calculated as the percentage of correctly classified samples. Additionally, it should be noted that the accuracy of ancestry prediction of the Greek population test sample is evaluated both in terms of being classified to Southeastern European group as well as collectively to European group.

3. Results

The results of the classification scores, when all groups of the 3D-ID's reference data set are included, are summarized in Table 2. Regarding classification of the Greek sample as Southeastern European the overall accuracy both for shape and form variables is improved when the 19-landmark configuration is used, since in 25-landmarks configuration all specimens were misclassified. When considering classification to the European group collectively, the 25-landmark configuration yielded 61.4% correct classification using shape variables and 63.9% using form variables, but the 19-landmark configuration still performs better for both shape and form variables, with 70.3% and 70.9% respectively, although only half of the classified samples as being Europeans is correctly classified in the Southeastern European group (34.2%). Remarkably, the 25-landmark configuration misclassified all specimens predicted being Europeans into the Southwestern European group instead of the Southeastern European group. It is also noted that form variables marginally perform better than shape variables.

Correct classification scores regarding sex follow an analogy to ancestry prediction in the sense that both 19-landmark

Table 1
List of digitized landmarks.

1, 2	Asterion L & R	15	Nasion
3	Basion	16*	Left Inferior Orbital Border
4	Bregma	17*	Left Superior Orbital Border
5, 6	Dacryon L & R	18*	Opisthion
7, 8	Ectoconchion L & R	19*	Subspinale
9, 10	Frontomale Anterior L & R	20, 21	Zygion L & R
11, 12	Frontomale Temporale L & R	22, 23	Zygomaxillare L & R
13	Glabella	24, 25	Zygoorbitale L & R
14	Lambda		

* excluded when analyzing explicitly for European ancestry groups.

Table 2

Software performance results regarding both sex and ancestry classification when the analysis contains all groups of reference samples.

Population Group ^a	Number of individuals			
	25-landmark configuration		19-landmark configuration	
	Shape variables	Form variables	Shape variables	Form variables
African	0	0	0	1 (0.6)
African_American	0	0	2 (1.3)	3 (1.9)
African_Brazilian	0	0	1 (0.6)	1 (0.6)
Brazilian	0	0	11 (7.0)	8 (5.1)
Circumcaribbean	0	0	3 (1.9)	5 (3.2)
Colombian	28 (17.7)	25 (15.8)	11 (7.0)	8 (5.1)
East_Asian	7 (4.4)	6 (3.8)	1 (0.6)	1 (0.6)
East_African	0	0	0	0
European_American	10 (6.3)	10 (6.3)	8 (5.1)	10 (6.3)
European_Central	0	0	43 (27.2)	43 (27.2)
European_Eastern	0	0	0	0
European_Southeastern	0	0	54 (34.2)	54 (34.2)
European_Southwestern	97 (61.4)	101 (63.9)	14 (8.9)	15 (9.5)
Japanese_Brazilian	0	0	1 (0.6)	1 (0.6)
Mesoamerican	0	0	1 (0.6)	1 (0.6)
Nigerian	0	0	0	0
South_American	11 (7.0)	12 (7.6)	7 (4.4)	6 (3.8)
Syrian	5 (3.2)	4 (2.5)	1 (0.6)	1 (0.6)
West_African	0	0	0	0
Sex Classification^b	117 (74.1)	126 (79.7)	120 (75.9)	135 (85.4)

^a Integer numbers correspond to the number of specimens classified in each population group of the reference sample. The values in parentheses show the respective percentage from the total sample size of 158 individuals.

^b Number of correctly classified specimens (and their respective percentage) concerns only correct sex prediction regardless of population group classification.

configurations and form variables outperform 25-landmark configurations and shape variables respectively. In particular, the 25-landmark configuration provides a sex classification accuracy of 74.1% when using shape variables and an accuracy of 79.7% when form variables are used. The classification rates of the 19-landmark configuration are 75.9% and 85.4% respectively.

Regarding the correct classification for both sex and ancestry collectively, when all the European population groups are considered as one group, the overall accuracy both for shape and form variables is also improved when the 19-landmark configuration is used. The 25-landmark configuration yields a correct classification of 50.63% when shape variables are used and 53.16% when form variables are used, while the corresponding scores for the 19-landmark configuration are 52.53% and 60.76% respectively.

The ancestry classification scores when testing individuals with known sex (sex determined reference sample) are summarized in Table 3. The results are very similar to Table 2 with respect to correct classification into European groups. The only notable difference lies with the distribution of misclassified specimens following a slightly different pattern.

Software's performance improves regarding both sex and ancestry classification when the analysis contains only European reference samples. More specifically, with respect to ancestry classification, the 43.7% of the Greek population sample is correctly classified as Southeastern Europeans when shape variables are used and the corresponding correct classification score for form variables is 46.8% (Table 4). It should be noted that none of the test subjects were classified Eastern Europeans. Regarding sex classification, the accuracy for shape variables is 76.6% and the accuracy for form variables is 86.7%. Finally, utilizing 19-landmark configuration for estimating sex with known ancestry (i.e. Southeastern European group) produced correct classification accuracy of 79.95% with shape variables and 89.87% with form variables.

Detailed results including reference sample size, D^2 , Posterior probability, and Typicality values for each tested individual from the Athens Collection is provided as supplementary material.

Reporting descriptive statistics on reference sample size and D^2 values along with the aforementioned classification results does not allow any statistically meaningful comparison between correctly classified and misclassified cases, since the varying availability of landmarks across the tested specimens directly affects reference sample size and especially D^2 values. However, it worths mentioning that D^2 values for classified specimens ranged between ~ 17.3 and ~ 163.6 throughout all analyses, which further corroborates the data integrity of the digitized 3D landmark coordinates of the tested specimens, since they fall within the normal range of variation for humans.

4. Discussion

The purpose of this study was to assess the reliability of 3D-ID software in estimating ancestry and sex of unknown subjects. Overall, the software exhibited moderate classification performance. Regarding ancestry estimation, our results show that the use of 19-landmark configuration enhanced accuracy, which supports the programmer's guideline to exclude specific landmarks in order to increase the reference sample size when dealing with European samples. 3D-ID is designed to force an assignment of the unknown specimen to the software's reference ancestry group, which it differs less from [12]. Hence, the poor to moderate ancestry classification accuracy may be attributed to the fact that 3D-ID's performance was tested against a poorly referenced sample with respect to the software's database. Although we cannot draw any definite conclusion based on our existing results that 3D-ID would have produced better classification results with the present test sample if the Greek population were well represented in its database, it may be considered an acceptable assumption that classifying test subjects from referenced populations would yield better classification scores.

One could argue that testing 3D-ID's classification accuracy, when knowingly classifying individuals not represented in the software's reference population, is not a valid study design. Nevertheless, the rationale behind testing an unreferenced sample is that the present study evaluates 3D-ID's performance in

Table 3

Software performance results regarding and ancestry classification with sex determined groups of reference samples.

Population Group ^a	Number of individuals			
	25-landmark configuration		19-landmark configuration	
	Shape variables	Form variables	Shape variables	Form variables
African	0	0	0	0
African_American	0	1 (0.6)	4 (2.5)	4 (2.5)
African_Brazilian	0	0	3 (1.9)	3 (1.9)
Brazilian	0	0	9 (5.7)	8 (5.1)
Circumcaribbean	1 (0.6)	1 (0.6)	5 (3.2)	4 (2.5)
Colombian	24 (15.2)	25 (15.8)	6 (3.8)	6 (3.8)
East_Asian	3 (1.9)	2 (1.3)	1 (0.6)	1 (0.6)
East_African	0	0	0	0
European_American	15 (9.5)	11 (7.0)	7 (4.4)	11 (7.0)
European_Central	0	0	42 (26.6)	42 (26.6)
European_Eastern	0	0	0	0
European_Southeastern	0	0	56 (35.4)	58 (36.7)
European_Southwestern	97 (61.4)	99 (62.7)	14 (8.9)	13 (8.2)
Japanese_Brazilian	0	0	0	0
Mesoamerican	0	0	1 (0.6)	1 (0.6)
Nigerian	0	0	0	0
South_American	14 (8.9)	13 (8.2)	10 (6.3)	7 (4.4)
Syrian	4 (2.5)	6 (3.8)	0	0
West_African	0	0	0	0

^a Integer numbers correspond to the number of specimens classified in each population group of the reference sample. The values in parentheses show the respective percentage from the total sample size of 158 individuals.

correctly classifying unknown individuals and by extent its utility in the forensic context. This means that an unidentified case is just that, unknown. In forensic practice, this translates to that there is no prior knowledge whether an unidentified case will be represented or not in the software's reference population. Utilizing 3D-ID or any other counterpart for estimating ancestry in such context is an exploratory process towards identification and not confirmation of known ancestral group. Of course, 3D-ID's performance evaluation is not limited into a mere correct/incorrect classification of the tested individuals into the Southeastern European group. It is also of interest what groups they are being classified into, which provides some insight to the phenotypic biological relations among populations as expressed by the cranial morphometric variation in 3D-ID's reference database, but also poses some questioning regarding the partitioning of ancestral groups.

One observation that stands out from the classification scores of the 25-landmark configuration in Table 2 is that although all tested individuals were misclassified, when considering that Greek population belongs to the Southeastern European group, about 70% were classified of European descent belonging either to the European_Southwestern or European_American groups. Moreover, most of the remaining misclassified sample was largely

assigned to Colombian and South American groups. On the other hand, when utilizing 19-landmark configurations, similar percentage is classified as European Americans (6%) and approximately another 70% is spread across more European subgroups instead of explicitly Southwestern Europeans, whereas samples previously classified as Colombians or South Americans are almost evenly spread across most ancestral groups present in 3D-ID population reference database. This pattern becomes even more evident when classifying ancestry based on known sex. However, the results based on sex determined reference samples imply that prior knowledge of sex of an individual does not improve ancestry classification.

Regarding the classification of Greek individuals into European groups but not explicitly as Southeastern Europeans, we could speculate that either Southeastern, Southwestern and Central European reference samples don't differ significantly regarding cranium variation and therefore discrimination between these three groups is marginal, or the cranial variation of the Greek population deviates from or overlaps those groups equally, while being more distant to all other reference samples. Regardless of what the actual case might be, we presume that if the reference database of 3D-ID included a Greek population sample in the European_Southeastern reference group, the software's performance would improve. Nevertheless, it would be safer to assume that repartitioning 3D-ID's reference population into wider ancestral groups would certainly provide much more reliable classification among Caucasians from other ancestral groups such as Africans or Asians.

The other point raised by the results of the present study concerns the distinct misclassification patterns between 25- and 19-landmark configurations. Considering that 1/6 of the Greek population sample is being classified as Colombian, when using 25-landmark configurations, one could argue that this is the result of phenotypic biological similarity to some extent between the two populations in the same manner that was earlier discussed for the European groups. On the other hand, classifying with 19-landmark configurations yields no such relation, which raises the question, whether such similarity could be configuration specific or just an artifact of geometric morphometric analysis with variable number of landmark configurations. 3D-ID guidelines about increasing

Table 4

Software performance results regarding both sex and ancestry classification when the analysis contains only European reference samples.

Population Group ^a	Number of individuals	
	19-landmark configuration	
	Shape variables	Form variables
European_Central	51 (32.3)	48 (30.4)
European_Eastern	0	0
European_Southeastern	69 (43.7)	74 (46.8)
European_Southwestern	38 (24.1)	36 (22.8)
Sex Classification^b	121 (76.6)	137 (86.7)

^a Integer numbers correspond to the number of specimens classified in each population group of the reference sample. The values in parentheses show the respective percentage from the total sample size of 158 individuals.

^b Number of correctly classified specimens (and their respective percentage) concerns only correct sex prediction regardless of population group classification.

sample size availability with 19-landmark configuration mainly concerns the European and Brazilian samples, which indicates that misclassification in other groups cannot be attributed to small sample. Would the same test subject be classified differently under different set of landmark configurations? If so, which specific configuration should be used? Unfortunately, these questions cannot be addressed in a feasible manner, since 3D-ID software does not provide any batch processing features or an external application programming interface.

Regarding sex estimation, 3D-ID's correct classification accuracy increases when form variables are used instead of shape variables, which is in agreement with the argument that the majority of the variation between males and females is size-related [24]. Using shape variables for sex estimation hardly reached 76.6% accuracy, which is quite low for forensic practice. Therefore, if 3D-ID is to be utilized for estimating sex, then the best practice would be to use form variables and the 19-landmark configuration regardless of the test subject's ancestral group, since both tests (that is across all groups and only European groups) yielded about 86% correct classification. Additionally, the highest observed sex classification score, when explicitly using European population reference groups and form variables from the 19-landmark configurations, is identical to previous research, which applied 3D geometric morphometrics on the upper face region of the same Greek population sample and reported sex estimation accuracy of 86.7% [25]. This agreement could be attributed to the suitability of the reduced set of landmarks when evaluating European samples, as 3D-ID's guidelines suggest. This is further supported by the results of the 25-landmark configuration, where sex classification accuracy decreases compared to the 19-landmark configuration. Moreover, the consistency of the cross-validated results regarding sex classification implies that the Greek population sample is well represented by the European groups in 3D-ID's reference database, at least with respect to the expression of sexual dimorphism. This further supports the aforementioned speculation regarding similarity across the European subgroups, which reduces the capability of 3D-ID to classify the Greek samples into the correct European subgroup, although it successfully assigns a reasonable percentage (~70%) into the European groups collectively.

Our results, when all the European population groups are considered as one group, are better regarding sex as well as ancestry estimation compared to previously reported results from other equivalent studies. In 2014, Urbanova and her colleagues [18] tested the accuracy of the FORDISC, COLIPR and 3D-ID software on a multi-ancestral Brazilian sample for estimating sex and ancestry. Their sample consisted of 174 documented human crania. According to their results, with respect to ancestry assessment, regardless of the software employed, approximately an accuracy of 50% was achieved. Additionally, tests of sex estimation showed moderate reliability with an accuracy ranging from 60 to 71%, while at the same time 3D-ID software showed a higher classification accuracy for males than females. In 2015, King also compared the accuracy rates of ancestry estimation of FORDISC 3.1 and 3D-ID software [26]. King's sample consisted of 385 South African individuals (186 black and 199 white) from the Pretoria Bone Collection. According to King's results, 3D-ID software has an overall 63.1% accuracy in estimating ancestry. However, it should be noted that Urbanova and her colleagues [18] as well as King [26], utilized a previous version of 3D-ID software with a much smaller reference population database, which signifies the importance and need for large reference population data for improving the classification performance, especially when estimating ancestry.

FORDISC [13] is an alternative software for estimating sex, ancestry and stature from skeletal material, which operates on traditional caliper-measured distances and has been evaluated more extensively than 3D-ID software. However, FORDISC does not

achieve high-reliability, especially for European samples. According to Guyomarc'h and Bruzek [10] as well as Ramsthaler and his colleagues [9], FORDISC cannot be used as a single method for the estimation of sex in European samples. Guyomarc'h and Bruzek evaluated FORDISC on French and Thai samples, while Ramsthaler and colleagues on a German sample. Furthermore, Elliott and Collard [17] concluded that FORDISC is problematic for ancestry determination after the evaluation of the software on various populations including a European. On the other hand, 3D-ID produced promising results, which makes it a useful exploratory tool in the forensic practice, although there is much space for further improvement. Nevertheless, we should point out that ancestral classification is not a binary classification problem such as sex estimation. Hence, classifying correctly 70% of unreferenced samples as Europeans yields some significance both in terms of 3D-ID's utility but also regarding the benefits of geometric morphometrics in ancestral classification as an analytic approach. On the other hand, sex estimation accuracy, which reaches a maximum of 89.9% when 3D-ID is used as population specific sex classifier, is not as high as recent studies with newly developed methods suggest, which produced cranial-based sex estimation on the same Greek population sample with cross-validated classification ranging from 91.5% [27] to 93.6% [28] and 95.2% [29]. This implies that 3D-ID, although adequate in performance, should not be used as a primary tool for estimating sex. Our suggestion is that 3D-ID's sex estimates should be cross-validated with available population specific methods based on its ancestral classification result considered with wider ancestral groups as previously discussed.

5. Conclusion

The application of a Greek sample to the 3D-ID software resulted in moderate reliability in ancestry estimation and adequate performance in sex estimation. The Greek population seems to deviate from the 3D-ID software's reference samples and therefore, as has been pointed out by other studies, caution should be taken in interpreting 3D-ID's results for unknown subjects, for which the software's reference sample database may not be representative. However, 3D-ID software could be utilized to successfully discriminate major ancestral groups. 3D-ID's guidelines for using 19-landmark configuration improves the accuracy of ancestry estimation and form variables should be preferred for sexing samples. Sex estimation should be cross-validated with population specific methods where possible.

6. Declarations of interest

None.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.forsciint.2019.02.004>.

References

- [1] D.A. Komar, J.E. Buikstra, *Forensic anthropology: Contemporary theory and practice*, Oxford University Press, Oxford, 2008.
- [2] A.H. Ross, E. Kimmerle, Contribution of quantitative methods in forensic anthropology: a new era, in: S. Blau, D. Ubelaker (Eds.), *Handbook of Forensic Anthropology and Archaeology*, Left Coast Press, Walnut Creek, 2009, pp. 479–489.
- [3] A.H. Ross, D.E. Slice, S.E. Williams, *Geometric Morphometric Tools for the Classification of Human Skulls (231195)*. Retrieved from, (2010). <https://www.ncjrs.gov/pdffiles1/nij/grants/231195.pdf>.
- [4] K.B. Burns, *Forensic Anthropology Training Manual*, Prentice-Hall, New Jersey, 1999.

- [5] S.K. Wärmländer, S.B. Sholts, Sampling and statistical considerations for the Suchey–Brooks method for pubic bone age estimation: Implications for regional comparisons, *Science & Justice* 51 (3) (2011) 131–134, doi:http://dx.doi.org/10.1016/j.scijus.2010.12.001.
- [6] T. Waldron, The relative survival of the human skeleton: implication for palaeopathology, in: A. Boddington, A. Garland, R. Janaway (Eds.), *Death, Decay and Reconstruction: Approaches to Archaeology and Forensic Science*, Manchester University Press, Manchester, 1987, pp. 55–64.
- [7] C.M. Stojanowski, R.M. Seidemann, G.H. Doran, Differential skeletal preservation at Windover Pond: Causes and consequences, *American Journal of Physical Anthropology* 119 (1) (2002) 15–26, doi:http://dx.doi.org/10.1002/ajpa.10101.
- [8] N. Wurmb-Schwark, H. Bosinski, S. Ritz-Timme, What do the X and Y chromosomes tell us about sex and gender in forensic case analysis? *Journal of Forensic and Legal Medicine* 14 (1) (2007) 27–30, doi:http://dx.doi.org/10.1016/j.jcfm.2005.09.003.
- [9] F. Ramsthaler, K. Kreutz, M.A. Verhoff, Accuracy of metric sex analysis of skeletal remains using Fordisc® based on a recent skull collection, *International Journal of Legal Medicine* 121 (6) (2007) 477–482, doi:http://dx.doi.org/10.1007/s00414-007-0199-x.
- [10] P. Guyomarc'h, J. Bruzek, Accuracy and reliability in sex determination from skulls: A comparison of Fordisc® 3.0 and the discriminant function analysis, *Forensic Science International* 208 (1–3) (2011), doi:http://dx.doi.org/10.1016/j.forsciint.2011.03.011 180.e1–180.e6.
- [11] T.D. White, M. Black, P.A. Folkens, *Human osteology*, Elsevier Acad. Press, Amsterdam, 2012 p. 424.
- [12] D.E. Slice, A. Ross, 3D-ID: geometric morphometric classification of crania for forensic scientists. Version 2018-05-27, (2009) . <http://www.3d-id.org>.
- [13] R.L. Jantz, S.D. Ousley, *FORDISC 3: Computerized Forensic Discriminant Functions. Version 3.0*, The University of Tennessee, Knoxville, 2005.
- [14] Richard Wright, Guide to using the CRANID programs Cr6bInd: for linear and nearest neighbours discriminant analysis Retrieved 1-6-2018 from, (2012) . <https://app.box.com/shared/h0674knjzl>.
- [15] P.L. Walker, Sexing skulls using discriminant function analysis of visually assessed traits, *American Journal of Physical Anthropology* 136 (1) (2008) 39–50, doi:http://dx.doi.org/10.1002/ajpa.20776.
- [16] D.H. Ubelaker, A.H. Ross, S.M. Graver, Application of forensic discriminant functions to a Spanish cranial sample, *Forensic Sci Commun* 4 (2002). Retrieved from <https://archives.fbi.gov/archives/about-us/lab/forensic-science-communications/fsc/july2002/ubelaker1.htm>.
- [17] M. Elliott, M. Collard, FORDISC and the determination of ancestry from cranial measurements, *Biology Letters* 5 (6) (2009) 849–852, doi:http://dx.doi.org/10.1098/rsbl.2009.0462.
- [18] P. Urbanová, A.H. Ross, M. Jurda, M. Nogueira, Testing the reliability of software tools in sex and ancestry estimation in a multi-ancestral Brazilian sample, *Legal Medicine* 16 (5) (2014) 264–273, doi:http://dx.doi.org/10.1016/j.legalmed.2014.06.002.
- [19] M. Steyn, M. Patriquin, Osteometric sex determination from the pelvis—Does population specificity matter? *Forensic Science International* 191 (1–3) (2009), doi:http://dx.doi.org/10.1016/j.forsciint.2009.07.009 113.e1–113.e5.
- [20] C. Eliopoulos, A. Lagia, S. Manolis, A modern, documented human skeletal collection from Greece, *HOMO - Journal of Comparative Human Biology* 58 (3) (2007) 221–228, doi:http://dx.doi.org/10.1016/j.jchb.2006.10.003.
- [21] P. Cignoni, M. Corsini, G. Ranzuglia, MeshLab: An open-source 3D mesh processing system, *ERCIM News* 73 (2008) 45–46.
- [22] J.W. Eaton, D. Bateman, S. Hauberg, R. Wehbring, *GNU Octave version 4.2.1 manual: a high-level interactive language for numerical computations*, (2017) . <https://www.gnu.org/software/octave/doc/v4.2.1/>.
- [23] [dataset] Bertatos, A., Chovalopoulou, M.-E. (2019). 3D cranial landmark coordinates from the Athens Collection. A modern Greek population reference dataset for testing 3D-ID software, Zenodo. doi:<https://doi.org/10.5281/zenodo.2553692>.
- [24] N. Lynnerup, Forensic anthropology and human identification, *Scandinavian Journal of Forensic Science* 19 (1) (2013), doi:http://dx.doi.org/10.2478/sjfs-2013-0005.
- [25] M.E. Chovalopoulou, E.D. Valakos, S.K. Manolis, Sex determination by three-dimensional geometric morphometrics of craniofacial form, *Anthropologischer Anzeiger - Journal of Biological and Clinical Anthropology* 73 (3) (2016) 195–206, doi:http://dx.doi.org/10.1127/anthranz/2016/0470.
- [26] R.E. King, Estimating ancestry in South Africa: a comparison of geometric morphometrics and traditional craniometrics (Master's Thesis, Boston University) Retrieved from, (2015) . <http://open.bu.edu>.
- [27] E. Nikita, E. Michopoulou, A quantitative approach for sex estimation based on cranial morphology, *American Journal of Physical Anthropology* 165 (3) (2018) 507–517, doi:http://dx.doi.org/10.1002/ajpa.23376.
- [28] E. Nikita, Quantitative Sex Estimation Based on Cranial Traits Using R Functions, *Journal of Forensic Sciences*. (2018), doi:http://dx.doi.org/10.1111/1556-4029.13833.
- [29] A. Bertatos, C. Papageorgopoulou, E. Valakos, M. Chovalopoulou, Investigating the sex-related geometric variation of the human cranium, *International Journal of Legal Medicine*. (2018), doi:http://dx.doi.org/10.1007/s00414-018-1790-z.

ORIGINAL ARTICLE

Open Access



Estimating sex using discriminant analysis of mandibular measurements from a modern Greek sample

Andreas Bertsatos, Konstantina Athanasopoulou and Maria-Eleni Chovalopoulou*

Abstract

Background: Sex determination is an integral and fundamental step in biological profile construction. The mandible, which holds many dimorphic traits, can be useful for sex estimation in the forensic context. However, reliable sex estimation usually relies on population-specific mandibular morphometric criteria. To date, no such criteria are available for the modern Greek population, and the present study aims to fill this lack of data by identifying the most sexually dimorphic mandibular traits on a modern Greek population sample and reporting the discriminant functions that can most effectively be used for estimating sex.

Materials and methods: For the purposes of this research, the 3D models of 194 adult mandibles (105 males and 89 females) from the Athens skeletal reference collection were used. A battery of 20 linear and 3 angular measurements was calculated from the 3D coordinates of anatomical landmarks positioned on the respective models and was analyzed by means of ANOVA and discriminant function analysis to investigate the expression of sexual dimorphism.

Results: The *coronoid height*, the *ramus height*, and the *maximum mandibular length* are the most sexually dimorphic metric traits of the mandible, while the produced sex discriminant functions yielded cross-validated classification accuracy up to 85.7% for the Greek sample. Furthermore, most of the examined combinations of measurements exhibited the same sex discriminant capacity between different reference samples, despite their respective discriminant functions being population specific.

Conclusions: Our findings indicate that the produced sex discriminant functions can be effectively used for sex determination in forensic casework and to verify the population specificity of these functions but also suggest that the expression of sexual dimorphism in the mandible shares certain features across different populations.

Keywords: Sex determination, Discriminant function analysis, Mandibular measurements, Greek population

Background

It is well established that the estimation of various biological profile parameters, such as age at death and stature, is affected by sex (Bidmos et al. 2010; Krishan et al. 2016). Therefore, sex determination is an integral and fundamental step in biological profile construction from skeletal remains (Lin et al. 2014;

Lopez-Capp et al. 2017). Although sex can be determined by forensic anthropologists with an accuracy of 90–100% when the entire skeleton is available (Byers 2016), in cases where the skeleton is incomplete or fragmented, sex determination can be more difficult (Passalacqua et al. 2013). Hence, the need of reliable sex discriminatory methods based on individual parts of the skeleton is evident.

Most human bones exhibit sexual dimorphism, while the pelvis is the most reliable sex estimator, followed by the skull and the humerus (France 1998). However, when the pelvis and the skull are fragmented or missing, the mandible can be useful for estimating sex in a

* Correspondence: marilenachov@biol.uoa.gr

Department of Animal and Human Physiology, Faculty of Biology, School of Sciences, National and Kapodistrian University of Athens, Panepistimiopolis, GR 15701 Athens, Greece

forensic context, since it is often well preserved and it holds many dimorphic traits (Franklin et al. 2007; Franklin et al. 2008; Saini et al. 2011). The mandible exhibits sexual dimorphism from an early stage of life (Kano et al. 2015). More specifically, earlier studies have shown that the expression of sexual dimorphism can be detected from as early as 5 months old (Schutkowski 1993; Loth and Henneberg 2001; Scheuer 2002). Nevertheless, mandibular sexual dimorphism is reduced during childhood (between the ages of 4 and 14), while it becomes again most prominent during adulthood (Coquerelle et al. 2011). Furthermore, it has been demonstrated that the sex-related morphological characteristics exhibit both inter- and intra-population variability, which is influenced by a range of factors including genetics, nutrition, secular change, growth, and morbidity as well as other general environmental factors (Ubelaker and DeGaglia 2017).

Various sex discriminating morphological and morphometric traits of the mandible have been described in the literature, such as the “mandibular ramus flexure” or the “ramus length,” respectively (Hu et al. 2006; Dong et al. 2015; Sharma et al. 2016). However, morphometric analyses, as opposed to qualitative sex discriminant methods, are more commonly used due to its inherent objectivity, accuracy, reproducibility, and lower level of inter- and intra-observer errors (Saini et al. 2011; Carvalho et al. 2013; Sharma et al. 2016).

Linear discriminant analysis is a simple-to-implement yet accurate classifier for normally distributed datasets (Dixon and Brereton 2009). Hence, it is commonly used in forensic applications for developing sex discriminant functions (Franklin et al. 2005; Dayal et al. 2008; Ogawa et al. 2013). Over the years, numerous sex discriminant functions based on mandibular measurements have been reported for various populations (Giles 1964; Steyn and İşcan 1998; Franklin et al. 2008; Saini et al. 2011; Lin et al. 2014; Dong et al. 2015). Nevertheless, the subtlety of metric assessments renders any derived sex discriminant function highly population specific as well as sensitive to secular trends (Dayal et al. 2008). Hence, their sex discriminating capacity usually diminishes when applied on samples different from the reference population. To date, no such study has been conducted on a modern Greek population sample and the aim of the present study was to fill this gap by identifying those mandibular morphometric traits that can be most effectively used in sex assessment. Additionally, this study investigated whether and to what extent discriminant functions derived from different population samples produce reliable sex estimation on the modern Greek population sample as a means to identify possible similarities in the expression of sexual dimorphism among diverse population groups. Sex discriminant functions representative for four contemporary populations from South Africa, India, and Korea have been examined to this end.

Materials and methods

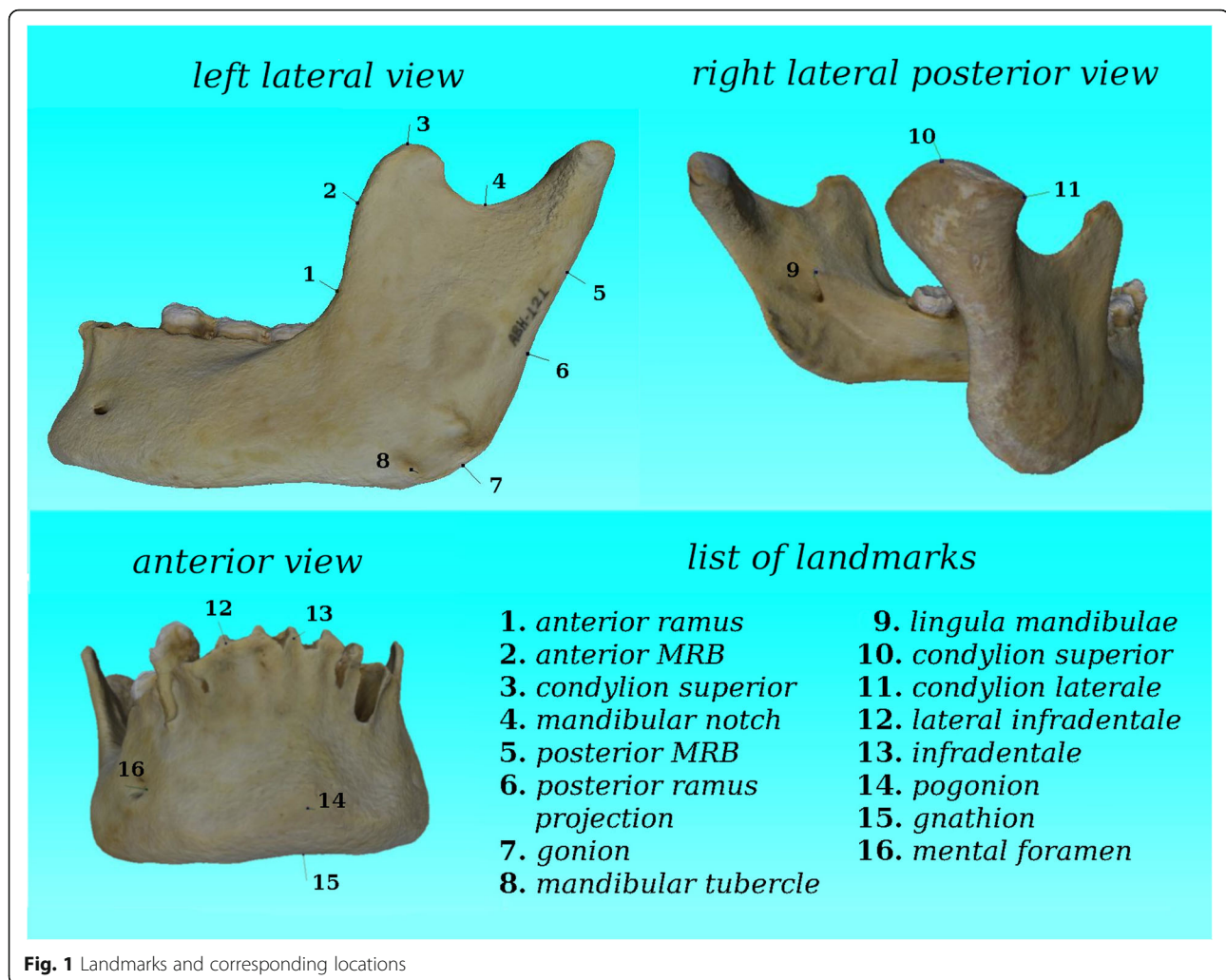
For the purposes of this research, 194 mandibles of adult Greek individuals were used. The sample consisted of 105 males and 89 females. Only individuals without any evidence of pathological deformation in the mandible and skull were selected. These individuals constitute part of the modern skeletal reference collection, also known as the Athens Collection, which is housed in the Department of Animal and Human Physiology at the National and Kapodistrian University of Athens. For each individual in the collection, sex, age at death, and cause of death as well as place and year of birth are known from death records (Eliopoulos et al. 2007). The age at death of these individuals ranges from 19 to 99 years old and their respective years of birth span from 1879 to 1965. Previous work has shown that the Athens collection is a fairly representative reference sample of the twentieth century Greek population (Bertsatos and Chovalopoulou 2017).

Data acquisition

The mandibles were previously digitized by means of 3D photogrammetry (Photoscan Pro version 1.4, Agisoft LLC, Russia) with a cross-validated model accuracy of 0.2 mm between repetitive digital and caliper-based measurements used for scaling. The ongoing digital documentation of the Athens Collection aims to the preservation of the physical reference collection; hence, the 3D models were utilized in the present study. The three-dimensional coordinates of 16 landmarks (3 midline and 13 bilateral) were extracted from the respective models with MeshLab software (Cignoni et al. 2008). All landmarks used in the present study are illustrated in Fig. 1, and their selection aimed to most appropriately capture the shape of the mandible but also facilitate comparisons with specific sets of measurements used in the existing literature. Most of the landmark definitions were obtained from Franklin et al. (2008), while *mental foramen* and *lingula mandibulae* were defined by Bejdová et al. (2013). The non-standard landmarks *mandibular tubercle*, *aMRB*, *pMRB*, and *prp* were incorporated to facilitate specific measurements. All measurements are listed in Table 1 along with their definitions. In total, 23 measurements were used in the subsequent analysis, from which three were angular and 11 bilateral linear measurements. Bilateral measurements were calculated for both sides of each mandible.

Data analysis

Intra- and inter-observer error was estimated using absolute and relative technical error measurement (TEM) as well as reliability coefficient R for each measurement based on a subsample of 20 randomly chosen individuals. The expression of sexual dimorphism was identified by means of ANOVA and univariate linear



discriminant analysis on each measurement to identify those that can be most effectively used in sex assessment regarding the Greek population. Subsequently, the measurements with correct cross-validated classification in excess of 75% were analyzed by means of multivariate discriminant function analysis, using both enter and stepwise methods, to identify the combination of metric traits that can be most effectively used in sex assessment.

In order to investigate whether and to what extent the mandible-based sex discriminant functions (DFs) are population specific, previously published DFs (from different population samples) were produced from the Greek population sample for comparing their cross-validated classification accuracy with that reported on the original population samples. More specifically, the combination of measurements of one DF from a South African Whites population sample (Steyn and İşcan 1998), three DFs from a South African Blacks population sample (Franklin et al. 2008), two DFs from an Indian

population sample (Saini et al. 2011), two DFs from a Korean population sample (Lin et al. 2014), and one DF from a contemporary Han Chinese population sample (Dong et al. 2015) were used for this comparison. Additionally, the classification accuracy of the previously reported DFs was also tested on the Greek sample as a measure of how their sex discriminant capacity changes when applied on a different sample than the reference population.

The evaluation of intra- and inter-observer error as well as the correct classification accuracy of applying previously published DFs on the Greek population sample were performed with the GNU Octave programming language (Eaton et al. 2016), while ANOVA and discriminant function analysis (DFA) were performed with SPSS (IBM SPSS version 23.0, Armonk, NY).

Results

Intra- and inter-observer error, by means of absolute technical error measurement (TEM), relative TEM and

Table 1 Definitions of the measurements used in this study

Measurement	Abbreviation	Definition ^a
Bi-condylar breadth	BicBR	The distance between right and left condyion laterale (cdl-cdl)
Bi-gonion breadth	BigBR	The distance between right and left gonion (go-go)
Bi-coronoid breadth	BicBR	The distance between right and left coronion (co-co)
Symphysis height	SyH	The distance between the gnathion and infradentale (gn-id)
Bi-mental foramina breadth	BimfBR	The distance between right and left mental foramen (mf-mf)
Symphysis breadth	SyBR	The distance between right and left lateral infradentale (lid-lid)
Bi-lingula breadth	BilBR	The distance between the lingula mandibulae (lm-lm)
Mandible breadth	MB	The distance between right and left anterior ramus (ar-ar)
Bi-notch breadth	BinBR	The distance between right and left mandibular notch (mn-mn)
Angle formed by bilateral gn-go lines	Agn-go	The angle formed by right gonion, gnathion and left gonion (go-gn-go)
Angle formed by bilateral gn-co lines	Agn-co	The angle formed by right coronion, gnathion and left coronion (co-gn-co)
Angle formed by bilateral gn-cdl lines	Agn-cdl	The angle formed by right condyion laterale, gnathion and left condyion laterale (cdl-gn-cdl)
Ramus height ^b	RH	The distance between the condyion superior and gonion (cs-go)
Maximum ramus height ^b	MRH	The distance between the most superior point on the mandibular condyle to the mandibular tubercle, or most protruding point of the inferior border of the ramus (cs-mtb)
Projective height of ramus ^b	PRH	Projective height of ramus between the highest point of the mandibular capitulum and lower margin of the bone (cs) ^c
Coronoid height ^b	CH	Projective distance between coronion and lower wall of the bone (co) ^c
Maximum ramus breadth ^b	MaxRBR	The distance between the most anterior point on the mandibular ramus and a line connecting the most posterior point on the condyle and the angle of jaw (aMRB-pMRB)
Minimum ramus breadth ^b	MinRBR	Smallest anterior-posterior diameter of the ramus (ar-prp)
Notch length ^b	NL	The distance between the coronion and condyion superior (co-cs)
Coronion-gonion length ^b	CGL	The distance between the coronion and gonion (co-go)
Maximum body length ^b	MaxBL	The distance between the gonion and pogonion (go-pg)
Body length ^b	BL	The distance between the gonion and gnathion (go-gn)
Maximum mandible length ^b	MaxML	The distance between the condyion superior and pogonion (cs-pg)

^aMost of landmarks definitions were obtained from Franklin et al. (2008), mental foramen and lingula mandibulae were described in Bejdová et al. (2013), while mandibular tubercle and non-standard landmarks aMRB, pMRB, and prp, are shown in Fig. 1

^bBilateral measurements were calculated for both sides of the mandible

^cThe projective distance is calculated by the formula $d = (v \cdot n)$, where n is the unit normal of the mandibular base plane calculated by three most inferior points of the mandibular body and v is a vector from any of these points to the landmark whose projective distance is being calculated

R coefficient, are presented in Table 2. Regarding the intra-observer error, absolute TEMs ranged within 0.16–0.43 mm, while the values of the relative TEM ranged within 0.27–1.36%. Our findings showed that the most precise measurement was BicBR and the least precise one was SyH. The corresponding ranges concerning the inter-observer error were 0.28–0.84 mm and 0.46–2.77%, respectively. According to the results, the most consistent measurement was MaxML (left side) while MinRBR (right side) provided the lowest consistency. R coefficients showed good reliability ($R > 0.95$) for almost all measurements with the exception of six measurements (Agn-cdl, BilBR, SyBR, left-side NL, left- and right-side MinRBR), whose R values ranged from 0.909 to 0.948.

Descriptive statistics for each measurement variable are presented in Table 3, whereas Table 4 lists the respective results of ANOVA and linear discriminant analysis. Males exhibit higher mean values in all linear measurements over females and lower mean values in angular measurements with the sole exception of the angle *Agn-go*. The univariate linear discriminant analysis revealed that only five metric traits exhibited a cross-validated correct classification percentage over 75%. These traits concerned the right-side measurements, RH, MRH, CH, CGL and MaxML, which were subsequently used in multivariate DFA.

Tables 5 and 6 summarize the statistics of the discriminant function analysis for all the DFs derived from the Greek population sample for the right- and the left-side

Table 2 Intra- and inter-observer error results

Measurement	Intra-observer TEM		R	Inter-observer TEM		R
	Absolute	Relative		Absolute	Relative	
BicBR	0.32	0.27	0.996	0.61	0.52	0.987
BigBR	0.33	0.35	0.997	0.69	0.72	0.986
BicrBR	0.40	0.43	0.995	0.63	0.68	0.989
SyH	0.39	1.36	0.982	0.58	2.03	0.963
BimfBR	0.30	0.67	0.991	0.56	1.26	0.971
SyBR	0.16	0.88	0.970	0.28	1.61	0.911
RH-right	0.38	0.60	0.994	0.50	0.78	0.989
MRH-right	0.34	0.50	0.995	0.70	1.02	0.978
PRH-right	0.36	0.62	0.997	0.65	1.10	0.991
CH-right	0.37	0.56	0.997	0.49	0.75	0.995
MaxRBR-right	0.35	0.82	0.989	0.66	1.55	0.967
MinRBR-right	0.28	0.91	0.991	0.84	2.77	0.929
NL-right	0.26	0.74	0.994	0.54	1.50	0.974
CGL-right	0.39	0.63	0.997	0.61	0.99	0.992
MaxBL-right	0.24	0.30	0.997	0.57	0.71	0.983
BL-right	0.31	0.40	0.995	0.62	0.79	0.981
MaxML-right	0.35	0.30	0.989	0.58	0.49	0.970
RH-left	0.38	0.61	0.993	0.70	1.13	0.976
MRH-left	0.25	0.37	0.998	0.37	0.55	0.996
PRH-left	0.28	0.50	0.999	0.49	0.85	0.996
CH-left	0.43	0.66	0.996	0.65	0.99	0.990
MaxRBR-left	0.31	0.72	0.991	0.68	1.61	0.955
MinRBR-left	0.39	1.25	0.963	0.61	1.97	0.909
NL-left	0.35	1.03	0.988	0.75	2.18	0.943
CGL-left	0.33	0.54	0.998	0.77	1.27	0.987
MaxBL-left	0.32	0.39	0.996	0.71	0.87	0.979
BL-left	0.27	0.34	0.997	0.61	0.76	0.985
MaxML-left	0.34	0.29	0.988	0.54	0.46	0.972
BiI BR	0.34	0.43	0.988	0.70	0.89	0.946
MB	0.31	0.37	0.991	0.59	0.70	0.969
BinBR	0.35	0.37	0.993	0.62	0.66	0.979
Agn-go	0.31	0.41	0.994	0.57	0.76	0.979
Agn-co	0.36	0.63	0.980	0.57	1.00	0.954
Agn-cdl	0.39	0.65	0.982	0.66	1.10	0.948

measurements, respectively. The first nine DFs in each table concern the corresponding functions based on previously published works, whereas the latter concern the statistically permissible DFs produced by enter and step-wise methods with the most sexually dimorphic measurements found on the Greek population sample. Although no statistically significant differences were observed between the measurements of the right and left side, the corresponding equations based on the left-side measurements were produced in order to detect any

differences in classification accuracy, but also for comparative consistency with previously published research.

The cross-validated classification accuracy of the examined DFs ranged from 77.1 to 85.7% with Function 7 (based on MRH and MinRBR measurements) exhibiting the lowest accuracy for both the left- and the right-side measurements and Function 4 (based on BigBR, CGL, MaxBL, and MaxML measurements) exhibiting the highest. Regarding the DFs produced by combining the most sexually dimorphic measurements from the Greek population sample, their accuracy was consistently over 82%. The DFs based on the right-side measurements exhibited slightly better overall performance, although Function 15 yielded the best classification score (84.3%) on the left-side measurements. Regarding the nine DFs produced from our sample, which correspond to previously published functions, in most of the cases, the equations based on the right-side measurements outperform their left-side counterparts. The only exceptions concern the Functions 5 and 7, which showed negligible improvement on the left side.

Table 7 summarizes the classification accuracy of the application of the previously reported DFs on the Greek sample. As expected, their discriminant capacity is reduced as compared to the classification accuracy of the respective DFs derived from the Greek sample. More specifically, the accuracy ranged from 63.78 to 83.84% when applied on the left-side measurements and from 68.11% to 83.84% when the right-side measurements were used. The highest sex discriminant capacity was observed with Function 3, which incorporated nine variables and reached 83.84% correct classification for both sides. The lowest correct classification accuracy was exhibited by Function 1, which correctly classified 63.79% and 68.11% of the Greek sample, when using the left- and the right-side measurements, respectively. Finally, it should be noted that the DF reported by Dong et al. (2015) could not be applied on our sample, since they omitted to report the DF's constant value in their paper.

Discussion and conclusions

In anthropometric studies, the determination of measurement error is essential (Ulijaszek and Lourie 1994). According to this study, all measurements exhibited an acceptable measurement error indicating high repeatability and precision, whereas their respective R coefficients demonstrated that all measurements may be regarded as reliable. These findings are in agreement with those reported by Toneva et al. (2016). Furthermore, most of the measurements are below the intra- and inter-observer %TEM thresholds of 1% and 1.5%, respectively, which are usually considered acceptable for skillful anthropometrists (Perini et al. 2005). Despite that the photogrammetric 3D modeling of the mandibles was

Table 3 Descriptive statistics for each measurement variable

Measurement	Males			Females		
	Mean	Std. deviation	Shapiro-Wilk	Mean	Std. deviation	Shapiro-Wilk
BicBR	118.106	6.674	0.602	113.002	4.691	0.230
BigBR	98.653	5.502	0.757	90.961	4.593	0.537
BicrBR	94.753	4.243	0.582	90.456	5.448	0.083
SyH	31.107	3.507	0.914	27.935	2.246	0.201
BimfBR	45.318	2.230	0.871	43.547	2.259	0.750
SyBR	20.084	1.759	0.395	19.251	1.239	0.107
RH-right	65.968	4.608	0.200 ^a	59.624	3.842	0.538
MRH-right	71.562	4.642	0.697	64.387	3.842	0.739
PRH-right	58.109	6.944	0.907	52.429	5.038	0.606
CH-right	68.257	4.982	0.975	59.891	4.485	0.814
MaxRBR-right	40.758	3.065	0.237	38.354	3.844	0.090
MinRBR-right	31.446	2.877	0.232	30.138	2.697	0.531
NL-right	34.537	3.094	0.001	32.730	3.472	0.214
CGL-right	63.835	5.001	0.799	56.159	4.325	0.770
MaxBL-right	84.172	4.740	0.268	79.383	4.087	0.766
BL-right	80.940	4.511	0.277	75.870	3.877	0.882
MaxML-right	124.716	5.181	0.286	116.618	4.542	0.881
RH-left	66.233	4.196	0.141	60.281	3.892	0.566
MRH-left	71.481	4.625	0.771	64.385	4.224	0.829
PRH-left	57.917	6.554	0.845	52.749	5.141	0.365
CH-left	68.125	4.458	0.412	60.255	4.563	0.336
MaxRBR-left	40.812	3.032	0.341	38.455	3.724	0.050
MinRBR-left	31.461	2.633	0.399	30.473	2.396	0.771
NL-left	34.563	3.294	0.380	33.143	3.626	0.486
CGL-left	63.867	4.286	0.657	56.649	4.457	0.077
MaxBL-left	83.343	4.507	0.211	78.557	4.449	0.203
BL-left	80.188	4.389	0.200 ^a	75.087	4.310	0.412
MaxML-left	124.710	5.780	0.488	116.576	4.311	0.959
BilBR	80.371	3.268	0.673	75.993	3.112	0.081
MB	83.984	2.834	0.346	80.665	3.499	0.553
BinBR	95.934	4.236	0.889	92.168	4.260	0.650
Agn-go	75.661	4.991	0.664	74.229	4.216	0.138
Agn-co	53.891	2.999	0.910	56.805	3.020	0.384
Agn-cdl	57.626	3.778	0.764	59.438	3.104	0.679

^anormality checked with KS test

not part of the present study, the reported cross-validated accuracy between digital and manual measurements, which was part of the digital documentation process, is smaller than the inter- and intra-observer absolute TEMs reported in the present study. Hence, the present results and DFs can be applied to either digital or manual measurements without any inter-method measurement error weakening their utility.

The aim of the present study has been twofold. We aimed to identify the mandibular measurements based on a modern Greek population sample that yield the highest sex discriminant capacity and use them to produce multivariate DFs that most successfully can estimate sex. Meanwhile, we investigated the performance of previously published DFs derived from different population samples. Population specificity is a long and well-established observation in osteometric studies also evident from our results (Giles 1964;

Table 4 ANOVA and linear discriminant analysis for each measurement variable

Measurement	Sample size Total (♂-♀)	ANOVA		Univariate DFAs		
		F	Sig.	Wilks' Lambda	Sig.	Cross-validated (%)
BicBR	187 (101-86)	34.128 ^a	< 0.001	0.850	< 0.001	65.8
BigBR	194 (105-89)	55.614	< 0.001	0.775	< 0.001	73.2
BicrBR	192 (103-89)	36.408	< 0.001	0.839	< 0.001	65.6
SyH	105 (58-47)	25.357 ^a	< 0.001	0.817	< 0.001	72.4
BimfBR	184 (101-83)	18.118	< 0.001	0.909	< 0.001	64.7
SyBR	112 (63-49)	4.026 ^a	0.047	0.968	0.059	60.7
RH-right	193 (105-88)	143.997 ^a	< 0.001	0.578	< 0.001	80.3
MRH-right	194 (105-89)	130.603	< 0.001	0.595	< 0.001	78.4
PRH-right	194 (105-89)	54.279	< 0.001	0.780	< 0.001	69.1
CH-right	192 (104-88)	92.520	< 0.001	0.673	< 0.001	75.5
MaxRBR-right	192 (104-88)	24.214	< 0.001	0.887	< 0.001	63.0
MinRBR-right	194 (105-89)	6.251	0.013	0.968	0.013	58.8
NL-right	191 (103-88)	-	-	0.928	< 0.001	59.7
CGL-right	192 (103-89)	101.686	< 0.001	0.651	< 0.001	77.6
MaxBL-right	192 (105-87)	35.913	< 0.001	0.841	< 0.001	66.7
BL-right	194 (105-89)	37.105	< 0.001	0.838	< 0.001	65.5
MaxML-right	191 (105-86)	117.120	< 0.001	0.617	< 0.001	82.7
RH-left	193 (104-89)	128.705 ^a	< 0.001	0.605	< 0.001	75.6
MRH-left	193 (104-89)	113.412	< 0.001	0.627	< 0.001	78.8
PRH-left	193 (104-89)	46.203	< 0.001	0.805	< 0.001	68.4
CH-left	192 (104-88)	91.658	< 0.001	0.675	< 0.001	76.0
MaxRBR-left	194 (105-89)	25.662	< 0.001	0.882	< 0.001	61.9
MinRBR-left	194 (105-89)	7.653	0.006	0.962	0.006	58.2
NL-left	193 (104-89)	10.675	0.001	0.947	0.001	57.5
CGL-left	194 (105-89)	99.576	< 0.001	0.658	< 0.001	77.8
MaxBL-left	192 (105-87)	33.815	< 0.001	0.849	< 0.001	64.6
BL-left	194 (105-89)	32.626	< 0.001	0.855	< 0.001	64.4
MaxML-left	191 (104-87)	120.559 ^a	< 0.001	0.624	< 0.001	78.5
BilBR	193 (104-89)	76.144	< 0.001	0.715	< 0.001	72.5
MB	189 (102-87)	33.036	< 0.001	0.850	< 0.001	64.0
BinBR	189 (102-87)	35.789	< 0.001	0.839	< 0.001	66.1
Agn-go	194 (105-89)	2.948	0.088	0.985	0.088	56.7
Agn-co	192 (103-89)	23.503	< 0.001	0.890	< 0.001	64.6
Agn-cdl	187 (101-86)	11.642	0.001	0.941	0.001	57.8

^aAsymptotically F-distributed values based on Welch ANOVA

İşcan and Steyn 1999; Franklin et al. 2007). Nevertheless, the rationale of our approach was not limited to a mere validation of population specificity but further aimed to identify particular morphometric traits that consistently yield high sex classification results among different population groups even though their expression patterns may vary among these groups.

Our findings regarding the most sex discriminant mandibular traits were consistent with earlier studies by

Franklin et al. (2006, 2008), who found that the *coronoid height*, the *ramus height*, and the *maximum mandibular length* univariately exhibited the most pronounced sexual dimorphism. Furthermore, the produced DF (Function 10, Table 5), when combining these measurements from the Greek population sample, yielded a cross-validated accuracy of 84.1%. Although the cranium and the pelvis provide much more reliable sex estimates (Oikonomopoulou et al. 2017; Bertsatos et al. 2018), our

Table 5 Discriminant functions based on the right-side measurements

Equation	Wilks' Lambda	Sig.	Group centroids and sectioning point	Correctly classified	
				Original	Cross-validated
Function 1 ^a (BicBR × 0.022) + (BigBR × 0.063) + (MinRBR × -0.026) + (BL × -0.043) + (MaxML × 0.183) - 26.447	0.571	< 0.001	♀-0.945, [-0.0795], ♂0.786	83.20%	81.60%
Function 2 ^b (BigBR × 0.052) + (CGL × 0.115) + (MaxML × 0.114) - 25.567	0.492	< 0.001	♀- 1.105, [- 0.091], ♂0.923	85.20%	85.20%
Function 3 ^b (BicBR × -0.037) + (BigBR × 0.103) + (BicrBR × 0.007) + (SyH × 0.032) + (SyBR × -0.041) + (RH × -0.017) + (CGL × 0.151) + (MaxBL × -0.047) + (MaxML × 0.126) - 25.838	0.381	< 0.001	♀- 1.411, [- 0.141], ♂1.129	89.90%	84.80%
Function 4 ^b (BigBR × 0.064) + (CGL × 0.119) + (MaxBL × -0.063) + (MaxML × 0.139) - 24.786	0.481	< 0.001	♀- 1.131, [- 0.0935], ♂0.944	85.70%	85.70%
Function 5 ^c (MRH × 0.214) + (PRH × -0.064) + (CH × 0.09) + (MaxRBR × 0.100) + (MinRBR × -0.124) - 16.856	0.532	< 0.001	♀- 1.015, [- 0.078], ♂0.859	82.30%	81.80%
Function 6 ^c (CH × 0.180) + (MaxRBR × 0.137) + (MinRBR × -0.133) - 12.807	0.633	< 0.001	♀- 0.824, [- 0.0635], ♂0.697	80.20%	79.70%
Function 7 ^d (MRH × 0.220) + (MaxRBR × 0.029) - 16.028	0.594	< 0.001	♀- 0.894, [- 0.0685], ♂0.757	77.10%	77.10%
Function 8 ^d (MRH × 0.247) + (MinRBR × -0.072) - 14.495	0.585	< 0.001	♀- 0.91, [- 0.069], ♂0.772	80.40%	80.40%
Function 9 ^e (BicBR × 0.008) + (BigBR × 0.049) + (MRH × 0.127) + (MaxML × 0.087) - 24.711	0.519	< 0.001	♀- 1.05, [- 0.0885], ♂0.873	85.90%	84.90%
Function 10 ^f (RH × 0.097) + (CH × 0.072) + (MaxML × 0.102) - 22.935	0.498	< 0.001	♀- 1.105, [- 0.101], ♂0.903	84.10%	84.10%
Function 11 ^f (MRH × 0.063) + (CH × 0.079) + (MaxML × 0.118) - 23.516	0.515	< 0.001	♀- 1.068, [- 0.0975], ♂0.873	84.10%	83.60%
Function 12 ^f (RH × 0.081) + (CGL × 0.086) + (MaxML × 0.108) - 23.085	0.494	< 0.001	♀- 1.102, [- 0.091], ♂0.92	84.10%	84.10%
Function 13 ^f (MRH × 0.064) + (CGL × 0.096) + (MaxML × 0.111) - 23.417	0.498	< 0.001	♀- 1.093, [- 0.0905], ♂0.912	84.10%	83.10%

^aFunction based on Steyn and İşcan (1998)

^bFunction based on Franklin et al. (2008)

^cFunction based on Saini et al. (2011)

^dFunction based on Lin et al. (2014)

^eFunction based on Dong et al. (2015)

^fA combination of the five variables with the highest univariate discriminating capacity. Their corresponding DFA analysis using the stepwise method yielded identical function and accuracy

results verify that the mandible can be useful for estimating sex in a forensic context, when cranial and pelvic elements are missing or deteriorated.

Franklin and colleagues in 2008 studied 225 individuals (120 male; 105 female) from five local populations of indigenous South Africans, in order to produce a series of mandibular metric standards for sex estimation, which resulted in Functions 2, 3, and 4 (see Table 7). The corresponding DFs based on the Greek sample (Tables 5 and 6) yielded the highest sex discriminant scores observed in the present study ranging from 84.3 to 85.7%. More specifically, Function 3, which includes nine mandibular measurements, yielded similar classification

accuracy on our sample (84.8%) with that reported by Franklin and colleagues on their population sample (84%). Furthermore, applying the original DF (Franklin et al. 2008) on the Greek population sample also resulted to similar accuracy (83.84%). However, the accuracy of Functions 2 and 4, which utilize three and four mandibular measurements, respectively, was higher for the DFs derived from the Greek population (Function 2: right side 85.2%, left side 84.3%; Function 4: right side 85.7%, left side 84.8%) as opposed to their counterparts derived from the indigenous South African population (Function 2: 81.8%; Function 4: 82.7%). Additionally, applying the original DFs on the Greek population sample

Table 6 Discriminant functions based on the left-side measurements

Equation	Wilks' Lambda	Sig.	Group centroids and sectioning point	Correctly classified	
				Original	Cross-validated
Function 1 ^a (BicBR × 0.028) + (BigBR × 0.065) + (MinRBR × -0.014) + (BL × -0.042) + (MaxML × 0.174) - 26.748	0.587	< 0.001	♀-0.915, [-0.077], ♂0.761	83.20%	81.60%
Function 2 ^b (BigBR × 0.056) + (CGL × 0.117) + (MaxML × 0.109) - 25.455	0.506	< 0.001	♀- 1.074, [- 0.088], ♂0.898	84.30%	84.30%
Function 3 ^b (BicBR × -0.017) + (BigBR × 0.101) + (BicrBR × 0.002) + (SyH × 0.024) + (SyBR × -0.020) + (RH × -0.010) + (CGL × 0.154) + (MaxBL × -0.034) + (MaxML × 0.099) - 26.007	0.399	< 0.001	♀- 1.36, [- 0.136], ♂1.088	89.90%	84.80%
Function 4 ^b (BigBR × 0.069) + (CGL × 0.122) + (MaxBL × -0.056) + (MaxML × 0.128) - 24.741	0.497	< 0.001	♀-1.095, [- 0.0895], ♂0.916	85.30%	84.80%
Function 5 ^c (MRH × 0.179) + (PRH × -0.047) + (CH × 0.098) + (MaxRBR × 0.123) + (MinRBR × -0.166) - 15.675	0.546	< 0.001	♀- 0.981, [- 0.0715], ♂0.838	83.80%	82.20%
Function 6 ^c (CH × 0.182) + (MaxRBR × 0.169) + (MinRBR × -0.159) - 13.474	0.619	< 0.001	♀-0.848, [- 0.065], ♂0.718	79.20%	78.60%
Function 7 ^d (MRH × 0.211) + (MaxRBR × 0.027) - 15.299	0.626	< 0.001	♀- 0.832, [-0.06], ♂0.712	77.70%	77.20%
Function 8 ^d (MRH × 0.245) + (MinRBR × -0.107) - 13.339	0.61	< 0.001	♀- 0.86, [-0.062], ♂0.736	80.30%	80.30%
Function 9 ^e (BicBR × 0.014) + (BigBR × 0.055) + (MRH × 0.116) + (MaxML × 0.082) - 24.524	0.545	< 0.001	♀- 0.997, [- 0.084], ♂0.829	84.30%	82.70%
Function 10 ^f (RH × 0.086) + (CH × 0.079) + (MaxML × 0.100) - 22.525	0.525	< 0.001	♀- 1.029, [- 0.0795], ♂0.87	83.20%	82.10%
Function 11 ^g (MRH × 0.045) + (CH × 0.088) + (MaxML × 0.118) - 22.900	0.538	< 0.001	♀- 1.003, [- 0.078], ♂0.847	83.70%	83.20%
Function 12 ^h (CH × 0.110) + (MaxML × 0.138) - 23.604	0.542	< 0.001	♀- 0.995, [- 0.0775], ♂0.84	84.20%	83.20%
Function 13 ^f (RH × 0.066) + (CGL × 0.095) + (MaxML × 0.109) - 22.949	0.518	< 0.001	♀- 1.05, [- 0.086], ♂0.878	83.20%	82.70%
Function 14 ^g (MRH × 0.041) + (CGL × 0.104) + (MaxML × 0.118) - 23.137	0.523	< 0.001	♀- 1.039, [- 0.085], ♂0.869	84.30%	83.80%
Function 15 ⁱ (CGL × 0.123) + (MaxML × 0.137) - 23.774	0.527	< 0.001	♀- 1.031, [- 0.0845], ♂0.862	84.80%	84.30%

^aFunction based on Steyn and İşcan (1998)

^bFunction based on Franklin et al. (2008)

^cFunction based on Saini et al. (2011)

^dFunction based on Lin et al. (2014)

^eFunction based on Dong et al. (2015)

^fA combination of the five variables with the highest univariate discriminating capacity. Their corresponding DFA analysis using the stepwise method yielded identical function and accuracy

^gA combination of the five variables with the highest univariate discriminating capacity

^hThe outcome of the stepwise DFA analysis using the variables of function 11

ⁱThe outcome of the stepwise DFA analysis using the variables of function 14

yielded even lower classification scores (see Table 7). Despite that Functions 2 and 4 resulted in differential performance between the two distinct population groups, which can be attributed to population specificity, the overall performance of these three functions implies some merit to the proposed metric standards by Franklin et al. (2008).

Steyn and İşcan in 1998 evaluated sexual dimorphism in the cranium and the mandible of South African Whites and developed osteometric standards to determine sex. They studied 91 South African Whites (44 males, 47 females) from cadaver collections housed at the Universities of Pretoria and Witwatersrand (Dart Collection). Although comparing their results to our

Table 7 The accuracy of other researchers' equations as well as their accuracy when applied on the Greek population sample

Date	Authors	Population	DF	DF's performance on reference population	DF's performance applied on the Greek sample	
1998	Steyn and İscan	South African Whites	Function 1	81.5%	Left side:	63.78%
					Right side:	68.11%
2008	Franklin et al.	South African Blacks	Function 2	81.8%	Left side:	72.77%
					Right side:	74.60%
2008	Franklin et al.	South African Blacks	Function 3	84.0%	Left side:	83.84%
					Right side:	83.84%
2008	Franklin et al.	South African Blacks	Function 4	82.7%	Left side:	78.53%
					Right side:	77.78%
2011	Saini et al.	Indian	Function 5	80.2%	Left side:	65.45%
					Right side:	68.75%
2011	Saini et al.	Indian	Function 6	80.2%	Left side:	75.00%
					Right side:	71.88%
2014	Lin et al.	Korean	Function 7	80.8%	Left side:	67.36%
					Right side:	68.75%
2014	Lin et al.	Korean	Function 8	80.0%	Left side:	68.84%
					Right side:	69.59%

corresponding DF results (Function 1) revealed similar classification accuracy (~ 81.5%), applying the original DF (Steyn and İscan 1998) on the modern Greek population sample exhibited much lower accuracy (left side: 63.78%; right side: 68.11%). The same pattern was observed on the rest of the comparisons between different population samples. The corresponding DFs produced similar classification accuracy between their respective reference population samples, but applying the original DFs (from other population samples) on the Greek sample resulted in significantly reduced accuracy.

More specifically, Saini et al. (2011), working on 116 dry adult mandibles of a Northern Indian population sample from the Department of Forensic Medicine in India, reported classification accuracy of 80.2% for both Functions 5 and 6, which on our sample yielded approximately 82% and 79%, respectively, with small deviations on each side. Lin et al. (2014) produced their DFs from cranial CT scans of 120 males and 120 females from Seoul St. Mary's Hospital. Despite most of their reported DFs were based on different set of measurements than these utilized in the present study, which restricted a direct comparison with their most accurate DFs in terms of correct sex classification, the available comparisons showed similar classification score for Function 7 (77.2% on Greek sample; 80.8% on Korean sample) and identical accuracy for Function 8 (80.4%). Similarly, the DF reported by Dong and colleagues based on a contemporary Han Chinese population sample yielded 83.3% classification accuracy (Dong et al. 2015), whereas the corresponding DF from the Greek sample produced

82.7% and 84.9% cross-validated classification scores for the left- and right-side measurements, respectively.

The present study produced a number of suitable DFs based on mandibular measurements that can be used for sexing unidentified individuals assumed to belong to the modern Greek population. Apart from verifying the fact that most often the morphometric DFs based on human bones exhibit population specificity, hence the need for studies on different population samples, the comparative part of this work also revealed some interesting aspects. The observation that the same combinations of measurements yield the same sex discriminant capacity between different population samples, despite their respective DFs being population specific, implies that the magnitude of expression of sexual dimorphism is similar on certain mandibular morphometric traits and shared across different populations, although their expression may follow different patterns in each population group. Regarding the identical results of Function 3 (Franklin et al. 2008) and especially the identical classification accuracy when applying the DF derived from indigenous South Africans on the Greek population, we cannot conclude whether this observation is a mere statistical coincidence between the two reference samples or the result of both population groups sharing similar environmental and developmental factors that led to similar expression of sexual dimorphism in their mandibles. Further work is necessary to this end, and studies on shared datasets comprising diverse population samples may expand our insight on this matter and provide more reliable sex discriminant tools for the forensic practice.

Abbreviations

DF: Discriminant function; DFA: Discriminant function analysis; TEM: Technical Error Measurement

Acknowledgements

Not applicable.

Funding

This research did not receive any grant from funding agencies in the public, commercial, or not-for-profit sectors.

Availability of data and materials

Please contact author for data requests.

Declarations

The authors declare that this manuscript is original, has not been published before, and is not currently being considered for publication elsewhere.

Authors' contributions

AB contributed to the data collection, statistical analysis, and manuscript preparation. KA contributed to the data collection and manuscript preparation. M-EC contributed to the experimental design, statistical analysis, and manuscript preparation. All authors have read and approved the final article.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 20 November 2018 Accepted: 15 May 2019

Published online: 24 May 2019

References

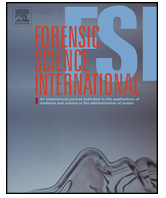
- Bejdová Š, Krajiček V, Velemínská J, Horák M, Velemínský P (2013) Changes in the sexual dimorphism of the human mandible during the last 1200 years in Central Europe. *HOMO*. 64(6):437–453. <https://doi.org/10.1016/j.jchb.2013.05.003>
- Bertsatos A, Chovalopoulou M-E (2017) Secular change in adult stature of modern Greeks. *Am J Hum Biol*. <https://doi.org/10.1002/ajhb.23077>
- Bertsatos A, Papageorgopoulou C, Valakos E, Chovalopoulou M-E (2018) Investigating the sex-related geometric variation of the human cranium. *Int J Legal Med* 132(5):1505–1514. <https://doi.org/10.1007/s00414-018-1790-z>
- Bidmos MA, Gibbon VE, Štrkalj G (2010) Recent advances in sex identification of human skeletal remains in South Africa. *S Afr J Sci*. <https://doi.org/10.4102/sajs.v106i11/12.238>
- Byers SN (2016) Introduction to forensic anthropology. Routledge, London
- Carvalho SP, Brito LM, Paiva LA, Bicudo LA, Crosato EM, Oliveira RN (2013) Validation of a physical anthropology methodology using mandibles for gender estimation in a Brazilian population. *J Appl Oral Sci*. <https://doi.org/10.1590/1678-775720130022>
- Cignoni P, Corsini M, Ranzuglia G (2008) MeshLab: an open-source 3D mesh processing system. *ERCIM News* 73:45–46
- Coquerelle M, Bookstein FL, Braga J, Halazonetis DJ, Weber GW, Mitteroecker P (2011) Sexual dimorphism of the human mandible and its association with dental development. *Am J Phys Anthropol*. <https://doi.org/10.1002/ajpa.21485>
- Dayal M, Spocter M, Bidmos M (2008) An assessment of sex using the skull of black South Africans by discriminant function analysis. *HOMO*. 59(3):209–221. <https://doi.org/10.1016/j.jchb.2007.01.001>
- Dixon SJ, Brereton RG (2009) Comparison of performance of five common classifiers represented as boundary methods: Euclidean distance to centroids, linear discriminant analysis, quadratic discriminant analysis, learning vector quantization and support vector machines, as dependent on data structure. *Chemom Intell Lab Syst*. <https://doi.org/10.1016/j.chemolab.2008.07.010>
- Dong H, Deng M, Wang W, Zhang J, Mu J, Zhu G (2015) Sexual dimorphism of the mandible in a contemporary Chinese Han population. *Forensic Sci Int*. <https://doi.org/10.1016/j.forsciint.2015.06.010>
- Eaton JW, Bateman D, Hausberg S, Wehbring R (2016) GNU Octave version 4.2.0 manual: a high-level interactive language for numerical computations. <https://octave.org/doc/interpreter/>
- Eliopoulos C, Lagia A, Manolis SK (2007) A modern, documented human skeletal collection from Greece. *HOMO*. <https://doi.org/10.1016/j.jchb.2006.10.003>
- France DL (1998) Observational and metric analysis of sex in the skeleton. In: Reichs KJ (ed) *Forensic osteology: advances in the identification of human remains*. Charles C Thomas Publisher, Springfield, pp 163–186
- Franklin D, Freedman L, Milne N (2005) Sexual dimorphism and discriminant function sexing in indigenous South African crania. *HOMO*. <https://doi.org/10.1016/j.jchb.2004.08.001>
- Franklin D, O'Higgins P, Oxnard CE, Dadour I (2008) Discriminant function sexing of the mandible of Indigenous South Africans. *Forensic Sci Int*. <https://doi.org/10.1016/j.forsciint.2008.03.014>
- Franklin D, O'Higgins P, Oxnard CE, Dadour I (2006) Determination of sex in South African Blacks by discriminant function analysis of mandibular linear dimensions: a preliminary investigation using the Zulu local population. *Forensic Sci Med Pathol*. <https://doi.org/10.1385/fmp.2:4:263>
- Franklin D, O'Higgins P, Oxnard CE, Dadour I (2007) Sexual dimorphism and population variation in the adult mandible. Forensic applications of geometric morphometrics. *Forensic Sci Med Pathol*. <https://doi.org/10.1385/fmp.2:4:263>
- Giles E (1964) Sex determination by discriminant function analysis of the mandible. *Am J Phys Anthropol*. <https://doi.org/10.1002/ajpa.1330220212>
- Hu KS, Koh KS, Han SH, Shin KJ, Kim HJ (2006) Sex determination using nonmetric characteristics of the mandible in Koreans. *J Forensic Sci*. <https://doi.org/10.1111/j.1556-4029.2006.00270.x>
- İşcan MY, Steyn M (1999) Craniometric determination of population affinity in South Africans. *Int J Legal Med*. <https://doi.org/10.1007/s004140050208>
- Kano T, Oritani S, Michiue T, Ishikawa T, Hishmat AM, Sogawa N et al (2015) Postmortem CT morphometry with a proposal of novel parameters for sex discrimination of the mandible using Japanese adult data. *Leg Med (Tokyo)*. <https://doi.org/10.1016/j.legalmed.2014.12.009>
- Krishan K, Chatterjee PM, Kanchan T, Kaur S, Baryah N, Singh R (2016) A review of sex estimation techniques during examination of skeletal remains in forensic anthropology casework. *Forensic Sci Int*. <https://doi.org/10.1016/j.forsciint.2016.02.007>
- Lin C, Jiao B, Liu S, Guan F, Chung N, Han S et al (2014) Sex determination from the mandibular ramus flexure of Koreans by discrimination function analysis using three-dimensional mandible models. *Forensic Sci Int*. <https://doi.org/10.1016/j.forsciint.2013.12.015>
- Lopez-Capp TT, Rynn C, Wilkinson C, De Paiva LA, Michel-Crosato E, Biazzevic MG (2017) Discriminant analysis of mandibular measurements for the estimation of sex in a modern Brazilian sample. *Int J Legal Med*. <https://doi.org/10.1007/s00414-017-1681-8>
- Loth SR, Henneberg M (2001) Sexually dimorphic mandibular morphology in the first few years of life. *Am J Phys Anthropol*. <https://doi.org/10.1002/ajpa.1067>
- Ogawa Y, Imaizumi K, Miyasaka S, Yoshino M (2013) Discriminant functions for sex estimation of modern Japanese skulls. *J Forensic Legal Med*. <https://doi.org/10.1016/j.jflm.2012.09.023>
- Oikononopoulou EK, Valakos E, Nikita E (2017) Population-specificity of sexual dimorphism in cranial and pelvic traits: evaluation of existing and proposal of new functions for sex assessment in a Greek assemblage. *Int J Legal Med*. <https://doi.org/10.1007/s00414-017-1655-x>
- Passalacqua NV, Zhang Z, Pierce SJ (2013) Sex determination of human skeletal populations using latent profile analysis. *Am J Phys Anthropol*. <https://doi.org/10.1002/ajpa.22295>
- Perini TA, de Oliveira GL, Ornellas JS, de Oliveira FP (2005) Technical error measurement in anthropometry. *Rev Bras Med Esporte*. <https://doi.org/10.1590/S1517-86922005000100009>
- Saini V, Srivastava R, Shamal SN, Singh T, Pandey AK, Tripathi SK (2011) Sex determination using mandibular ramus flexure: a preliminary study on Indian population. *J Forensic Legal Med*. <https://doi.org/10.1016/j.jflm.2011.02.014>
- Scheuer L (2002) A blind test of mandibular morphology for sexing mandibles in the first few years of life. *Am J Phys Anthropol*. <https://doi.org/10.1002/ajpa.10098>

- Schutkowski H (1993) Sex determination of infant and juvenile skeletons: I. Morphognostic features. *Am J Phys Anthropol*. <https://doi.org/10.1002/ajpa.1330900206>
- Sharma M, Gorea R, Gorea A, Abuderman A (2016) A morphometric study of the human mandible in the Indian population for sex determination. *Egypt J Forensic Sci*. <https://doi.org/10.1016/j.ejfs.2015.01.002>
- Steyn M, İşcan M (1998) Sexual dimorphism in the crania and mandibles of South African whites. *Forensic Sci Int*. [https://doi.org/10.1016/s0379-0738\(98\)00120-0](https://doi.org/10.1016/s0379-0738(98)00120-0)
- Toneva DS, Nikolova S, Georgiev I, Tchorbadjieff A (2016) Intra- and interobserver measurement error of linear measurements on three-dimensional computed tomography models of dry mandibles. *Acta morphologica et anthropologica* 23
- Ubelaker DH, DeGaglia CM (2017) Population variation in skeletal sexual dimorphism. *Forensic Sci Int*. <https://doi.org/10.1016/j.forsciint.2017.06.012>
- Ulijaszek SJ, Lourie JA (1994) Intra- and inter-observer error in anthropometric measurement. In: Ulijaszek SJ, Mascie-Taylor CGN (eds) *Anthropometry: the individual and the population*. University Press, Cambridge

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)



Technical Note

A novel method for analyzing long bone diaphyseal cross-sectional geometry. A GNU Octave CSG Toolkit



Andreas Bertatos*, Maria-Eleni Chovalopoulou

Department of Animal and Human Physiology, Faculty of Biology, School of Sciences, National and Kapodistrian University of Athens, Panepistimiopolis, GR 157 01 Athens, Greece

ARTICLE INFO

Article history:

Received 14 December 2018

Received in revised form 16 January 2019

Accepted 29 January 2019

Available online 5 February 2019

Keywords:

Cross sectional geometry

Long bones

GNU Octave language

Second moments of area

3D models

ABSTRACT

The diaphyseal cross-sectional geometric properties of the humerus, femur and tibia have been extensively used for studying their adaptation to mechanical loading. To date common practices for such studies involve either computed tomography or the latex cast method in conjunction with image analysis for calculating such properties. With the advent of modern laser scanning and photogrammetry technologies in biological anthropology, the computation of the cross-sectional geometric properties directly from 3D models is a viable and sensible alternative. Nevertheless, such method has not been properly implemented as yet. A dedicated toolkit, named long-bone-diaphyseal-CSG-Toolkit, comprising a set of functions for the GNU Octave programming language, is presented here. Offering a robust analytical implementation and an easy to follow application either for a single bone or in batch-processing mode, the toolkit requires minimum user intervention and also provides functionality for graphical representation of the calculated periosteal contours and their respective cross-sectional geometric properties. Finally, the long-bone-diaphyseal-CSG-Toolkit utilizes advanced optimization algorithms, which eliminate intra- and inter-observer error by reliably orienting the cross-sectional contours to a well-defined orientation and close to the bone's true anatomical position, which provides a significant advantage over the latex cast method.

© 2019 Elsevier B.V. All rights reserved.

1. Introduction

Cross-sectional geometric (CSG) properties, such as area and second moments of area, of the long bone diaphyses have been extensively used for studying the mechanical loading of the long bone shafts [1–6]. CSG properties, which emerged from engineering-based beam modeling theory and applications [7], provide a shape-related quantified measure of the bone's mechanical performance, which relates to its resistance to bending forces [8]. Proper calculation of CSG properties requires the quantification of both periosteal and endosteal contours [2], which is paramount when comparing individuals [9] or studying factors such as growth and aging [10]. This implies that computed tomography (CT) is the only available non-destructive method that can provide accurate images of both endosteal and periosteal contours [11]. Nevertheless, CSG properties can be approximated by quantifying only the periosteal cross-sectional contour [9,12]. Furthermore, CSG properties of the periosteal contours alone can have significant value in studying the behaviorally determined

mechanical environment of past populations [9], since periosteal apposition may be more dependent on mechanical stimuli, while endosteal apposition may be more sensitive to nutritional influences [10,13]. Finally, studying CSG of periosteal contours relaxes the requirement for CT facilities and allows alternative methods for quantifying CSG properties, such as the latex cast method (LCM), which involves moulding the periosteal contour, digital tracing of the latex cast and image analysis of the periosteal cross-section with suitable software such as ImageJ's BoneJ plugin [14] for calculating the CSG properties. Despite LCM's popularity, the method requires significant labor in terms of cast preparation and image preprocessing prior to analysis and is prone to observer error either during cast digitization or image orientation.

The past decade witnessed the increased utilization of bone 3D modeling in anthropological studies. The advent of affordable 3D laser scanning technologies as well as the maturation of 3D photogrammetry software has attracted the interest of many researchers in the field due to their advantages in capturing in detail the morphological variation of the bones. The field of forensic anthropology adds no exception to this trend, where 3D models are increasingly being used for developing methods for biological profiling [15] and more recently they have been even utilized as demonstrative evidence in court cases [16]. Accordingly,

* Corresponding author.

E-mail address: abertatos@biol.uoa.gr (A. Bertatos).

the availability of digital 3D models of skeletal elements has led to the development of novel algorithms usually accompanied by software implementations that aim to quantify the morphological variation captured by laser scanners and 3D photogrammetry modeling. In the forensic context, such methods are of utmost importance, since they most often supersede the accuracy of visual observation methods while maintaining the benefits associated with 3D modeling regarding reproducibility of the results and observer bias.

The present paper introduces a novel method for analyzing the periosteal diaphyseal CSG properties of the humerus, femur and tibia bones. Although the idea of analyzing laser scanner data for estimating the CSG properties of long bone diaphyses is certainly not new and a previous implementation has been reported [17], the software introduced in the present paper comprises a robust implementation of functions in GNU Octave programming language [18] that automatically analyze the periosteal diaphyseal CSG properties for any of the aforementioned long bones. The developed set of functions can work on a single bone or in batch processing mode and handles in a concise manner all matters regarding the analysis of periosteal diaphyseal CSG properties including the automatic alignment of cross-sectional areas, the optimization of the anatomical orientation for each bone, the output of the results as well as the visualization of the cross-sectional areas. The true novelty of the presented algorithm, though, is that it eliminates any observer-induced error from the analytical processing of the triangular mesh of any given long bone and allows the user to easily retrieve the CSG properties of the available skeletal sample and continue with the anthropological question at hand rather than worrying about the accuracy of tracing a latex cast or properly orienting a two-dimensional cross-sectional area image in black and white about the anatomical axis of the bone it originated. The accuracy of the CSG properties calculated by the novel method and accompanied software implementation presented in this paper is bounded only by the accuracy of the bone's 3D model representation as a triangular mesh and the double-precision floating point arithmetic limitations of modern computers.

2. Software requirements and algorithm description

The algorithm implemented in the “long-bone-diaphyseal-CSG-Toolkit: A GNU Octave CSG Toolkit” [19] requires minimum intervention from the user, who is only expected to define two points on the bone's triangular mesh as an initial alignment of its

anatomical position. This can be easily accomplished by using the point picking tool from Meshlab software [20] and saving the point coordinates in the appropriate (.pp) file format. The implementation of the CSG Toolkit was based on the assumption that the user needs no programming skills or prior knowledge of the GNU Octave programming environment in order to use the toolkit and calculate the CSG properties of the long bone at hand. The only requirements for successfully utilizing the CSG Toolkit are that GNU Octave along with the ‘io’, ‘statistical’ and ‘geometry’ packages must be installed, each bone to be analyzed must be represented as a single triangular mesh in a separate (.obj) file, which must conform to Wavefront OBJ file format, and finally, each bone's triangular mesh file and its Meshlab picked point complement file must share a common file name and be in the same working directory with the CSG Toolkit functions, which may be freely downloaded from <https://github.com/pr0m1th3as/long-bone-diaphyseal-CSG-Toolkit>.

The CSG Toolkit loads the faces and their corresponding vertices from the .obj file and calculates the long bone's maximum distance, which is used as an initial estimate of the proximal-distal axis. Subsequently, it slices the mesh at 20%, 35%, 50%, 65% and 80% of its maximum distance and calculates the centroid of each cross-sectional area. The centroids of the 1st and 5th cross-sections are utilized for defining the optimal vertical anatomical position of the bone, since the longitudinal axis of the bone's diaphysis provides a better estimate of its vertical anatomical position rather than its maximum distance (especially in femur). However, it should be pointed out that the optimal vertical axis, which in practice is the neutral axis of bending, does not usually coincide with the bone's true anatomical vertical axis, but it was implemented in the CSG Toolkit as a sensible, well-defined good approximation of the latter. The proximal-to-distal (transverse plane) along with the posterior-anterior (coronal plane) orientation of the bone's anatomical position is then calculated by utilizing the user defined points provided in the relevant .pp file. The user defined points, although specific for each type of bone, that is for humerus, femur and tibia, do not need to be accurately defined since the CSG Toolkit performs automatic optimization of their location according to the explicit geometry of each bone. Nevertheless, for humerus the two points need to be located on the anterior side of the trochlea and capitulum respectively. For femur the points need to be defined posteriorly of the medial and lateral condyles, whereas the points for the tibia need to be located posteriorly of the medial and lateral condyles. The order by which the user defines the two points is irrelevant and this applies to both left and right side bones. The

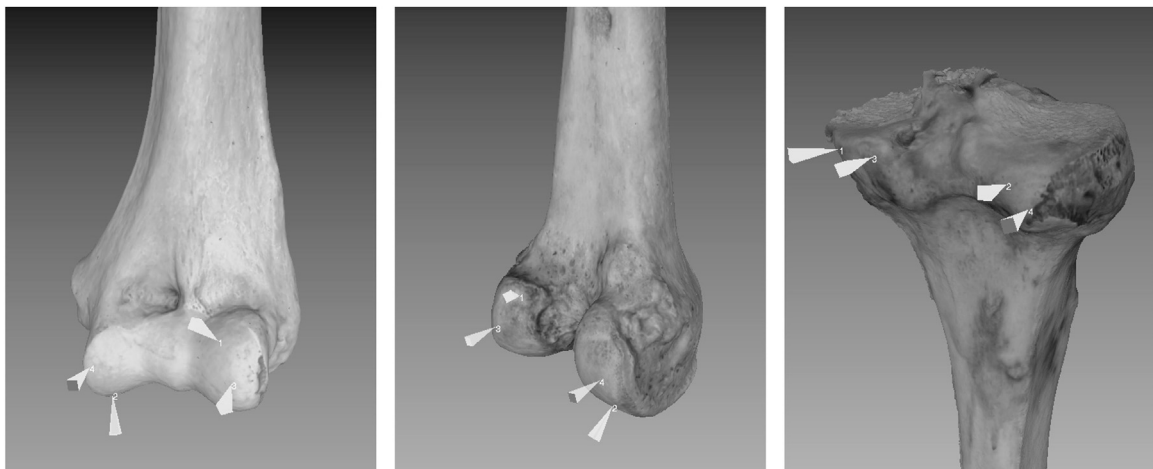


Fig. 1. Alignment points retrieved from Meshlab's '.pp' file. Left-hand side: under-defined (1, 2) and optimized (3, 4) points of alignment for humerus; middle: under-defined (1, 2) and optimized (3, 4) points of alignment for femur; right-hand side: under-defined (1, 2) and optimized (3, 4) points of alignment for tibia.

CSG Toolkit will optimize the locations of the two points and consequently the orientation of the bone's anatomical position as long as the initial points are picked anywhere near the aforementioned locations. The optimized point locations for each bone are explicitly described as follows. For the humerus, the two most anterior points of the trochlea and capitulum respectively. For the femur, the two most posterior points of the medial and lateral condyles respectively. For the tibia, the two most posterior points of the bone's surface located on the back and below the medial and lateral facets respectively.

Fig. 1 illustrates three examples of initial and optimized points on a humerus, femur and tibia. Points numbered 1 and 2 are user defined, whereas 3 and 4 are the optimized corresponding points. The optimized points are considered to lie on the coronal plane and since for each bone, these points are defined either on the distal (for humerus and femur) or the proximal (for tibia) end, the CSG Toolkit calculates the exact orientation of the bone's anatomical position. The optimization algorithm utilizes the already defined vertical axis in conjunction with the user defined points and finds the corresponding local extreme points (anteriorly or posteriorly depending to the bone selected for analysis), which are regarded as the optimal points that lie on the coronal plane. It should be noted that although for humerus and femur the optimized points properly approximate the bone's true anatomical orientation, this may not be true for the tibia, since the medial condyle usually protrudes behind the coronal plane when the bone is oriented in its true anatomical position. Nevertheless, it was decided to use this optimization alignment for the tibia as well, since it provides a robust and reproducible method that eliminates observer error from the calculation of the anatomical position, which is subsequently used for calculating axis oriented parameters of second moments of area, such as I_x and I_y , from the bone's cross-sectional contours.

After calculating the optimal orientation of the bone's anatomical position, the algorithm implemented in the CSG Toolkit determines the order of the centroids of the initial cross-sections from proximal to distal with the 1st centroid located at 20% length from the proximal end and the 5th at 80% towards the distal end. Subsequently, new cross-sectional contours are calculated at the "height" of each centroid and locally perpendicular to the diaphyseal axis as defined by each consecutive pair of centroids. More specifically, the normal of the 1st slicing plane is defined by the 1st and 2nd centroids, the 2nd slicing plane normal is defined by centroids 1 and 3, the 3rd is determined by centroids 2 and 4, etc. Fig. 2 illustrates the cross-sectional contours calculated on a left humerus bone analyzed with the CSG toolkit. Each cross-sectional contour is mapped onto local 2D coordinates with the local x-axis co-planar with the coronal plane and its positive side facing towards the right side and the local y-axis co-planar with the sagittal plane and its positive side facing towards the front.

Finally, after calculating the anatomically oriented diaphyseal cross-sections, the CSG Toolkit calculates the perimeter and area of

each cross-sectional contour along with its centroid coordinates and the respective unit normal vectors for the slicing and coronal planes in R3 (3 dimensional space) of the original model. Additionally, the $I_x, I_y, I_{xy}, I_{MIN}, I_{MAX}$ and θ_m parameters of the second moment of area are calculated, whereas the points of polygons that constitute each cross-sectional contour are returned both in local aligned 2D coordinates as well as in original R3. All aforementioned calculated values are returned in comma separated value (CSV) files, which are explained in detail in the following section, whereas the optimized alignment points are appended in the original Meshlab picked-points file.

3. A working example of the CSG Toolkit

As mentioned above, the CSG Toolkit relies on certain GNU Octave extra packages that need to be installed on the GNU Octave core environment. These necessary packages are the 'io', 'statistics' and 'geometry' packages and their latest release can be directly downloaded and installed by issuing a simple command on the GNU Octave command window. For example, installing the latest 'statistics' package requires the user to issue the following command:

```
>> pkg install -forge statistics
```

If all necessary extra packages are installed on the GNU Octave environment, the user is only required to navigate through GNU Octave's file browser to the working folder where the CSG Toolkit functions along with the relevant .obj and .pp files for analysis reside and issue the following command:

```
>> longbone_Analysis
```

The software will open a dialog box and ask the user for the type of bone to be analyzed. The user may select either 'Humerus', 'Femur' or 'Tibia' and subsequently the CSG Toolkit will enter batch mode by scanning the working folder for all available .obj files and analyze them sequentially. When using the CSG Toolkit in batch mode the user must ensure that all available .obj files correspond to a specific type of bone, since the software will regard all available .obj files containing the specific bone selected at the beginning of the batch processing. For each triangular mesh available in the working directory the software will return four CSV files with the following filename convention. The files 'geometry-name.csv', 'inertia-name.csv', 'polyline2D-name.csv' and 'polyline3D-name.csv' contain all the available results of the CSG analysis, where 'name' is the filename of the 'name.obj' file containing the analyzed triangular mesh. The structure of the aforementioned CSV files is summarized in Table 1. Additionally, the user may display the help file regarding the usage of the 'longbone_Analysis.m' script and detailed structure of the generated CSV files by issuing the command:

```
>> help longbone_Analysis
```

Except for the batch processing mode, the CSG Toolkit provides the user with the capacity of analyzing single bones by direct utilization of the 'longbone_Geometry.m' function, which relaxes

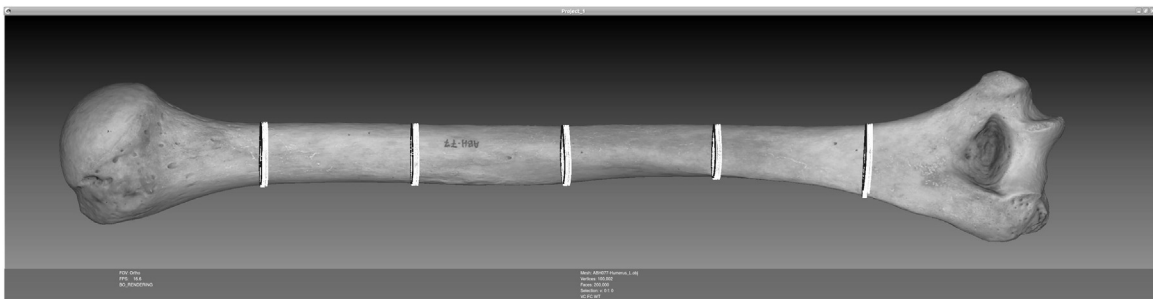


Fig. 2. Cross-sectional contour points of a humerus retrieved from 'polyline3D-name.csv' file.

Table 1
Structure of the CSV files produced by the 'longbone_Analysis.m' script.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	
<i>'geometry-name.csv'</i>																
Contour at 20%	1	Area	Per.	C _X	C _Y	C _Z	SP _X	SP _Y	SP _Z	CP _X	CP _Y	CP _Z				
–	–	–	–	–	–	–	–	–	–	–	–	–				
Contour at 80%	5	–	–	–	–	–	–	–	–	–	–	–				
<i>'inertia-name.csv'</i>																
Contour at 20%	1	I _X	I _Y	I _{XY}	I _{MIN}	I _{MAX}	θ _m									
–	–	–	–	–	–	–	–									
Contour at 80%	5	–	–	–	–	–	–									
<i>'polyline2D-name.csv'</i>																
Point 1	1	CS1	CS1	CS2	CS2	CS3	CS3	CS4	CS4	CS5	CS5					
–	–	x1	y1	x1	y1	x1	y1	x1	y1	x1	y1					
Point N	N	–	–	–	–	–	–	–	–	–	–					
<i>'polyline3D-name.csv'</i>																
Point 1	1	CS1	CS1	CS1	CS2	CS2	CS2	CS3	CS3	CS3	CS4	CS4	CS4	CS5	CS5	CS5
–	–	x1	y1	z1	x1	y1	z1	x1	y1	z1	x1	y1	z1	x1	y1	z1
Point N	N	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–

C_{XYZ}: centroid coordinates in R3; SP_{XYZ}: normal vector of the sectioning plane; CP_{XYZ}: normal vector of the coronal plane;
CS1-5X,Y: cross sectional points in local 2D coordinates for each contour; CS1-5X,Y,Z: cross sectional points in R3 for each contour.
* Row numbers (1–5) and column letters (A–O) illustrate the .csv structure as opened in a regular spreadsheet.

the restriction of a single type of bone that may be present in the working directory. Additionally, instead of saving the results in the relevant CSV files, the 'longbone_Geometry.m' function returns all results in the workspace of the GNU Octave for any further statistical (or other) analysis the user may wish to conduct. The 'longbone_Geometry.m' function may be utilized as in the following example:

```
>> [CS_Geometry, SMoA, polyline] = longbone_Geometry("name.obj", "Humerus")
```

Where "name.obj" contains a humerus and of course the required Meshlab point file named "name.pp" is also present in the working directory. Additional help for the usage of the 'longbone_Geometry.m' function may be accessed with the following command:

```
>> help longbone_Geometry
```

Finally, the CSG Toolkit contains a function, namely 'visualize_CrossSections.m', which plots the cross-sectional contours along with their calculated CSG properties. The function works as a stand-alone application and requires only the produced CSV files to read the outcome of the CSG analysis and produce 5 plots, one for each cross-sectional contour. The function may be called by typing:

```
>> visualize_CrossSections("name")
```

Alternatively, the function may also be utilized to visualize the results that are saved in GNU Octave's workspace, when a particular bone has been analyzed with the 'longbone_Geometry.m' function. You may type "help visualize_CrossSections" at the command window for further information regarding its usage. Fig. 3 shows collectively the 5 cross-sectional contours of the humerus illustrated in Fig. 2, which have been plotted with 'visualize_CrossSections.m' function.

4. Testing the accuracy and reliability of the CSG Toolkit

A series of tests have been performed to assess the integrity as well as the reproducibility of the results of the CSG Toolkit. Three 3D mesh bone models (a humerus, a femur and a tibia, which are part of the Athens modern reference skeletal collection) and one 3D mesh "ground-truth" model (fabricated with known geometric properties with CAD software) were used for testing purposes. All four models along with their respective "longbone_Analysis"

generated CSV files are included in the testing dataset [21], which is freely available as supplementary material of the CSG Toolkit. A fully detailed description of these tests along with their associated results and illustrative figures including plots, renderings and screenshots are reported in a separate validation report [22], which is also freely available as supplementary material to the CSG Toolkit along with the aforementioned corresponding testing dataset.

The first test concerns the optimization algorithm for finding the optimal locations of the user defined orientation points for each type of bone. A humerus, a femur and a tibia 3D model were analyzed repeatedly with the CSG Toolkit each time with a different pair of user defined initial points to validate the stability of the algorithm. Each time the optimized points, which are also reported in GNU Octave's command window by the CSG Toolkit, were exactly the same. Although proof of repeatability was made on the bone models, the integrity of CSG Toolkit's calculations regarding the optimized orientation points was positively tested on the "ground-truth" model, whose optimal points are known, with the returned values being a perfect match. The exact numbers and related figures of this test are provided in the "optimizing anatomical orientation" section of the supplementary validation report (pp. 1–6).

The next test regards the proof of correct calculation of the CSG properties and was based on the ground-truth model. Calculated and corresponding theoretical values are provided in the respective section of the supplementary validation report (pp. 7–12). The ground-truth model comprises of 5 basic shapes (square, ellipse, hexagon, rectangle and circle), which are properly aligned at the maximum length intervals where the CSG Toolkit calculates the cross-sectional contours from. For the regular polygons of the ground-truth model, all calculated CSG properties were identical with the theoretical values based on the known dimensions of each polygon, whereas small deviations were observed for the circle and elliptical shapes, since the latter are represented in the triangular mesh 3D model as high-count polygon approximations inscribed in the theoretical shapes. Nevertheless, the outcome of this test proves that the CSG Toolkit implementation accurately calculates all CSG properties and its precision is only bound by the accuracy of the 3D mesh model that is supplied. As an extension to this test, the cross-sectional contours calculated with the CSG Toolkit from the

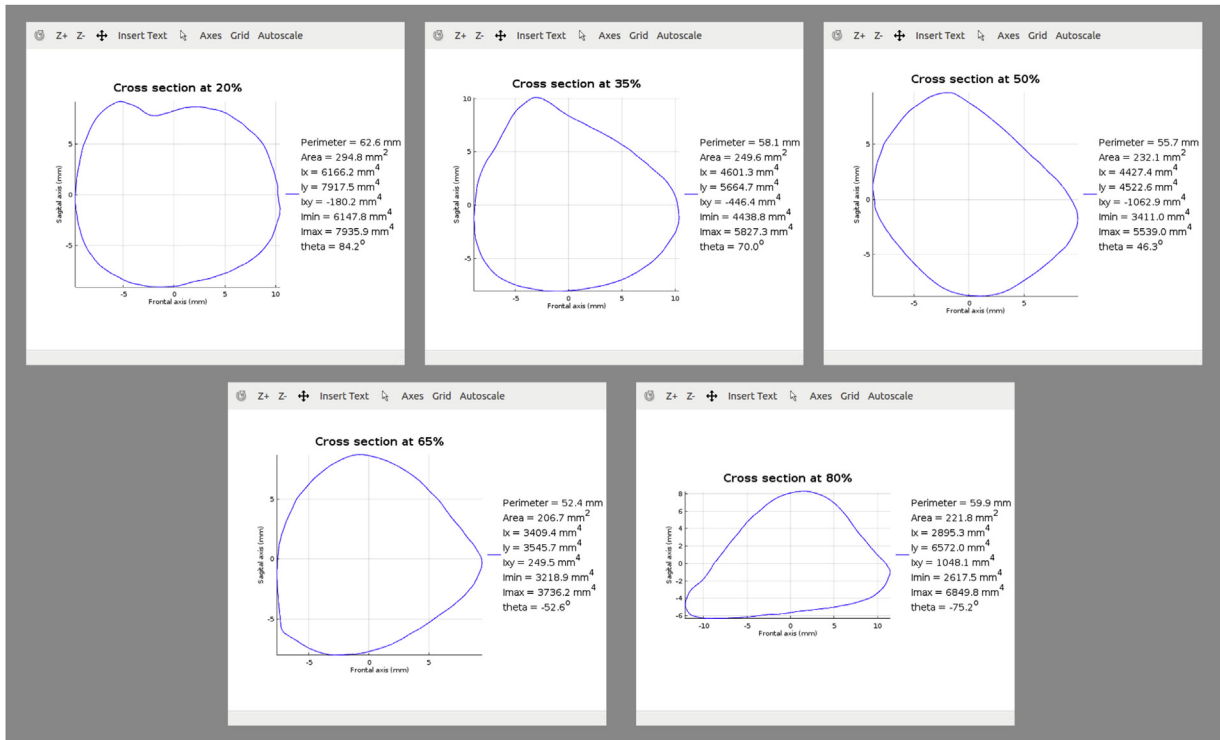


Fig. 3. Cross-sectional contour plots and related CSG properties of a humerus produced with ‘visualize_CrossSections.m’ function.

ground-truth model and plotted with the “visualize_CrossSections.m” function were gray-scaled and analyzed with ImageJ software. This test aimed to demonstrate relative consistency with an alternative method of calculating CSG properties but also to gain some insight about how much error is introduced from image-based CSG calculations. Table 2 summarizes the relative errors of CSG Toolkit and ImageJ results calculated as the percentage of deviation from the theoretical values of each shape. Note that values of I_x , I_y and I_{xy} are not reported by ImageJ software although it computes them internally, hence they are not included in the comparison. Additionally, error regarding angle θ is reported as absolute deviation in degrees.

Testing so far has established the reliability of the algorithm implementation in the CSG Toolkit as well as some measure regarding the error induced by the inherent noise of image pre-processing and image-based calculations as opposed to the analytical geometry approach of the CSG Toolkit, which is only bound by floating point precision errors. The last test of this series makes a direct comparison of calculating the CSG properties from real bones with the LCM and subsequently with ImageJ as opposed to analyzing their respective 3D models

with the CSG Toolkit. As with the previous test, the cross-sectional contours of the 3D models are plotted and further analyzed with ImageJ to establish some comparison measure. Detailed results are provided in the relevant section of the supplementary validation report (pp. 13–19). Our findings showed that both methods produce relatively consistent results. More specifically, both scanned images from LCM moulds and images from plotted cross-sectional contours yielded similar relative errors with respect to the values calculated with the CSG Toolkit. Moreover, these relative errors are comparable to those observed with the ground-truth model, which implies that most of the error relates to image pre-processing (i.e. scaling) and image-based analysis of the CSG properties instead of the moulding procedure with the LCM. Of course, some observer-error relates to manual measurement of the slicing distance along the bone’s diaphysis as well as manually slicing the LCM mould perpendicular to the bone’s shaft. The only exception regarding the consistency of the observed errors concerns the angle θ , which yielded much greater variability for the LCM among the cross-sectional contours of the same bone. The absolute errors of angle θ are summarized in Table 3. The observed variation is due

Table 2

Relative errors in calculating the CSG properties of the “ground-truth” model’s cross-sectional contours between CSG Toolkit and ImageJ.

CSG properties	long-bone-diaphyseal-CSG-Toolkit					ImageJ with BoneJ plugin				
	Perimeter	Area	I_{MIN}	I_{MAX}	θ^a	Perimeter	Area	I_{MIN}	I_{MAX}	θ^a
Circle	0.27%	0.80%	1.74%	1.42%	–	4.72%	1.75%	5.17%	4.87%	–
Rectangle	0	0	0	0	0	6.86%	0.44%	2.86%	2.68%	0
Regular hexagon	0	0	0	0	0.4 ^b	5.33%	1.11%	0.06%	1.23%	1.4°
Ellipse	0.10%	0.66%	1.91%	0.72%	0	5.23%	0.58%	4.02%	2.95%	0.1°
Square	0	0	0	0	0	0.33%	1.02%	1.64%	2.35%	1.1°

^a Angle θ is reported as absolute error in degrees.

^b The slight difference observed in θ angle of the hexagon is an artifact of rounding errors due to the way the triangular mesh of the “ground-truth” model was segmented from the initially fabricated 3D solid in AutoCAD.

Table 3

Absolute errors of angle θ between image-based CSG calculations with ImageJ and CSG Toolkit's results.

	Humerus		Femur		Tibia	
	θ_{LCM}^a	θ_{PLOT}^b	θ_{LCM}^a	θ_{PLOT}^b	θ_{LCM}^a	θ_{PLOT}^b
Contour at 20%	2.1	0.1	16.5	0.1	8.2	0.2
Contour at 35%	3.7	0.1	11.2	0.1	2.3	0.1
Contour at 50%	1.6	0	12.9	0.2	4.4	0.1
Contour at 65%	4.1	0.6	7.4	0	3.3	0.1
Contour at 80%	3.1	0.2	6.6	0.1	13.7	0.7

^a Angle θ_{LCM} is reported as the absolute error in degrees between scanned images of LCM moulds and CSG Toolkit's reported values.

^b Angle θ_{PLOT} is reported as the absolute error in degrees between plotted images of cross-sectional contours and CSG Toolkit's reported values.

to observer error in properly and consistently aligning the LCM contours with the bone's anatomical position, which emphasizes the benefits of the automatic alignment performed with the CSG Toolkit.

5. Conclusion

The CSG Toolkit introduced in this technical note provides the user with an efficient, user-friendly and analytically robust alternative to the LCM for calculating the periosteal diaphyseal CSG properties of long bones. Moreover, the implementation of specific optimization algorithms eliminating the observer error in "manual" orientation of the Ix and Iy second moments of area, which is inevitable when LCM is applied, not only allows for direct comparison between data sets analyzed by different researchers, but also allows a far more accurate interpretation of these values, which reflect how the bone's cross-sectional area (and by extension its mass) is distributed along its anatomical axes. The increased accuracy and repeatability of our proposed method allows researchers to extract geometric information from long bones while quantitatively capturing subtle variation between specimens. Such precision is not yet available by routinely used methods involving measurements or visual observation, which may provide advancements beyond the limited scope of mechanical loading studies. In forensic applications, the CSG Toolkit, a fast yet reliable method for retrieving information of the periosteal diaphyseal morphology, can facilitate the development of novel methods in element pair-matching and osteometric sorting, which would make significant contributions to cases of commingled burials and mass disaster victim identification. Furthermore, evaluating diaphyseal periosteal contour geometry for sex identification may provide advantages over sex discriminant functions based on maximum length, which are prone to secular change. The open-source nature of the presented implementation provides easy adaptation of the source code to specific needs such as arbitrary cross-sections along the bone and allows further expansion of its use to other skeletal elements or to non-human species.

Great effort has been made to properly implement and debug the underlying code of the CSG toolkit as well as test its output on a number of bones from the Athens Collection. Nevertheless, there is always space for improvement and it is highly recommended to utilize the 'visualize_CrossSections.m' function for visual inspection of the calculated cross-sectional contours before further analysis of the CSG properties retrieved from the CSV files. As with any newly implemented optimization algorithm, there is always a chance that some extreme case of input values may produce irrational results. The CSG Toolkit is published under the terms of the GNU General Public License and any further contributions or even requests for implementing extra functionality are welcomed and encouraged.

Competing interests

The author have no conflict of interest to declare.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Author contributions

A.B. conceived and developed the method, made the software implementation and produced the required models for testing. M.-E.C. contributed in testing the repeatability and reliability of the results, independent platform testing of the software. Both authors contributed in writing the present manuscript as well as the supplementary validation report.

Acknowledgements

We would like to thank Ioannis Galyfas for his insight into the appropriate orientation of principal and rotated axes of second moments of area, Dr. Efthymia Nikita for her advice on how to properly orient the LCM moulds to the bone's anatomical position and Dr. Afrodite Kakaboura for supplying the material (HydroXtreme putty soft) for the LCM moulds used in validating the CSG Toolkit's results.

References

- [1] J. Stock, S. Pfeiffer, Linking structural variability in long bone diaphyses to habitual behaviors: foragers from the southern African Later Stone Age and the Andaman Islands, *Am. J. Phys. Anthropol.* 115 (4) (2001) 337–348, doi:http://dx.doi.org/10.1002/ajpa.1090.
- [2] M.C. O'Neill, C.B. Ruff, Estimating human long bone cross-sectional geometric properties: a comparison of noninvasive methods, *J. Hum. Evol.* 47 (4) (2004) 221–235, doi:http://dx.doi.org/10.1016/j.jhevol.2004.07.002.
- [3] I.S. Maggiano, M. Schultz, H. Kierdorf, T.S. Sosa, C.M. Maggiano, V. Tiesler Blos, Cross sectional analysis of long bones, occupational activities and long distance trade of the Classic Maya from Xcambó—archaeological and osteological evidence, *Am. J. Phys. Anthropol.* 136 (4) (2008) 470–477, doi:http://dx.doi.org/10.1002/ajpa.20830.
- [4] L. Shackelford, F. Marshall, J. Peters, Identifying donkey domestication through changes in cross-sectional geometry of long bones, *J. Archaeol. Sci.* 40 (12) (2013) 4170–4179, doi:http://dx.doi.org/10.1016/j.jas.2013.06.006.
- [5] M.E. Cameron, S. Pfeiffer, Long bone cross-sectional geometric properties of Later Stone Age foragers and herder–foragers, *South Afr. J. Sci.* 110 (9/10) (2014) 1–11, doi:http://dx.doi.org/10.1590/sajs.2014/20130369.
- [6] M.E. Zaki, A.A. Azab, W. Yousef, E.Y. Wassal, H.T. El-Bassyouni, Cross-sectional analysis of long bones in a sample of ancient Egyptians, *Egypt. J. Radiol. Nucl. Med.* 46 (3) (2015) 675–681, doi:http://dx.doi.org/10.1016/j.ejnm.2015.03.008.
- [7] R. Huiskes, On the modeling of long bones in structural analyses, *J. Biomech.* 15 (1) (1982) 65–69, doi:http://dx.doi.org/10.1016/0021-9290(82)90036-7.
- [8] C.B. Ruff, E. Trinkaus, A. Walker, C.S. Larsen, Postcranial robusticity in Homo. I: temporal trends and mechanical interpretation, *Am. J. Phys. Anthropol.* 91 (1) (1993) 21–53, doi:http://dx.doi.org/10.1002/ajpa.1330910103.
- [9] V. Sparacello, O. Pearson, The importance of accounting for the area of the medullary cavity in cross-sectional geometry: a test based on the femoral midshaft, *Am. J. Phys. Anthropol.* 143 (4) (2010) 612–624, doi:http://dx.doi.org/10.1002/ajpa.21361.
- [10] C.B. Ruff, E. Garofalo, M.A. Holmes, Interpreting skeletal growth in the past from a functional and physiological perspective, *Am. J. Phys. Anthropol.* 150 (1) (2013) 29–37, doi:http://dx.doi.org/10.1002/ajpa.22120.
- [11] C.B. Ruff, Biomechanical analyses of archaeological human skeletons, in: M.A. Katzenberg, S.R. Saunders (Eds.), *Biological Anthropology of the Human Skeleton*, 2nd ed., Wiley, Hoboken, NJ, 2008, pp. 183–206.
- [12] A.A. Macintosh, T.G. Davies, T.M. Ryan, C.N. Shaw, J.T. Stock, Periosteal versus true cross-sectional geometry: a comparison along humeral, femoral, and tibial diaphyses, *Am. J. Phys. Anthropol.* 150 (3) (2013) 442–452, doi:http://dx.doi.org/10.1002/ajpa.22218.
- [13] S. Mays, R. Ives, M. Brickley, The effects of socioeconomic status on endochondral and appositional bone growth, and acquisition of cortical bone in children from 19th century Birmingham, England, *Am. J. Phys. Anthropol.* 140 (3) (2009) 410–416, doi:http://dx.doi.org/10.1002/ajpa.21076.
- [14] M. Doube, M.M. Kłosowski, I. Arganda-Carreras, F. Cordelières, R.P. Dougherty, J. Jackson, B. Schmid, J.R. Hutchinson, S.J. Shefelbine, BoneJ: free and extensible

- bone image analysis in ImageJ, *Bone* 47 (6) (2010) 1076–1079, doi:<http://dx.doi.org/10.1016/j.bone.2010.08.023>.
- [15] K.L. Colman, J.G. Dobbe, K.E. Stul, J.M. Ruijter, R. Oostra, R.R. Van Rijn, A.E. van der Merwe, H.H. de Boer, G.J. Streekstra, The geometrical precision of virtual bone models derived from clinical computed tomography data for forensic anthropology, *Int. J. Legal Med.* 131 (4) (2017) 1155–1163, doi:<http://dx.doi.org/10.1007/s00414-017-1548-z>.
- [16] W. Baier, J.M. Warnett, M. Payne, M.A. Williams, Introducing 3D printed models as demonstrative evidence at criminal trials, *J. Forensic Sci.* 63 (4) (2018) 1298–1302, doi:<http://dx.doi.org/10.1111/1556-4029.13700>.
- [17] T.G. Davies, C.N. Shaw, J.T. Stock, A test of a new method and software for the rapid estimation of cross-sectional geometric properties of long bone diaphyses from 3D laser surface scans, *Archaeol. Anthropol. Sci.* 4 (4) (2012) 277–290, doi:<http://dx.doi.org/10.1007/s12520-012-0101-8>.
- [18] J.W. Eaton, D. Bateman, S. Hauberg, R. Wehbring, GNU Octave Version 4.2.1 Manual: A High-level Interactive Language for Numerical Computations URL, (2017) . <https://www.gnu.org/software/octave/doc/v4.2.1/>.
- [19] A. Bertsatos, long-bone-diaphyseal-CSG-Toolkit: a GNU Octave CSG Toolkit (Version v1.0.1), Zenodo (2018), doi:<http://dx.doi.org/10.5281/zenodo.1467045>.
- [20] P. Cignoni, M. Callieri, M. Corsini, M. Dellepiane, F. Ganovelli, G. Ranzuglia, MeshLab: an open-source mesh processing tool, Sixth Eurographics Italian Chapter Conference, (2008) , pp. 129–136.
- [21] A. Bertsatos, M.-E. Chovalopoulou, Testing dataset for the "long-bone-diaphyseal-CSG-Toolkit", Zenodo (2018), doi:<http://dx.doi.org/10.5281/zenodo.1466962>.
- [22] A. Bertsatos, M.-E. Chovalopoulou, Validation study of correct operation for the "long-bone-diaphyseal-CSG-Toolkit", Zenodo (2018), doi:<http://dx.doi.org/10.5281/zenodo.1466135>.



Advancements in sex estimation using the diaphyseal cross-sectional geometric properties of the lower and upper limbs

Andreas Bertsatos¹ · Nefeli Garoufi¹ · Maria-Eleni Chovalopoulou^{1,2}

Received: 11 June 2020 / Accepted: 30 September 2020
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

Abstract

This paper introduces an automated method for estimating sex from the lower and upper limbs based on diaphyseal CSG properties. The proposed method was developed and evaluated using 389 femurs, 412 tibias, and 404 humeri of adult individuals from a modern Greek reference sample, the Athens Collection. The skeletal properties, which were extracted with the CSG-Toolkit, were analyzed with step-wise DFA (evaluated with LOOCV) and subsequently with RBF kernel SVM supervised learning. SVM cross-validation was based on a 20-fold stratified random sample splitting as well as a chronological split based on year of birth to further assess the effect of secular change in sex estimation capacity. Maximum cross-validated classification accuracy from step-wise DFA reached 94.8% for the femur, 94.7% for the tibia, and 97.3% for the humerus, whereas SVM cross-validated results were similar although slightly lower, mainly due to the more strict cross-validation scheme. Our results suggest that the proposed sex estimation method is reasonably robust to secular change, since there was limited loss in classification accuracy between different chronological groups, despite the presence of secular change in stature of the Greek population during the examined period. The proposed method has been implemented as a function for the GNU Octave environment, named `estimate_sex`, which comprises a self-intuitive graphical user interface for facilitating sex estimation and is freely available under a suitable license.

Keywords Lower and upper limbs · CSG properties · Sex estimation software · 3D models · Machine learning · Greek population

Introduction

Sex estimation is one of the most important steps towards profiling unidentified human remains. When the pelvic bones are available, forensic anthropologists typically rely on the os coxa as the most reliable source of sex information by either visual assessment [1–3] or metric evaluation [4–7]. However, it is often the case that sex identification has to rely on the cranium or other postcranial elements. Their importance and utility in sex identification are further corroborated by the plethora of literature on bone-specific sex estimation methods,

although, admittedly, this abundance also relates to their population-specific expression of sexual dimorphism. Despite Spradley's and Jantz's [8] argument that the postcranial skeleton provides more reliable sex predictors than the cranium and a number of relevant works reporting high classification accuracy for specific appendicular elements [9–13], recent works on cranial sex estimation have demonstrated even higher sex diagnostic capacity [14–16].

Beyond the “Skull Versus Postcranial Elements” debate and the relevant rally of classification scores between skeletal elements, there is one other important differentiation. Recent works on the cranium have attracted an array of advanced analytical methods contributing to the advancements of its sex-diagnostic capacity. These methods range from geometric morphometrics, computational geometry, and computer vision for 3D mesh analysis to deep neural networks and support vector machine (SVM) supervised learning algorithms for statistical analysis [16–19]. At the same time, most research related to sex estimation based on the postcranial skeleton has been limited to traditional measurements [9, 13, 20–22], while few exceptions have focused on alternative

✉ Andreas Bertsatos
abertsatos@biol.uoa.gr

¹ Department of Animal and Human Physiology, Faculty of Biology, School of Sciences, University of Athens, Panepistimiopolis, GR 157 01 Athens, Greece

² Science and Technology in Archaeology and Culture Research Center, The Cyprus Institute, 2121 Aglantzia, Nicosia, Cyprus

methods [23–25]. This is especially relevant to the bones of the lower and upper limbs, whose sexual dimorphism has been typically assessed based on standard osteometric measurements [9, 11–13, 21, 22, 26–31]. Despite their ease of use, standard osteometric measurements do not necessarily provide the best sex discriminant variables [14] and their utilization poses further issues regarding inter-observer measurement error when it comes to applying sex discriminant functions in the forensic context. Moreover, the high correlation of linear measurements with body size and stature further amplifies the effect of population specificity, which is an inherent problem in forensic anthropology and the ever changing demographic structure of modern populations [30].

The present work focuses on the long bones of the lower (femur, tibia) and upper limbs (humerus) and it introduces a novel sex estimation method, which tries to overcome the previously discussed limitations of standard osteometric protocols and linear measurements while maintaining their ease of use for the forensic anthropologists, who have to rely on these bones for estimating sex of unidentified cases. The method primarily utilizes the long bones' diaphyseal cross-sectional geometric (CSG) properties, which have already proven their advantages in osteometric sorting by accurately quantifying delicate morphological variations [32]. This newly introduced sex estimation method relies on the CSG-Toolkit software [33], which can automatically compute all necessary sex-diagnostic metric parameters from any available 3D digital representation of a femur, tibia, or humerus bone, and it is accompanied by a freely available open-source application to facilitate sex estimation through an easy to operate graphical user interface.

Materials and methods

The present study utilizes 1205 skeletal elements of the lower and upper limbs of 220 adult individuals of known sex and age-at-death from the Athens Collection, which is housed at the Department of Animal and Human Physiology of the National and Kapodistrian University of Athens. More specifically, 389 femurs (194 left and 195 right) derived from 210 individuals, 116 males and 94 females, another 412 tibias (204 left and 208 right) derived from 214 individuals, 117 males and 97 females, and 404 humeri (201 left and 203 right) derived from 213 individuals, 118 males and 95 females were utilized. Only complete, well-preserved bones with no pathological deformation or surgical implants have been included in the present sample. The age-at-death of the included individuals spans from 18 to 99 years and they were born between 1879 and 1965 according to the death certificates [34]. All skeletal elements have been previously digitized into 3D triangular mesh models by means of photogrammetry with the Photoscan Pro v1.4 software (Agisoft LLC, Russia). The

produced 3D models have a rated accuracy of ~ 0.2 mm along their maximum length distance [35].

Each skeletal element has been processed with the latest version (v1.1.0) of the long-bone-diaphyseal-CSG-Toolkit, which facilitates the computation of the CSG properties directly from 3D models of femur, tibia, and humerus bones [36]. The CSG-Toolkit, which comprises a set of functions for the GNU Octave programming language [37], establishes the anatomical orientation of any given sample, it extracts five cross sections along its shaft at 20%, 35%, 50%, 65%, and 80% of its maximum length and it calculates the perimeter, the area, and the I_X , I_Y , I_{XY} , I_{MIN} , I_{MAX} , θ_m parameters of the second moments of area for every cross section. Furthermore, it calculates the maximum length of the sample as well as the cross-sectional normal vectors, which were used to calculate the dihedral angles between each pair of adjacent slicing planes. These dihedral angles and their total sum reflect the bending of the diaphyseal shaft. Further information regarding the CSG-Toolkit's usage and operation can be found in the original publication [33] and its supplementary validation report [38]. After initial testing of the 46 variables computed by the CSG-Toolkit for sexual dimorphism, the I_{XY} and θ_m parameters of each cross section as well as the bending angles of the diaphyseal shaft were excluded. Complementary to the remaining variables, another three parameters were calculated for each cross section. The first parameter is the *ArPerIndex*, which is calculated as $\frac{Area \times 4\pi}{Perimeter^2}$ and provides a measure of circularity and smoothness of the cross section [39]. The other two parameters are the ratios of I_X/I_Y and I_{MAX}/I_{MIN} , which reflect the amount of cross-sectional elongation along the frontal (coronal) anatomical axis and the principal (maximum) axis of the 2nd moment of area, respectively. Finally, after the removal and addition of parameters, an updated list of 46 variables, which are

Table 1 List of variables utilized for sex estimation from the femur, tibia, and humerus 3D models

#	Variables
1	<i>Max distance</i> ^a
2, 11, 20, 29, 38	<i>Area</i> (20%, 35%, 50%, 65%, 80%) ^b
3, 12, 21, 30, 39	<i>Perimeter</i> (20%, 35%, 50%, 65%, 80%) ^a
4, 13, 22, 31, 40	<i>ArPerIndex</i> (20%, 35%, 50%, 65%, 80%) ^c
5, 14, 23, 32, 41	I_X (20%, 35%, 50%, 65%, 80%) ^d
6, 15, 24, 33, 42	I_Y (20%, 35%, 50%, 65%, 80%) ^d
7, 16, 25, 34, 43	I_X/I_Y (20%, 35%, 50%, 65%, 80%) ^c
8, 17, 26, 35, 44	I_{MIN} (20%, 35%, 50%, 65%, 80%) ^d
9, 18, 27, 36, 45	I_{MAX} (20%, 35%, 50%, 65%, 80%) ^d
10, 19, 28, 37, 46	I_{MAX}/I_{MIN} (20%, 35%, 50%, 65%, 80%) ^c

^a Measured in mm

^b Measured in mm²

^c Unit-less

^d Measured in mm⁴

summarized in Table 1, has been adopted for further statistical processing and development of appropriate sex classifiers for the femur, tibia, and humerus bones. Each side was separately evaluated for maximizing the efficiency of the classifiers resulting to six distinct datasets, which were evaluated sequentially with the following statistical workflow.

Initially, group box-plots with the typical whisker's size of 1.5 times the interquartile range were used to identify potential outliers, which were case-wise removed in each variable. Subsequently, each variable was tested with Student's t test for sexual dimorphism and its expression was evaluated with Cohen's d effect size. Welch's t test and non-parametric statistics were used accordingly. Prior to the multivariate analyses, all variables were normalized across the entire sample (pooled sexes) and their respective z -scores were utilized in subsequent statistical processing. The variables exhibiting sexual dimorphism were incorporated in step-wise discriminant function analysis (DFA) to identify the subset of variables that produce the optimal classification results. Since step-wise DFA evaluates only the F -statistics for the selection of the optimal variables, a trial-and-error repetitive procedure was applied by further considering the amount of multicollinearity (as expressed by the variable inflation factor, VIF) between the optimal variables, their original expression of sexual dimorphism (effect size), and the resulting leave-one-out cross-validated (LOOCV) classification. During DFA analysis, certain samples were further removed from the dataset to meet the necessary assumption of multivariate normality. The three most optimal variable configurations for each side and each bone were chosen for further analysis with supervised machine learning algorithms and the development of SVM models for sex estimation. Although machine learning does not require the strict statistical assumptions of DFA, the same datasets (removed outliers) were utilized to facilitate comparisons between the resulted classification scores.

Supervised machine learning was performed for all 18 optimal variable configurations produced by step-wise DFA. SVM models with a radial basis function (RBF) kernel ($C = 1$, $\gamma = 1 / \text{number of variables}$) were applied on the already normalized datasets with a K -fold (20 iterations) stratified random sample splitting (70% training, 30% testing). Cross-validated classification accuracy was calculated as the mean classification percentage over the 20 iterations and a 95% confidence interval being two standard deviations of the distributed classification scores. To further evaluate the accuracy of the proposed configurations of variables in the trained models, the entire sample was split into two equally-sized groups, conveniently named "old_yob" and "new_yob", according to the samples' year of birth and the same machine learning protocol was applied on each group while the other was used as a leave-out cross-validation group. The resulting chronological division of the entire sample was from 1879

until 1919 for the "old_yob" group and from 1920 until 1965 for the "new_yob" group. Machine learning was implemented in Python 3 using the scikit-learn 0.21.3 library [40], univariate sexual dimorphism was evaluated with JASP [41], multivariate DFA was done with SPSS (IBM SPSS version 23, Armonk NY), and general data handling, outlier removal, calculating SVM posterior probabilities, and result aggregation were performed in GNU Octave [37].

The final production-ready classifiers were trained on the entire sample and original group testing was utilized to produce a discrete probability density function (PDF) based on a sliding threshold range of the SVM's decision function results. These classifiers are intended for estimating sex in cases of unidentified individuals and the discrete PDF can be used as a lookup table for providing an estimate of posterior probability of the predicted outcome. Both SVM and DFA classifiers have been integrated in a standalone GNU Octave function, named "estimate_sex", which allows the user to apply them on new datasets for predicting sex through a self-intuitive graphical user interface. The function is provided as open-source software (GPLv3) and it is freely available at <https://github.com/pr0m1th3as/skeletal-sex-estimation>.

Results

The univariate statistics regarding all three bones revealed that most variables exhibit statistically significant sexual dimorphism. These results are summarized for the left and right side bones in Tables 2 and 3, respectively. The only variables that do not consistently yield statistically significant sexual dimorphism are the I_X/I_Y and I_{MAX}/I_{MIN} ratios. This is evident in both sides and across all cross sections and bones. Additionally, statistically non-significant results were found for the *ArPerIndex* variable at 35% and 65% cross sections of the right femur and the *ArPerIndex* 80% of the right tibia and humerus. More interesting results were revealed by the effect size estimates, although it should be noted that the effect size values between parametric (Cohen's d) and non-parametric (rank biserial correlation) tests are not directly comparable. As expected, the maximum length (*Max Distance*) of each bone was consistently found among the most sexually dimorphic traits along with the *Area* and *Perimeter* variables at every cross section (20%, 35%, 50%, 65%, 80%) with the latter most often exhibiting even more pronounced sexual dimorphism than the *Max Distance* variable. This is most evident in the cross sections near the mid shaft (35%, 50%, 65%) and especially in the humerus. Moderate to large effect sizes of sexual dimorphism were also observed for the I_X , I_Y and I_{MAX} , I_{MIN} variables demonstrating considerable differences in the distribution of the cross-sectional area along the anatomical axis as well as the cross-sectional longitudinal axis, respectively, between males and

Table 2 Univariate sexual dimorphism on left side femur, tibia, and humerus bones

Variables	Femur			Tibia			Humerus		
	Statistic	Sig. ^a	Effect Size ^b	Statistic	Sig. ^a	Effect Size ^b	Statistic	Sig. ^a	Effect size ^b
<i>Max Distance</i>	13.281 ^c	< .001	1.932	12.896 ^c	< .001	1.824	8988 ^{c, M}	< .001	0.870
<i>Area 20%</i>	8022 ^{e, M}	< .001	0.760	12.660 ^c	< .001	1.777	15.513 ^d	< .001	2.198
<i>Perimeter 20%</i>	7823 ^{e, M}	< .001	0.720	14.474 ^c	< .001	2.045	15.035 ^c	< .001	2.146
<i>ArPerIndex 20%</i>	3472 ^{e, M:F}	0.002	-0.254	-3.165 ^c	0.002	-0.451	5707 ^{e, M}	0.002	0.199
<i>I_X 20%</i>	7380 ^{e, M}	< .001	0.736	9161 ^{e, F}	< .001	0.805	9024 ^{e, M:F}	< .001	0.874
<i>I_Y 20%</i>	7493 ^{e, M:F}	< .001	0.753	13.743 ^d	< .001	1.929	14.940 ^d	< .001	2.096
<i>I_X/I_Y 20%</i>	4236 ^{e, F}	0.282	-0.090	-0.664 ^c	0.508	-0.094	6499 ^{e, F}	0.508	0.325
<i>I_{MIN} 20%</i>	7573 ^{e, M:F}	< .001	0.747	8896 ^{e, F}	< .001	0.769	9018 ^{e, F}	< .001	0.873
<i>I_{MAX} 20%</i>	7575 ^{e, M}	< .001	0.755	14.043 ^d	< .001	1.947	15.247 ^d	< .001	2.134
<i>I_{MAX}/I_{MIN} 20%</i>	4864 ^{e, M}	0.591	0.045	1.236 ^c	0.218	0.175	3679 ^{e, M:F}	0.218	-0.244
<i>Area 35%</i>	13.243 ^c	< .001	1.922	15.304 ^c	< .001	2.159	15.094 ^d	< .001	2.130
<i>Perimeter 35%</i>	13.846 ^c	< .001	2.022	9473 ^{e, F}	< .001	0.899	9148 ^{e, M}	< .001	0.828
<i>ArPerIndex 35%</i>	5324 ^{e, M:F}	0.026	0.188	-5.426 ^c	< .001	-0.768	5774 ^{e, M:F}	< .001	0.210
<i>I_X 35%</i>	8329 ^{e, F}	< .001	0.789	9297 ^{e, F}	< .001	0.829	8725 ^{e, M}	< .001	0.929
<i>I_Y 35%</i>	7924 ^{e, F}	< .001	0.738	16.019 ^d	< .001	2.229	8640 ^{e, M:F}	< .001	0.798
<i>I_X/I_Y 35%</i>	5331 ^{e, M}	0.023	0.192	-0.433 ^c	0.666	-0.061	7590 ^{e, F}	0.666	0.516
<i>I_{MIN} 35%</i>	13.318 ^d	< .001	1.900	9092 ^{e, F}	< .001	0.808	16.080 ^d	< .001	2.249
<i>I_{MAX} 35%</i>	12.293 ^d	< .001	1.759	9740 ^{e, M:F}	< .001	0.881	8854 ^{e, M:F}	< .001	0.863
<i>I_{MAX}/I_{MIN} 35%</i>	3228 ^{e, M:F}	< .001	-0.278	5910 ^{e, M}	0.046	0.163	-5.202 ^d	0.046	-0.756
<i>Area 50%</i>	14.636 ^c	< .001	2.129	9413 ^{e, F}	< .001	0.852	9335 ^{e, M}	< .001	0.865
<i>Perimeter 50%</i>	12.880 ^c	< .001	1.869	9618 ^{e, F}	< .001	0.892	9206 ^{e, M}	< .001	0.839
<i>ArPerIndex 50%</i>	5204 ^{e, M}	0.039	0.175	-8.434 ^c	< .001	-1.193	7009 ^{e, M:F}	< .001	0.400
<i>I_X 50%</i>	14.906 ^d	< .001	2.101	8972 ^{e, F}	< .001	0.838	9046 ^{e, M:F}	< .001	0.886
<i>I_Y 50%</i>	7967 ^{e, F}	< .001	0.802	9300 ^{e, F}	< .001	0.889	8890 ^{e, F}	< .001	0.864
<i>I_X/I_Y 50%</i>	5957 ^{e, F}	< .001	0.280	1.395 ^c	0.165	0.198	-2.324 ^d	0.165	-0.337
<i>I_{MIN} 50%</i>	14.982 ^d	< .001	2.146	9022 ^{e, F}	< .001	0.794	9091 ^{e, M}	< .001	0.935
<i>I_{MAX} 50%</i>	8304 ^{e, F}	< .001	0.822	9398 ^{e, F}	< .001	0.884	8790 ^{e, F}	< .001	0.863
<i>I_{MAX}/I_{MIN} 50%</i>	4405 ^{e, F:M}	0.880	-0.013	6411 ^{e, M}	< .001	0.290	2520 ^{e, M}	< .001	-0.461
<i>Area 65%</i>	13.355 ^c	< .001	1.940	9048 ^{e, F}	< .001	0.833	19.152 ^d	< .001	2.692
<i>Perimeter 65%</i>	13.028 ^c	< .001	1.897	9285 ^{e, F}	< .001	0.862	17.940 ^c	< .001	2.571
<i>ArPerIndex 65%</i>	5130 ^{e, M}	0.015	0.207	2763 ^{e, M}	< .001	-0.451	4.561 ^c	< .001	0.655
<i>I_X 65%</i>	8207 ^{e, F}	< .001	0.835	8732 ^{e, F}	< .001	0.806	9260 ^{e, M}	< .001	0.953
<i>I_Y 65%</i>	7896 ^{e, F}	< .001	0.780	9067 ^{e, F}	< .001	0.840	18.648 ^d	< .001	2.585
<i>I_X/I_Y 65%</i>	2.399 ^c	0.017	0.352	5390 ^{e, M}	0.614	0.041	-3.485 ^c	0.614	-0.496
<i>I_{MIN} 65%</i>	8148 ^{e, F}	< .001	0.865	8657 ^{e, F}	< .001	0.754	9464 ^{e, M}	< .001	0.951
<i>I_{MAX} 65%</i>	12.306 ^d	< .001	1.758	8949 ^{e, F}	< .001	0.868	9253 ^{e, M}	< .001	0.951
<i>I_{MAX}/I_{MIN} 65%</i>	4313 ^{e, M:F}	0.523	-0.054	1.456 ^c	0.147	0.207	-3.439 ^c	0.147	-0.494
<i>Area 80%</i>	11.071 ^c	< .001	1.609	10.379 ^c	< .001	1.466	17.595 ^d	< .001	2.461
<i>Perimeter 80%</i>	9.779 ^c	< .001	1.424	8717 ^{e, M}	< .001	0.684	15.458 ^d	< .001	2.176
<i>ArPerIndex 80%</i>	7.789 ^c	< .001	1.145	-0.571 ^c	0.568	-0.081	5089 ^{e, F}	0.568	0.077
<i>I_X 80%</i>	13.360 ^d	< .001	1.934	8603 ^{e, M:F}	< .001	0.662	16.586 ^d	< .001	2.333
<i>I_Y 80%</i>	7149 ^{e, F}	< .001	0.580	8259 ^{e, F}	< .001	0.690	14.878 ^d	< .001	2.052
<i>I_X/I_Y 80%</i>	6565 ^{e, M}	< .001	0.461	5021 ^{e, F}	0.711	-0.030	4184 ^{e, M}	0.711	-0.147
<i>I_{MIN} 80%</i>	8126 ^{e, M}	< .001	0.856	8332 ^{e, F}	< .001	0.655	16.047 ^d	< .001	2.244
<i>I_{MAX} 80%</i>	7170 ^{e, F}	< .001	0.585	8237 ^{e, F}	< .001	0.702	14.601 ^d	< .001	2.022
<i>I_{MAX}/I_{MIN} 80%</i>	-5.807 ^c	< .001	-0.844	0.636 ^c	0.525	0.090	2.533 ^c	0.525	0.361

^a *p* value for corresponding statistic

^b Calculated either with Cohen's *d* (Student's *t* test; Welch's *t* test) or with rank biserial correlation (Mann-Whitney)

^c Statistic of Student's *t* test

^d Statistic of Welch's *t* test

^e Statistic of Mann-Whitney test, the following superscripts (M;F) indicate the groups without normality

females. Another notable finding concerns the *ArPerIndex* variables in cases of statistically significant outcomes. Despite exhibiting small to moderate positive effect size for the femur and humerus, they produced moderate to large negative effect size for the tibia. On the contrary, the

corresponding statistically significant *I_{MAX}/I_{MIN}* ratios followed an inverse pattern yielding mostly negative effect size for femur and humerus while positive effect size for tibia.

The multivariate step-wise DFA produced cross-validated classification results ranging from 88.8 up to 97.3%. All DFA

Table 3 Univariate sexual dimorphism on right side femur, tibia, and humerus bones

Variables	Femur			Tibia			Humerus		
	Statistic	Sig. ^a	Effect Size ^b	Statistic	Sig. ^a	Effect Size ^b	Statistic	Sig. ^a	Effect Size ^b
<i>Max Distance</i>	12.195 ^c	< .001	1.764	15.627 ^d	< .001	2.191	13.758 ^c	< .001	1.969
<i>Area 20%</i>	8148 ^{e, F}	< .001	0.784	15.905 ^c	< .001	2.247	16.005 ^d	< .001	2.248
<i>Perimeter 20%</i>	7954 ^{e, M;F}	< .001	0.725	16.909 ^c	< .001	2.384	15.129 ^c	< .001	2.158
<i>ArPerIndex 20%</i>	3185 ^{e, M;F}	< .001	-0.300	-3.657 ^c	< .001	-0.514	6011 ^{e, M}	< .001	0.278
<i>I_X 20%</i>	7822 ^{e, M;F}	< .001	0.787	9675 ^{e, F}	< .001	0.855	9063 ^{e, F}	< .001	0.882
<i>I_Y 20%</i>	7589 ^{e, M;F}	< .001	0.815	14.190 ^d	< .001	1.963	15.457 ^d	< .001	2.143
<i>I_X/I_Y 20%</i>	-0.624 ^c	0.533	-0.090	5630 ^{e, F}	0.545	0.049	2.697 ^c	0.008	0.385
<i>I_{MIN} 20%</i>	7748 ^{e, F}	< .001	0.809	9322 ^{e, F}	< .001	0.784	9022 ^{e, F}	< .001	0.873
<i>I_{MAX} 20%</i>	7546 ^{e, M;F}	< .001	0.800	15.666 ^d	< .001	2.169	9172 ^{e, F}	< .001	0.883
<i>I_{MAX}/I_{MIN} 20%</i>	4249 ^{e, M;F}	0.541	-0.052	1.597 ^c	0.112	0.222	4699 ^{e, M}	0.525	-0.052
<i>Area 35%</i>	13.837 ^c	< .001	2.006	16.490 ^c	< .001	2.311	9076 ^{e, F}	< .001	0.863
<i>Perimeter 35%</i>	13.815 ^c	< .001	2.005	10098 ^{e, F}	< .001	0.915	14.938 ^c	< .001	2.131
<i>ArPerIndex 35%</i>	4834 ^{e, M;F}	0.229	0.102	-5.820 ^c	< .001	-0.816	3.462 ^d	< .001	0.507
<i>I_X 35%</i>	12.142 ^d	< .001	1.716	9580 ^{e, F}	< .001	0.867	8902 ^{e, F}	< .001	0.909
<i>I_Y 35%</i>	8222 ^{e, F}	< .001	0.780	16.420 ^d	< .001	2.271	13.466 ^d	< .001	1.877
<i>I_X/I_Y 35%</i>	0.035 ^c	0.972	0.005	5199 ^{e, F}	0.930	0.007	6.868 ^c	< .001	0.993
<i>I_{MIN} 35%</i>	8095 ^{e, F}	< .001	0.807	9535 ^{e, F}	< .001	0.828	9026 ^{e, F}	< .001	0.896
<i>I_{MAX} 35%</i>	13.171 ^d	< .001	1.868	10223 ^{e, M;F}	< .001	0.905	14.754 ^d	< .001	2.057
<i>I_{MAX}/I_{MIN} 35%</i>	3757 ^{e, M;F}	0.033	-0.179	2.030 ^c	0.044	0.286	2697 ^{e, M}	< .001	-0.440
<i>Area 50%</i>	15.350 ^d	< .001	2.194	9812 ^{e, F}	< .001	0.861	16.518 ^d	< .001	2.309
<i>Perimeter 50%</i>	13.788 ^c	< .001	1.997	10037 ^{e, F}	< .001	0.904	9247 ^{e, M}	< .001	0.833
<i>ArPerIndex 50%</i>	5286 ^{e, M}	0.035	0.178	-8.777 ^c	< .001	-1.222	7024 ^{e, F}	< .001	0.481
<i>I_X 50%</i>	8404 ^{e, F}	< .001	0.840	9099 ^{e, F}	< .001	0.890	9051 ^{e, F}	< .001	0.913
<i>I_Y 50%</i>	8261 ^{e, F}	< .001	0.809	9457 ^{e, F}	< .001	0.890	9073 ^{e, F}	< .001	0.846
<i>I_X/I_Y 50%</i>	2.451 ^d	0.015	0.351	5574 ^{e, F}	0.634	0.038	-1.228 ^d	0.221	-0.178
<i>I_{MIN} 50%</i>	8397 ^{e, F}	< .001	0.838	9459 ^{e, F}	< .001	0.813	9082 ^{e, M}	< .001	0.920
<i>I_{MAX} 50%</i>	8383 ^{e, F}	< .001	0.857	9615 ^{e, M;F}	< .001	0.894	8984 ^{e, F}	< .001	0.871
<i>I_{MAX}/I_{MIN} 50%</i>	5318 ^{e, M;F}	0.047	0.167	2.874 ^c	0.004	0.402	2899 ^{e, F}	< .001	-0.410
<i>Area 65%</i>	13.278 ^c	< .001	1.919	13.701 ^c	< .001	1.919	19.495 ^d	< .001	2.710
<i>Perimeter 65%</i>	13.384 ^c	< .001	1.939	9637 ^{e, F}	< .001	0.844	18.204 ^c	< .001	2.600
<i>ArPerIndex 65%</i>	4964 ^{e, M}	0.102	0.139	-6.117 ^c	< .001	-0.852	5.293 ^c	< .001	0.760
<i>I_X 65%</i>	8407 ^{e, F}	< .001	0.841	9214 ^{e, F}	< .001	0.796	19.014 ^d	< .001	2.591
<i>I_Y 65%</i>	8231 ^{e, F}	< .001	0.802	9273 ^{e, F}	< .001	0.833	18.267 ^d	< .001	2.491
<i>I_X/I_Y 65%</i>	1.811 ^c	0.072	0.262	5477 ^{e, F}	0.632	0.039	-3.391 ^c	< .001	-0.485
<i>I_{MIN} 65%</i>	8241 ^{e, F}	< .001	0.843	9238 ^{e, F}	< .001	0.790	18.841 ^d	< .001	2.573
<i>I_{MAX} 65%</i>	13.353 ^d	< .001	1.894	9500 ^{e, F}	< .001	0.857	18.271 ^d	< .001	2.490
<i>I_{MAX}/I_{MIN} 65%</i>	4607 ^{e, F}	0.721	0.030	1.391 ^c	0.166	0.196	3341 ^{e, M}	< .001	-0.306
<i>Area 80%</i>	10.850 ^d	< .001	1.555	11.141 ^d	< .001	1.587	18.177 ^d	< .001	2.508
<i>Perimeter 80%</i>	9.479 ^c	< .001	1.370	11.072 ^d	< .001	1.575	15.364 ^d	< .001	2.145
<i>ArPerIndex 80%</i>	6.900 ^c	< .001	1.004	0.177 ^d	0.860	0.025	5229 ^{e, M}	0.286	0.089
<i>I_X 80%</i>	12.970 ^d	< .001	1.853	8928 ^{e, F}	< .001	0.709	17.501 ^d	< .001	2.389
<i>I_Y 80%</i>	7376 ^{e, M;F}	< .001	0.597	10.136 ^c	< .001	1.432	9136 ^{e, M}	< .001	0.869
<i>I_X/I_Y 80%</i>	6843 ^{e, M}	< .001	0.484	1.494 ^c	0.137	0.210	-1.143 ^c	0.254	-0.162
<i>I_{MIN} 80%</i>	8150 ^{e, F}	< .001	0.840	10.829 ^c	< .001	1.517	17.154 ^d	< .001	2.350
<i>I_{MAX} 80%</i>	7376 ^{e, M;F}	< .001	0.597	8525 ^{e, F}	< .001	0.693	9103 ^{e, M}	< .001	0.885
<i>I_{MAX}/I_{MIN} 80%</i>	-6.405 ^d	< .001	-0.934	-1.200 ^c	0.232	-0.168	2.266 ^c	0.024	0.321

^a *p* value for corresponding statistic

^b Calculated either with Cohen's *d* (Student's *t* test; Welch's *t* test) or with rank biserial correlation (Mann-Whitney)

^c Sf Student's *t* test

^d Statistic of Welch's *t* test

^e Statistic of Mann-Whitney test, the following superscripts (M;F) indicate the groups without normality

results are summarized in Tables 4 and 5 for the left and right sides, respectively. More specifically, the most sexually dimorphic long bone is the humerus, whose DFs yielded LOOCV classification between 95.2 and 97.3% with the left side performing slightly better. The femur produced LOOCV

classification between 91.1% and 94.8% with the left side performing once again somewhat better. Surprisingly, the tibia displayed the most bilateral discrepancy in classification accuracy with the right side ranging from 94 to 94.7% and even outperforming most femur DFs, whereas the left side

Table 4 Optimal multivariate discriminant functions for left side femur, tibia, and humerus bones

Function	Sample size	Wilks' lambda	Sig.	Used variables	Correctly classified	
					Original	Cross-validated
Femur #1	174	0.287	< 0.001	1. <i>Max Distance</i> 3. <i>Area</i> 50% 5. <i>ArPerIndex</i> 80%	2. <i>ArPerIndex</i> 20% 4. <i>Perimeter</i> 50% 6. I_X/I_Y 80%	95.4% 94.8%
Femur #2	177	0.320	< 0.001	1. <i>Max Distance</i> 3. <i>Area</i> 50% 5. I_X/I_Y 80%	2. <i>ArPerIndex</i> 20% 4. <i>ArPerIndex</i> 80%	93.2% 92.1%
Femur #3	180	0.319	< 0.001	1. <i>Max Distance</i> 3. I_{MAX} 50%	2. <i>Area</i> 50% 4. <i>ArPerIndex</i> 80%	92.2% 91.1%
Tibia #1	188	0.345	< 0.001	1. <i>Max Distance</i> 3. <i>ArPerIndex</i> 50% 5. <i>Perimeter</i> 65%	2. I_{MAX} 35% 4. I_Y 50%	90.4% 90.4%
Tibia #2	188	0.345	< 0.001	1. <i>Max Distance</i> 3. <i>ArPerIndex</i> 50% 5. <i>Area</i> 65%	2. I_{MAX} 35% 4. I_Y 50%	89.9% 89.9%
Tibia #3	188	0.346	< 0.001	1. <i>Max Distance</i> 3. <i>ArPerIndex</i> 50% 5. <i>Perimeter</i> 65%	2. I_Y 35% 4. I_{MAX} 50%	89.9% 88.8%
Humerus #1	187	0.224	< 0.001	1. <i>Max Distance</i> 3. I_X/I_Y 35% 5. <i>Area</i> 65%	2. I_X/I_Y 20% 4. I_X 50% 6. I_{MIN} 80%	97.9% 97.3%
Humerus #2	189	0.241	< 0.001	1. <i>Max Distance</i> 3. I_X/I_Y 35% 5. I_{MAX} 65%	2. I_X/I_Y 20% 4. <i>Area</i> 65%	97.4% 96.8%
Humerus #3	189	0.275	< 0.001	1. <i>Max Distance</i> 3. I_X/I_Y 35% 5. I_{MIN} 80%	2. I_X/I_Y 20% 4. <i>Area</i> 65%	96.8% 96.8%

Table 5 Optimal multivariate discriminant functions for right side femur, tibia, and humerus bones

Function	Sample size	Wilks' lambda	Sig.	Used variables	Correctly classified	
					Original	Cross-validated
Femur #1	176	0.300	< 0.001	1. <i>Max Distance</i> 3. <i>Area</i> 50% 5. <i>ArPerIndex</i> 80%	2. <i>ArPerIndex</i> 20% 4. I_{MIN} 65% 6. I_{MIN} 80%	94.3% 92.6%
Femur #2	178	0.303	< 0.001	1. <i>Max Distance</i> 3. <i>Area</i> 50% 5. I_{MIN} 80%	2. <i>ArPerIndex</i> 20% 4. <i>ArPerIndex</i> 80%	93.8% 92.1%
Femur #3	184	0.342	< 0.001	1. <i>Max Distance</i> 3. <i>Area</i> 35%	2. <i>ArPerIndex</i> 20% 4. <i>ArPerIndex</i> 80%	91.8% 91.8%
Tibia #1	187	0.299	< 0.001	1. <i>Max Distance</i> 3. <i>Perimeter</i> 35%	2. <i>ArPerIndex</i> 20% 4. I_Y 35%	94.7% 94.7%
Tibia #2	184	0.297	< 0.001	1. <i>Max Distance</i> 3. <i>Perimeter</i> 35% 5. I_{MAX} 50%	2. <i>ArPerIndex</i> 20% 4. I_Y 35%	94.0% 94.0%
Tibia #3	182	0.301	< 0.001	1. <i>Max Distance</i> 3. <i>Perimeter</i> 35%	2. <i>ArPerIndex</i> 20% 4. I_X 50%	94.0% 93.4%
Humerus #1	183	0.247	< 0.001	1. <i>Max Distance</i> 3. <i>Area</i> 65% 5. <i>ArPerIndex</i> 65%	2. I_X/I_Y 35% 4. <i>Perimeter</i> 65% 6. <i>Perimeter</i> 80%	96.2% 96.2%
Humerus #2	184	0.274	< 0.001	1. <i>Max Distance</i> 3. <i>Perimeter</i> 65% 5. <i>Perimeter</i> 80%	2. I_X 35% 4. <i>ArPerIndex</i> 65%	96.2% 96.2%
Humerus #3	186	0.304	< 0.001	1. <i>Max Distance</i> 3. I_X/I_Y 35%	2. I_X/I_Y 20% 4. <i>Area</i> 65%	95.2% 95.2%

produced the poorest results in this study with LOOCV classification between 88.8 and 90.4%. Another finding that further strengthens the split-side design of this study regards the differences in the optimal variables between left and right side bones. Despite certain variables being used consistently in left-side and right-side DFs, some of the used variables in each DF, shown in Tables 4 and 5, differ from one side to another. It should be noted that during the multivariate analysis, the optimal list of variables of each DF has been applied to the complementary side by enter-method DFA and found sub-optimal to those reported here. What has shown to be fairly consistent between sides, though, is the portion of the diaphysis that the DFs of each bone focus upon. More specifically, there is an observed trend in the femoral DFs focusing mainly on the proximal and distal ends of the diaphysis (20% and 80% cross sections), whereas the DFs of the tibia and the humerus utilize more variables near the midshaft (35%, 50%, and 65% cross sections). Additionally to that, the DFs of the right tibia shift further to the proximal part rather than these of the left side. Another observation of the multivariate results worth mentioning is that *ArPerIndex* variables are more frequently used in the optimal DFs than any other variable with the sole exception of the *Max Distance*, despite exhibiting much lower univariate effect size than other variables with more prominent sexual dimorphism. Moreover, it also became apparent from the resulting

VIFs during the statistical analysis that substituting *Area* or *Perimeter* with *ArPerIndex* from the same cross section not only increased the performance but also reduced the multi-collinearity among the optimal variables resulting in more robust estimators.

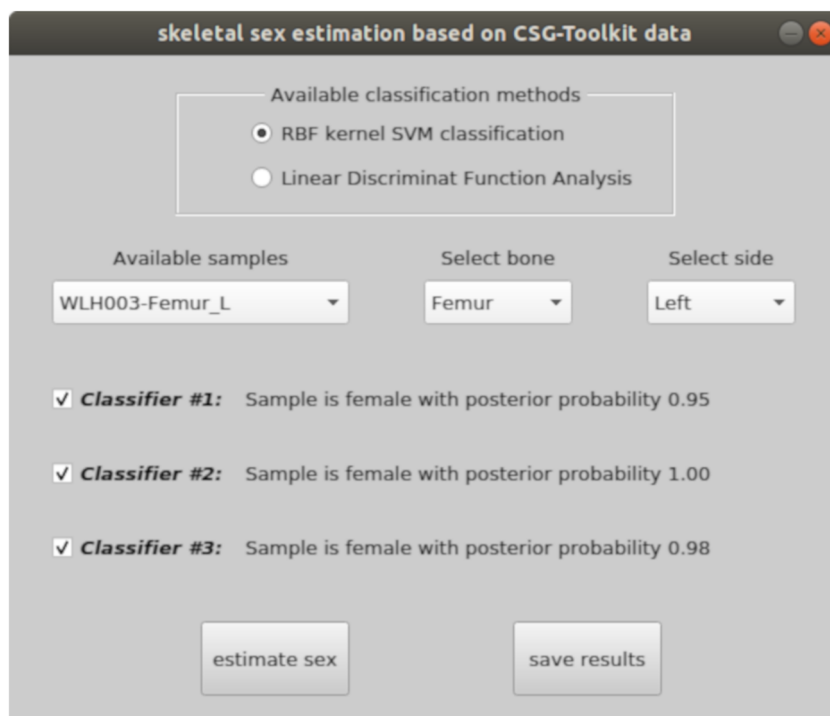
The machine learning results from applying the same optimal sets of variables on the entire sample as well as on the two sample splits based on chronological year of birth are summarized in Table 6. The cross-validated (CV) classification results from the entire sample follow the same pattern with the DFA results although being slightly lower. Once again, the most sex diagnostic bone was the right humerus while the left tibia yielded the lowest performance in sex discrimination. The original group classification results, which are mostly similar to their DFA counterparts, are also included in Table 6 for cross-referencing. Regarding the results of the chronologically split groups, the CV classification scores are mostly lower with few exceptions. Additionally, the CV scores are accompanied by fairly larger confidence interval values due to the smaller sample size of each respective training/testing group. More interestingly, the original group classification scores for both *old_job* and *new_job* groups are consistently higher than their correspondences in the entire sample training. However, when testing the trained models from the chronologically split samples on their complements, the classification accuracy falls with the only notable

Table 6 Aggregated results from SVM classifiers

Functions	Full sample			Training on old job			Testing new job	Training on new job			Testing old job
	OR	CV	CI	OR	CV	CI		OR	CV	CI	
Left femur #1	0.931	0.908	0.071	0.948	0.893	0.111	0.909	0.909	0.898	0.116	0.935
Left femur #2	0.932	0.911	0.063	0.936	0.893	0.131	0.949	0.937	0.913	0.089	0.923
Left femur #3	0.911	0.917	0.071	0.911	0.880	0.120	0.888	0.913	0.908	0.111	0.911
Right femur #1	0.949	0.925	0.066	0.961	0.935	0.095	0.922	0.974	0.948	0.074	0.909
Right Femur #2	0.949	0.926	0.065	0.962	0.943	0.073	0.910	0.962	0.950	0.063	0.897
Right femur #3	0.924	0.921	0.070	0.926	0.914	0.098	0.926	0.938	0.905	0.095	0.914
Left tibia #1	0.915	0.902	0.070	0.902	0.879	0.115	0.890	0.927	0.883	0.111	0.890
Left tibia #2	0.915	0.903	0.072	0.902	0.879	0.115	0.890	0.890	0.881	0.111	0.915
Left tibia #3	0.899	0.871	0.078	0.915	0.893	0.108	0.866	0.939	0.902	0.118	0.854
Right tibia #1	0.947	0.935	0.074	0.963	0.917	0.116	0.927	0.927	0.902	0.136	0.951
Right tibia #2	0.946	0.937	0.055	0.950	0.930	0.166	0.900	0.950	0.925	0.087	0.950
Right tibia #3	0.945	0.913	0.089	0.900	0.883	0.135	0.938	0.963	0.923	0.128	0.925
Left humerus #1	0.984	0.968	0.037	0.975	0.960	0.087	0.988	0.988	0.960	0.075	0.938
Left humerus #2	0.974	0.963	0.045	0.976	0.943	0.083	0.976	0.976	0.940	0.066	0.963
Left humerus #3	0.979	0.966	0.048	0.976	0.967	0.080	0.988	0.988	0.974	0.077	0.963
Right humerus #1	0.962	0.960	0.050	0.975	0.975	0.067	0.975	0.975	0.950	0.084	0.962
Right humerus #2	0.962	0.949	0.057	0.938	0.925	0.128	0.963	0.975	0.965	0.071	0.950
Right humerus #3	0.962	0.940	0.069	0.950	0.925	0.132	0.963	0.988	0.973	0.080	0.950

OR original group classification, CV cross-validation classification, CI 95% confidence interval of CV

Fig. 1 The graphical user interface of the “estimate_sex” function



exception being the humerus. The trained humerus models from the *old_yob* perform even better in the *new_yob* group. Nevertheless, the opposite is not true, since all trained humerus models from the *new_yob* did not perform as good on the *old_yob* group.

All DFs reported in Tables 4 and 5 as well as their SVM counterparts trained on the entire sample have been assembled in a GNU Octave data text container, named “long_bone_classifier.mat”. This can be loaded by the “estimate_sex” function, whose graphical user interface (GUI) is shown in Fig. 1, to facilitate sex estimation on new femur, tibia, or humerus samples. The only requirement is that any new testing datasets must be created with the CSG-Toolkit, whose latest version (v.1.1.0) has an additional function, called “inspect_CSG”, which will aggregate any processed 3D models into the appropriate tabular format and save it as a csv file. As shown in Fig. 1, the GUI allows the selection of bone, side, and classification method for any available sample and the function will utilize any of the selected classifiers to make a prediction along with its posterior probability. Additionally, the selected results can be saved in a csv file and any further requests will be appended into the same results file. The “estimate_sex” function can be invoked from within the GNU Octave environment. A testing dataset (“sampledata.csv”) containing all relevant measurements from three long bones of one male and one female individual from the Athens collection is also bundled with the function for demonstrating purposes.

Discussion

One primary goal of the present work has been to explore the sex diagnostic utility of metric variables beyond the standard osteometric measurements, since the latter are mostly representative of the size of the bone rather than its shape. To this end, both CSG properties (I_X , I_Y , I_{MIN} , I_{MAX}) and custom ratios (I_X/I_Y , I_{MAX}/I_{MIN} , $ArPerIndex$) have been evaluated with the former representing general measures of robusticity and the latter representing particular shape attributes. The rationale of this approach was that despite the well-established size-related sexual dimorphism, shape differences are also apparent between male and female long bones [42]. In this aspect, CSG properties can be regarded as shape variables portraying the area distribution along the anatomical and principal axes of the cross section, whereas $ArPerIndex$ is a shape measure of its exterior boundary. As expected, male long bones are larger than their female counterparts, as shown by *Max Distance*, *Area*, and *Perimeter* variables. Beyond this well-established pattern of sexual dimorphism, though, the CSG properties also exhibit substantial levels of sexual dimorphism. As shown from the respective results (Tables 2 and 3), the variables representing the 2nd moments of cross-sectional area exhibit comparable levels of sexual dimorphism with the size variables and even higher in certain cases, especially the I_X and I_{MAX} variables. Another important finding of the univariate analysis concerns the statistically significant occurrences of the $ArPerIndex$ and I_{MAX}/I_{MIN} variables. According to the results, the diaphyseal cross section of the tibia is less elongated

and more smooth in females rather than males. On the contrary, females have more elongated femoral and humeral diaphyseal cross sections as revealed by the negative effect size of the corresponding I_{MAX}/I_{MIN} variables (i.e., ratio is larger in females). Naturally, the higher the I_{MAX}/I_{MIN} ratio, the lower the *ArPerIndex* should be, since they are inversely proportional by definition. Nevertheless, the respective *ArPerIndex* variables in male femur and humerus are slightly larger than females, which indicates that the corresponding cross-sectional boundary is still smoother although less circular in females than males (i.e., the positive size effect in *ArPerIndex* would be much larger if male cross sections were smoother or if female cross sections were less elongated). This is mostly related to the *deltoid tuberosity* of the humerus and the *gluteal tuberosity* of the femur being more pronounced in males, which implies increased perimeter for a given area (i.e., lower *ArPerIndex*) and consequently counterbalances for the lower circularity (i.e., lower *ArPerIndex*) of the cross-sectional boundary observed in females.

Although the univariate results have already demonstrated the sex-diagnostic capacity of non-standard osteometric measurements in skeletal sex estimation, the multivariate statistics made their usefulness even more evident, since the non-standard *ArPerIndex* is the second most utilized variable in step-wise DFA and most of the optimal DFs (Tables 4 and 5) further include variables related to CSG properties. The advantages of the proposed method are further corroborated by the classification results, which confirm a higher sex-diagnostic capacity over earlier methods [e.g., 8, 9, 12, 13, 20, 22, 26, 28, 43]. For example, Charisi and her colleagues in 2011, who also studied the Athens Collection for sexing Greeks by standard osteometric measurements of the arm bones [10], reported a maximum cross-validated classification accuracy of 95.7% for the humerus whereas the present method yielded 97.3% for the same bone, albeit on a different side and with slightly larger sample size.

The present work utilized both DFA and SVM learning for constructing sex prediction models. Although this approach appears to induce some redundancy, it was adopted because both statistical methods have strengths and weaknesses that complement each other. DFA relies on certain statistical assumptions regarding the sample, which although restrictive, once they are met it provides a statistically robust way of identifying the optimal subset of variables through the step-wise method. Limiting the number of variables is important in order to avoid model over-fitting and considering the number of variables extracted by the CSG-Toolkit makes it evidently a necessary step. However, a small number of variables in a DF does not fully compensate for the limitations imposed by other statistical assumptions of DFA, which may result to over-fitting, hence sample specificity, regardless of the number of variables. This could possibly be an additional explanation for Kranioti's and Tzanakis' work [27], where they reported high

misclassification for certain DFs by Charisi and her colleagues [10] when applied to their Greek Cretan sample and they concluded that “most probably it represents a sampling effect rather than population differences between Cretans and mainland Greeks”. On the other hand, SVM is much more permissive on the statistical assumptions regarding training data [44]. This was instrumental for devising the much more rigid cross-validation strategy compared with DFA's LOOCV scheme since splitting the already limited skeletal sample into training/testing subsets did not pose any considerable limitations. Nevertheless, SVM remains prone to model over-fitting, hence the utilization of the subset of optimal variables identified by the step-wise DFA.

As shown from the full sample SVM classification results (Table 6), step-wise DFA and SVM produced fairly similar sex-diagnostic accuracy for all three bones under evaluation. Most importantly, the CV scores are very consistent with their corresponding original group scores while the respective confidence intervals are considerably low given the limited sample sizes. These results manifest that the trained SVM models are well-fitted into the population sample and considering the 20-fold stratified random sample splitting cross-validation strategy further implies good fitting for the represented population. This is also corroborated by the classification results based on the chronological sample splitting. Previous work on the Athens collection has shown that the Greek population experienced a moderate positive secular change in stature during the period 1879–1965, which is more prominent in females [45]. This dictates that the length of the lower limbs must have changed and more importantly at a different pace between males and females, which expectedly should affect the sex-diagnostic capacity of the trained models from one chronological group when applied to another. However, this is only partially evident in the classification results between the *old_yob* and the *new_yob* groups and vice versa, since they only show a modest tendency of classification accuracy loss between the chronological groups. It should be noted that the larger deviations in accuracy between original and cross-validated groups with respect to the full sample SVM results may be mostly attributed to the much smaller sub samples as a result of the chronological split. A possible explanation for the modest differentiation in classification accuracy is that all trained classifiers further rely on shape attributes of the respective bones and these shape attributes, such as *ArPerIndex*, are most likely unaffected by secular changes. Although this assertion cannot be verified with the existing dataset, it would imply that the utilized shape variables, directly or indirectly extracted from the diaphyseal CSG properties, not only produce more accurate sex classifiers, but also that these classifiers might be less population specific.

Another important goal of this study was the automatic evaluation of posterior probability for any sex estimate made on new unidentified cases. Despite posterior probability being

often underappreciated in skeletal sex estimation research, since most works explicitly focus on the achieved classification accuracy of their reported DFs and fail to provide sufficient parameters for its calculation, we argue that the posterior probability is even more important when it comes to utilizing a DF in forensic casework. Classification accuracy merely represents the sex discriminatory capacity of a set of measurements on a given population, but it does not provide any measure of validity over a DF's outcome on individual cases. Fortunately, the posterior probability of membership can be analytically calculated from the DF's outcome for each group when dealing with DFA due to its statistical properties. In general, the posterior probability of a sample y belonging to group j is $\pi_{jy} = Pr[X = y] = \frac{\exp[df_j(y)]}{\sum_{h=1}^k \exp[df_h(y)]}$, $j = 1, \dots, k$. [46].

In sex estimation, which is a binary classification problem [47] and thus the group DFs are complementary as additive inverse scores, the calculation of posterior probability is further simplified to $\frac{e^{LCF}}{e^{LCF} + e^{-LCF}}$, where LCF is the outcome of the DF, assuming the sectioning point is zero. It should be noted that this last assumption is not always true, especially when group sizes differ [48], and this is why group centroids should always be reported along with their respective DFs. Contrary to DFA, SVM learning does not provide the luxury of a statistical way of calculating posterior probabilities from the trained models, since RBF kernel-based SVMs make no assumptions about the distribution properties for the trained data. Although various complex techniques have been proposed to obtain posterior probabilities from SVM models [49–51], usually by cluster sampling and estimating the PDF of the inverse of the separating margin of SVM, these are limited to large datasets [50]. Hence, in the present work, we adopted a simplified strategy based on a sliding window over the range of the SVM's decision function scores for every available sample and their distance from the separating margin in order to estimate what we conveniently named a “discrete PDF”. The concept is that the further away a sample's score is from the separating margin, the more likely the sample belongs to the predicted group. Despite not being analytically sound from a statistical perspective, this approach provides a good approximation of posterior probability that can be used to assess the validity of a prediction in an unidentified case.

The present study has focused on improving the sex diagnostic utility of the femur, tibia, and humerus bones. This was achieved by incorporating recent technologies such as the CSG-Toolkit for exploring metric traits beyond the standard osteometric measurements; by blending statistical analyses to overcome particular sample constraints while obtaining optimal classification results; even by improvising on certain methods to estimate posterior probabilities. The most important aim, however, has been to produce a sex estimation method that can be easily applied on forensic or archaeological

context. This was the motivation for implementing all the necessary computations (data normalization, decision scores, posterior probabilities) into an application with a self-intuitive graphical user interface following the paradigm of relevant osteometric software like DSP2. Despite the merits of osteometric over visual observation methods that have been extensively discussed in the literature, the latter are still often favored in practice because of their ease of use. Acknowledging this practical advantage, we introduce this measurement-free osteometric method for estimating sex, implemented in the “estimate_sex” program, which solely relies on the GNU Octave environment and the availability of skeletal 3D models. Furthermore, it can be readily used in virtual anthropology, whether it concerns a modern virtopsy case or archaeological human remains such as bog bodies or mummies, which are usually examined with medical imaging techniques due to preservation.

As a final remark, all the necessary classification parameters contained in the *long_bone_classifier.mat* file have been kept separate from the “estimate_sex” program on purpose to allow the implemented software the flexibility of loading classification parameters from different population samples. The GNU Octave scripts used in the present study to calculate and store the classification parameters into the aforementioned file are available upon request to any researchers willing to produce a respective file for a different population sample. The authors would be glad to provide feedback for the minor changes that might be required to adapt to a different working sample.

Conclusions

The newly introduced sex estimation method fully eliminates observer error, while it facilitates accurate sex estimation from virtual 3D bone models. Our results show that blending standard osteometric measurements such as maximum length with custom variables based on the diaphyseal CSG properties such as the I_{MAX}/I_{MIN} ratio and the *ArPerIndex* further improves classification performance compared with existing methods. The maximum cross-validated classification accuracy in the Greek sample reached 94.8% for the femur, 94.7% for the tibia, and 97.3% for the humerus. Our results further suggest that the proposed method is reasonably robust to secular changes regarding its sex-diagnostic capacity.

Finally, the proposed method has been implemented in a graphical user interface program for the GNU Octave environment, which allows the user to select any of the available classifiers presented in this study in order to estimate sex from a femur, tibia, or humerus and further obtain the posterior probability of each classification result. Since each classifier utilizes a partially different set of variables and each classification method has different statistical properties, we advise

the user to utilize them all for estimating sex of any given unidentified sample and comprehensively take into consideration the reported posterior probabilities. Furthermore, it should be kept in mind that the proposed method has been developed on a modern Greek population sample and thus its classification outcomes may not be directly applicable to different populations.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Code availability All code of custom built software is licensed under the GNU GPL v3 open source license and is available at GitHub. See manuscript for more details.

References

- Bruzek J (2002) A method for visual determination of sex, using the human hip bone. *Am J Phys Anthropol* 117:157–168. <https://doi.org/10.1002/ajpa.10012>
- Phenice TW (1969) A newly developed visual method of sexing the os pubis. *Am J Phys Anthropol* 30:297–301. <https://doi.org/10.1002/ajpa.1330300214>
- Ubelaker DH, Volk CG (2002) A test of the phenice method for the estimation of sex. *J Forensic Sci* 47:15200J. <https://doi.org/10.1520/JFS15200J>
- Brůžek J, Santos F, Dutailly B, Murail P, Cunha E (2017) Validation and reliability of the sex estimation of the human os coxae using freely available DSP2 software for bioarchaeology and forensic anthropology. *Am J Phys Anthropol* 164:440–449. <https://doi.org/10.1002/ajpa.23282>
- Chapman T, Lefevre P, Semal P et al (2014) Sex determination using the Probabilistic Sex Diagnosis (DSP: Diagnose Sexuelle Probabiliste) tool in a virtual environment. *Forensic Sci Int* 234:189.e1–189.e8. <https://doi.org/10.1016/j.forsciint.2013.10.037>
- Murail P, Bruzek J, Houët F, Cunha E (2005) DSP: a tool for probabilistic sex diagnosis using worldwide variability in hip-bone measurements. *Bull Mém Société D'Anthropologie Paris* 17:167–176
- Quatrehomme G, Radoman I, Nogueira L, du Jardin P, Alumni V (2017) Sex determination using the DSP (probabilistic sex diagnosis) method on the coxal bone: efficiency of method according to number of available variables. *Forensic Sci Int* 272:190–193. <https://doi.org/10.1016/j.forsciint.2016.10.020>
- Spradley MK, Jantz RL (2011) Sex estimation in forensic anthropology: skull versus postcranial elements. *J Forensic Sci* 56:289–296. <https://doi.org/10.1111/j.1556-4029.2010.01635.x>
- Anzellini A, Toyne JM (2019) Estimating sex using isolated appendicular skeletal elements from Chachapoyas, Peru. *Int J Osteoarchaeol* 29:961–973. <https://doi.org/10.1002/oa.2810>
- Charisi D, Eliopoulos C, Vanna V, Koiliias CG, Manolis SK (2011) Sexual dimorphism of the arm bones in a modern Greek population. *J Forensic Sci* 56:10–18. <https://doi.org/10.1111/j.1556-4029.2010.01538.x>
- Duangto P, Mahakkanukrauh P (2020) Sex estimation from upper limb bones in a Thai population. *Anat Cell Biol* 53:36–43. <https://doi.org/10.5115/acb.19.179>
- Kranioti EF, Michalodimitrakis M (2009) Sexual dimorphism of the humerus in contemporary Cretans—a population-specific study and a review of the literature. *J Forensic Sci* 54:996–1000. <https://doi.org/10.1111/j.1556-4029.2009.01103.x>
- Moore MK, DiGangi EA, Niño Ruiz FP et al (2016) Metric sex estimation from the postcranial skeleton for the Colombian population. *Forensic Sci Int* 262:286.e1–286.e8. <https://doi.org/10.1016/j.forsciint.2016.02.018>
- Bertsatos A, Papageorgopoulou C, Valakos E, Chovalopoulou M-E (2018) Investigating the sex-related geometric variation of the human cranium. *Int J Legal Med* 132:1505–1514. <https://doi.org/10.1007/s00414-018-1790-z>
- Nikita E (2019) Quantitative sex estimation based on cranial traits using r functions. *J Forensic Sci* 64:175–180. <https://doi.org/10.1111/1556-4029.13833>
- Yang W, Liu X, Wang K, Hu J, Geng G, Feng J (2019) Sex determination of three-dimensional skull based on improved backpropagation neural network. *Comput Math Methods Med* 2019:1–8. <https://doi.org/10.1155/2019/9163547>
- Bertsatos A, Chovalopoulou M-E, Brůžek J, Bejdová Š (2020) Advanced procedures for skull sex estimation using sexually dimorphic morphometric features. *Int J Legal Med* 134:1927–1937. <https://doi.org/10.1007/s00414-020-02334-9>
- Čechová M, Dupej J, Brůžek J, Bejdová Š, Horák M, Velemínská J (2019) Sex estimation using external morphology of the frontal bone and frontal sinuses in a contemporary Czech population. *Int J Legal Med* 133:1285–1294. <https://doi.org/10.1007/s00414-019-02063-8>
- Yang W, Zhou M, Zhang P, Geng G, Liu X, Zhang H (2020) Skull sex estimation based on wavelet transform and Fourier transform. *Biomed Res Int* 2020:1–10. <https://doi.org/10.1155/2020/8608209>
- Curate F, Coelho J, Gonçalves D et al (2016) A method for sex estimation using the proximal femur. *Forensic Sci Int* 266:579.e1–579.e7. <https://doi.org/10.1016/j.forsciint.2016.06.011>
- Frutos LR (2005) Metric determination of sex from the humerus in a Guatemalan forensic sample. *Forensic Sci Int* 147:153–157. <https://doi.org/10.1016/j.forsciint.2004.09.077>
- Šlaus M, Bedić Ž, Strinović D, Petrovečki V (2013) Sex determination by discriminant function analysis of the tibia for contemporary Croats. *Forensic Sci Int* 226:302.e1–302.e4. <https://doi.org/10.1016/j.forsciint.2013.01.025>
- Ammer S, d'Oliveira Coelho J, Cunha EM (2019) Outline shape analysis on the trochlear constriction and Olecranon fossa of the humerus: insights for sex estimation and a new computational tool. *J Forensic Sci* 64:1788–1795. <https://doi.org/10.1111/1556-4029.14096>
- Kranioti EF, Nathana D, Michalodimitrakis M (2011) Sex estimation of the Cretan humerus: a digital radiometric study. *Int J Legal Med* 125:659–667. <https://doi.org/10.1007/s00414-010-0470-4>
- Tallman SD, Blanton AI (2020) Distal humerus morphological variation and sex estimation in modern Thai individuals. *J Forensic Sci* 65:361–371. <https://doi.org/10.1111/1556-4029.14218>
- İşcan MY, Miller-Shaivitz P (1984) Determination of sex from the Tibia. *Am J Phys Anthropol* 64:53–57. <https://doi.org/10.1002/ajpa.1330640104>
- Kranioti EF, Tzanakis N (2015) Estimation of sex from the upper limb in modern Cretans with the aid of ROC-analysis: a technical report. *Forensic Res Criminol Int J*:1. <https://doi.org/10.15406/frcij.2015.01.00008>
- Robinson MS, Bidmos MA (2011) An assessment of the accuracy of discriminant function equations for sex determination of the femur and tibia from a South African population. *Forensic Sci Int* 206:212.e1–212.e5. <https://doi.org/10.1016/j.forsciint.2010.12.009>
- Ross AH, Manneschi MJ (2011) New identification criteria for the Chilean population: estimation of sex and stature. *Forensic Sci Int* 204:206.e1–206.e3. <https://doi.org/10.1016/j.forsciint.2010.07.028>
- Spradley MK, Jantz RL, Robinson A, Peccerelli F (2008) Demographic change and forensic identification: problems in

- metric identification of Hispanic skeletons. *J Forensic Sci* 53:21–28. <https://doi.org/10.1111/j.1556-4029.2007.00614.x>
31. Steyn M, İşcan MY (1997) Sex determination from the femur and tibia in South African whites. *Forensic Sci Int* 90:111–119. [https://doi.org/10.1016/S0379-0738\(97\)00156-4](https://doi.org/10.1016/S0379-0738(97)00156-4)
 32. Bertsatos A, Chovalopoulou M-E (2020) Advances in osteometric sorting: utilizing diaphyseal CSG properties for lower limb skeletal pair-matching. *J Forensic Sci* 65:1400–1405. <https://doi.org/10.1111/1556-4029.14480>
 33. Bertsatos A, Chovalopoulou M-E (2019) A novel method for analyzing long bone diaphyseal cross-sectional geometry. A GNU Octave CSG Toolkit. *Forensic Sci Int* 297:65–71. <https://doi.org/10.1016/j.forsciint.2019.01.041>
 34. Eliopoulos C, Lagia A, Manolis S (2007) A modern, documented human skeletal collection from Greece. *HOMO* 58:221–228. <https://doi.org/10.1016/j.jchb.2006.10.003>
 35. Bertsatos A, Gkaniatsou E, Papageorgopoulou C, Chovalopoulou M-E (2019) “What and how should we share?” An inter-method inter-observer comparison of measurement error with landmark-based craniometric datasets. *Anthropol Anz*:109–120. <https://doi.org/10.1127/anthranz/2019/1047>
 36. Bertsatos A (2020) long-bone-diaphyseal-CSG-Toolkit: A GNU Octave CSG Toolkit. Zenodo
 37. Eaton JW, Bateman D, Hauberg S, Wehbring R (2019) GNU Octave version 5.1.0 manual: a high-level interactive language for numerical computations
 38. Bertsatos A, Chovalopoulou M-E (2018) Validation study of correct operation for the “long-bone-diaphyseal-CSG-Toolkit.” Zenodo
 39. Bertsatos A, Chovalopoulou M-E, Giannaki K, Valakos E (2018) Morphological variation of the femoral head fovea capitis. *Eur J Anat* 22:397–402
 40. Pedregosa F, Varoquaux G, Gramfort A et al (2011) Scikit-learn: machine learning in Python. *J Mach Learn Res* 12:2825–2830
 41. JASP Team (2020) JASP (Version 0.12.2)[Computer software]
 42. Brzobohatá H, Krajiček V, Horák Z, Velemínská J (2016) Sexual dimorphism of the human tibia through time: insights into shape variation using a surface-based approach. *PLoS ONE* 11:e0166461. <https://doi.org/10.1371/journal.pone.0166461>
 43. Cuzzullin MC, Curate F, Freire AR, Costa ST, Prado FB, Daruge Junior E, Cunha E, Rossi AC (2020) Validation of anthropological measures of the human femur for sex estimation in Brazilians. *Aust J Forensic Sci*:1–14. <https://doi.org/10.1080/00450618.2020.1729411>
 44. Chang C-C, Lin C-J (2011) LIBSVM: a library for support vector machines. *ACM Trans Intell Syst Technol* 2:1–27. <https://doi.org/10.1145/1961189.1961199>
 45. Bertsatos A, Chovalopoulou M-E (2018) Secular change in adult stature of modern Greeks. *Am J Hum Biol* 30:e23077. <https://doi.org/10.1002/ajhb.23077>
 46. Flury B (1997) A first course in multivariate statistics. Springer, New York
 47. Duda RO, Hart PE, Stork DG (2001) Pattern classification, 2nd edn. Wiley, New York
 48. Patriquin ML, Steyn M, Loth SR (2005) Metric analysis of sex differences in South African black and white pelves. *Forensic Sci Int* 147:119–127. <https://doi.org/10.1016/j.forsciint.2004.09.074>
 49. Gonen M, Tanugur AG, Alpaydin E (2008) Multiclass posterior probability support vector machines. *IEEE Trans Neural Netw* 19:130–139. <https://doi.org/10.1109/TNN.2007.903157>
 50. Li H-X, Yang J-L, Zhang G, Fan B (2013) Probabilistic support vector machines for classification of noise affected data. *Inf Sci* 221:60–71. <https://doi.org/10.1016/j.ins.2012.09.041>
 51. Tao Q, Wu G-W, Wang F-Y, Wang J (2005) Posterior probability support vector machines for unbalanced data. *IEEE Trans Neural Netw* 16:1561–1573. <https://doi.org/10.1109/TNN.2005.857955>

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**RESEARCH ARTICLE**

The impact of activity on pelvic age-at-death estimation

Andreas Bertsatos¹ | Maria-Eleni Chovalopoulou^{1,2} |
 Nikoletta-Maria Boskovits¹ | Nefeli Garoufi¹ | Efthymia Nikita²

¹Department of Animal and Human Physiology, Faculty of Biology, School of Sciences, University of Athens, Athens, Greece

²Science and Technology in Archaeology and Culture Research Center, The Cyprus Institute, Nicosia, Cyprus

Correspondence

Efthymia Nikita, Science and Technology in Archaeology and Culture Research Center, The Cyprus Institute, 20 Konstantinou Kavafi Street, 2121 Aglantzia, Nicosia, Cyprus.
 Email: efi.nikita@gmail.com

Funding information

European Regional Development Fund and the Republic of Cyprus, Grant/Award Number: EXCELLENCE/1216/0023; European Union's Horizon 2020 research and innovation programme, Grant/Award Number: 811068

Abstract

This study assesses whether greater levels of physical activity result in increased rates of degenerative change in the pubic symphysis and the iliac auricular surface. The sample comprises 131 skeletons from the Athens Collection. Skeletal age-at-death was estimated from the pubic symphysis and the iliac auricular surface. Skeletal activity was assessed using femoral cross-sectional geometric (CSG) properties and fibrocartilaginous entheses. The association between skeletal age stages, CSG properties and enthesal changes (EC) was tested using Spearman correlation followed with partial Spearman correlation controlling for the effect of documented age-at-death and estimated body mass, as well as generalised linear models. Moreover, Kruskal-Wallis tests were used to compare the EC and CSG values among individuals who were underaged, correctly aged and overaged using the pubic symphysis and the iliac auricular surface. Our results show a negative correlation of skeletal age stages with particular CSG properties, implying a decrease in skeletal rigidity as age increases, and a positive correlation with EC scores, suggesting that older individuals exhibit more pronounced EC. Controlling for documented age-at-death and body mass produced low correlation coefficients and very few statistically significant results. The difference in EC and CSG values among underaged, correctly aged and overaged individuals was significant only for ECs in individuals over 50 years old. The present study highlights that the effect of activity on pelvic age markers is not pronounced, because a limited association between activity and skeletal degeneration in the pubic symphysis and iliac auricular surface was found. Considering the difficulty in identifying past occupations skeletally, our study supports that this often missing parameter in past osteobiographies does not affect skeletal age estimation. Moreover, our results suggest that taking into account the EC scores when evaluating skeletal age can provide further insights in age-at-death estimation in older individuals.

KEYWORDS

biomechanics, enthesal changes, pelvis, skeletal age-at-death

1 | INTRODUCTION

Estimation of age-at-death holds a central role in skeletal analysis, whether the remains come from archaeological or modern contexts.

In the case of bioarchaeological assemblages, age-at-death estimation is key to palaeodemographic reconstruction, and age is also an important factor to consider when examining palaeopathology, activity and other osteobiographic markers (Larsen, 2015). In

forensic anthropology, age-at-death is one of the main elements of the biological profile of an individual, along with sex, stature and ancestry, contributing greatly to the individuation process (Christensen, Passalacqua, & Bartelink, 2014).

Age-at-death estimation in adult remains focuses on different parts of the skeleton, such as the pelvis, thorax and cranium. Some of the most frequently employed methods utilise the morphology of the pubic symphysis and the iliac auricular surface, and these are the anatomical areas that the current paper will also examine (Berg, 2008; Brooks & Suchey, 1990; Buckberry & Chamberlain, 2002; Chen, Zhang, & Tao, 2008; Gilbert & McKern, 1973; Hanihara & Suzuki, 1978; Hartnett, 2010; Igarashi, Uesu, Wakebe, & Kanazawa, 2005; Katz & Suchey, 1986; Kimmerle, Konigsberg, Jantz, & Baraybar, 2008; Lottering, Reynolds, MacGregor, Meredith, & Gregory, 2014; Lovejoy, Meindl, Pryzbeck, & Mensforth, 1985; McKern & Stewart, 1957; Meindl, Lovejoy, Mensforth, & Walker, 1985; Osborne, Simmons, & Nawrocki, 2004; Todd, 1920, 1921a, 1921b). It should be noted that the above-mentioned methods are macroscopic and nondestructive, therefore, easily applicable at a very low cost, and not imposing any of the legal and ethical issues linked with destructive analysis.

Despite the importance of accurate age-at-death estimation and the multiplicity of available methods, the correlation between skeletal age indicators and chronological age is imperfect, an issue that tends to be aggravated as an individual's chronological age increases. This phenomenon is due to the fact that in adults, skeletal age is largely assessed based on the degeneration of the skeleton, and this degeneration is controlled by a number of factors beyond age (Mays, 2015; Nawrocki, 2010). As a result, there is great interindividual and interpopulation variability in the rate of skeletal degeneration and associated skeletal ageing, which renders current skeletal age estimation methods imperfect and limits the applicability of ageing standards in different populations (e.g. Bocquet-Appel & Masset, 1982; Cowie, Bekvalac, & Kausmally, 2008; Falys, Schutkowski, & Weston, 2006; Rissech, Wilson, Winburn, Turbón, & Steadman, 2012). The latter issue has given rise to Bayesian applications in age-at-death estimation during the past years (Bolsen, Milner, Konigsberg, & Wood, 2002; Konigsberg & Frankenberg, 1992; Nikita & Nikitas, 2019; Prince, Kimmerle, & Konigsberg, 2008).

In his review of factors that may affect the expression of skeletal age markers besides age itself, Mays (2015) highlighted that the morphological changes seen on the pubic symphysis can be linked to two primary factors: hormones and activity. Increased oestrogen levels, such as during pregnancy, lead to bony resorption at the pubis (Gardner, 1936; Hall, 1950; Tague, 1988; Vix & Ryu, 1971). Relaxin, a hormone secreted during pregnancy, increases joint laxity in order to allow the joint space at the pubic symphysis to widen sufficiently during childbirth (Dehghan et al., 2014; Keriakos, Bhatta, Morris, Mason, & Buckley, 2011; Marnach et al., 2003). Although joint laxity is greatly reduced after parturition, there remains greater mobility in the symphysis in parous women, which accelerates degenerative changes in this region (Tague, 1988). With regard to the role of activity, stress injuries are often found at the pubic symphysis (*osteitis*

pubis) of individuals regularly involved in strenuous activities and in athletes (Cunningham et al., 2007). Although such injuries may be of an acute nature, most often they result from long-term mechanical stresses applied upon the pelvis (Budak & Oliver, 2013). Bone alterations include erosion of the subchondral bone, sclerosis, cyst formation and, less often, periarticular osteophytes (Budak & Oliver, 2013; Cunningham et al., 2007; Lentz, 1995).

Mays (2015) identified sex and mechanical strain as the key factors affecting also iliac auricular surface morphology. In specific, the increased laxity that characterises the sacroiliac articulation during pregnancy and the greater range of motion of the sacroiliac joints in females compared with males (Damen et al., 2001) must be linked to an earlier occurrence and more rapid progress of degenerative changes in females (Shibata, Shirai, & Miyamoto, 2002). In addition, physically demanding tasks generally increase the risk for osteoarthritis and acute joint injury (Richmond et al., 2013). In the latter case, a notable example is *osteitis condensans ilii*, a stress injury of the sacro-iliac articulation often seen in athletes and occasionally in other groups (Mitra, 2010). In the case of the iliac auricular surface, Mays (2015) identified two additional factors that may affect the rate of degeneration, that is, genetic predisposition and obesity. With regard to the latter, obesity has been linked to the development of osteoarthritis as it imposes increased mechanical loads and biochemical alterations (Pottie et al., 2006; Sowers, 2001), and it also affects the posture of individuals and their motion patterns in sitting and standing, which ultimately also affects the patterns of joint degeneration (Gilleard & Smith, 2007). Hence, body mass may also play a role in the morphology of the iliac auricular surface and by extent affect the relevant skeletal aging markers.

The aim of the current study is to assess whether greater levels of physical activity result in an increase of the rate of degenerative change in the pubic symphysis and the iliac auricular surface. The pelvic joints are associated with locomotion and weight bearing. Thus, a more physically demanding lifestyle is anticipated to lead to an increase in pressure that can affect the rate of degenerative change of the pubic symphysis and iliac auricular surface. Consequently, individuals with more physically active lifestyles should appear older than they are. At the same time, the pubic symphysis and sacroiliac joints are not experiencing identical stress and subsequent joint degeneration; hence, it is interesting to explore how age-at-death estimates based on each of them may be differentially affected. Our study builds on earlier work by Campanacho, Santos and Cardoso (2012), who examined the effect of physical activity on the pubic symphysis, but extends this further by examining also the iliac auricular surface, employing a range of skeletal markers of robusticity and including among the predictors sex and body mass.

2 | MATERIALS AND METHODS

2.1 | The Athens Collection

The dataset used to test the effect of activity on pelvic skeletal age included 131 skeletons from the University of Athens Modern Human

Skeletal Reference Collection (hereafter Athens Collection), housed at the Department of Animal and Human Physiology at the National and Kapodistrian University of Athens, Greece. The Athens Collection comprises the skeletal remains of individuals who lived mostly in the 20th century and were buried in cemeteries in the area of Athens. The age-at-death, sex, occupation, place and year of birth, as well as the cause of death for most of these individuals are documented in their death certificates (Eliopoulos, Lagia, & Manolis, 2007; Nikita, 2020). The distribution of the material per sex and age group is given in Table 1. The selection of the skeletons was made on the ground of good preservation of (a) the pubic symphysis and iliac auricular surface, (b) full length of the femur, (c) femoral head diameter and (d) bone surfaces so that the presence or absence of femoral enthesal changes (EC) could be recorded. All selected individuals lacked any evidence of genetic disorders that could potentially act as confounding variables to this sample. In addition, we tried to keep the sample as balanced as possible with regard to the representation of males and females as well as different age groups.

2.2 | Skeletal age-at-death estimation methods

The skeletal age-at-death estimation methods used included the Brooks and Suchey (1990) method for the pubic symphysis and the Lovejoy et al. (1985) as well as the Buckberry and Chamberlain (2002) methods for the iliac auricular surface. These methods will be denoted as SB, L and BC, respectively. Data collection took place in the context of an earlier study (Xanthopoulou, Valakos, Youlatos, & Nikita, 2018) where intrarater and interrater agreement were also examined and found to be statistically significant (Cohen's $\kappa > 0.7$, p -value < 0.05). In the context of this earlier study, the correlation between documented age-at-death (chronological age) and the skeletal stages used in the above methods was tested, and it was found to be statistically significant for both sexes and for all three methods ($r_s = 0.32$ – 0.79 , p -value < 0.01 ; Table 6 in Xanthopoulou et al., 2018). Similarly, the percentage of individuals assigned to the correct age class using these methods was around 70% and over 80% for males and females, respectively, when using SB, at least 78% for both sexes when using L, but merely 56% when adopting BC (Table 2 in Xanthopoulou et al., 2018).

TABLE 1 Sample composition

Age group	Males	Females
20–30	13	4
31–40	7	6
41–50	13	10
51–60	16	9
61–70	14	6
71–80	7	11
>81	6	9
Total	76	55

2.3 | Activity reconstruction

As mentioned above, the profession/occupation of many individuals that form the Athens Collection is also documented. However, we decided to omit this information and assess the physical activity levels skeletally because documented occupations in modern reference collections are often misleading (Cardoso & Henderson, 2013). In particular, in the Athens Collection, many occupations were vague, such as ‘civil servant’ or ‘housekeeper’. Furthermore, death certificates usually state only the last known occupation. These limitations do not allow the assessment of how mechanically strenuous daily life was for these individuals. Hence, the activity levels of each individual were assessed skeletally using the cross-sectional geometric (CSG) properties of the femur, as well as femoral fibrocartilaginous EC.

The CSG properties used in the analyses included the total subperiosteal area (TA), which is a measure of bone robusticity particularly with respect to tension, compression or shear, as well as ratios of second moments of area I_x/I_y and I_{max}/I_{min} , whereby I_x , I_y , I_{max} and I_{min} express resistance to bending loads, and their ratios are indicative of cross-sectional shape as they represent the distribution of cortical bone within the section (Ruff, 1987, 2008). These properties were calculated from virtual 3D bone models of the Athens Collection with the long-bone-diaphyseal-CSG-Toolkit (Bertsatos & Chovalopoulou, 2019). The CSG Toolkit is a free computer software that automatically analyses 3D models of human long bones and provides highly accurate measurements of their CSG properties along the diaphyseal shaft at regular intervals (20%, 35%, 50%, 65% and 80% of the maximum bone length). These properties are estimated based on the subperiosteal contours of the femur, that is, they exclude the medullary cavity. This approach has been found to exhibit a very high correlation with true cross-sectional geometry (e.g. $r^2 = 0.998$ in Stock & Shaw, 2007). It must be noted that with increasing age, there is endosteal resorption and cortical thinning, which might affect the estimation of CSG (Ruff & Hayes, 1983), especially when using a method that excludes the medullary cavity. Nonetheless, Sparacello and Pearson (2010) showed that cortical thickness has minimal impact on average biomechanical properties; thus, the data obtained from subperiosteal contours may accurately reflect bending rigidity.

To avoid including too many intercorrelated parameters in our statistical tests, we only used the properties estimated at the midshaft (50% of maximum length). The practice of using the midshaft value as a sufficient measure to express diaphyseal rigidity has been adopted by numerous scholars (e.g. Sparacello & Pearson, 2010; Stock & Shaw, 2007). The femoral 3D models used in the present study have been reconstructed from their respective dry bone counterparts by means of 3D photogrammetry with a rated accuracy of ~ 0.2 mm (Bertsatos, Gkaniatsou, Papageorgopoulou, & Chovalopoulou, 2019). The rated accuracy corresponds to the maximum femoral length measurement, which implies that the expected measurement error in CSG calculated values is much less.

The femoral fibrocartilaginous entheses examined included *gluteus minimus*, *psaos major*, *iliacus* and *gluteus medius*. All entheses were preferably recorded on the left side limbs, unless these were

damaged, in which case the right side was preferred. EC were recorded using the New Coimbra Method (Henderson, Mariotti, Pany-Kucera, Villotte, & Wilczak, 2016). An estimate of repeatability of the recorded values was established by interval-by-interval intraobserver agreement on a subsample of 22 femurs, which were recorded twice with a gap of 3 weeks between each recording session. Intraobserver agreement was calculated for each individual variable based on the scored difference and an average of 92.02% was established over all entheses examined. Given that this recording scheme involves multiple variables, such as bone formation and erosion, microporosity and macroporosity, cavitations, and textural change, in most statistical analyses, we summed the values of these variables so that each enthesis would be represented by a single composite value. We acknowledge that this approach has the limitation that certain variables (EC) may be more strongly linked to activity levels and this effect may be lost within a composite score. For this reason, we also performed selected analyses using every single EC variable, as outlined below, in order to make sure that we have avoided missing any significant activity effect. Note that the effectivity of EC as skeletal activity markers has been criticised in a number of studies (e.g. Cardoso & Henderson, 2010; Michopoulou, Nikita, & Valakos, 2015; Milella, Belcastro, Zollikofer, & Mariotti, 2012; Nikita, Xanthopoulou, Bertsatos, Chovalopoulou, & Hafez, 2019; Weiss, 2003, 2004), many of which have stressed that the primary factors affecting EC expression are age and body size/mass rather than mechanical stress (Cardoso & Henderson, 2010; Milella et al., 2012; Niinimäki, 2011; Weiss, 2003, 2004, 2007; Weiss, Corona, & Schultz, 2012; Wilczak, 1998). Nonetheless, recent research has experimentally reaffirmed the link between activity levels and EC (Karakostis, Jeffery, & Harvati, 2019; Karakostis, Wallace, Konow, & Harvati, 2019). On this ground, we have included EC as a skeletal activity marker, bearing in mind its inherent limitations. Note that individuals exhibiting seronegative spondyloarthropathies, DISH or other diseases that may accentuate the expression of EC were not included in our sample (Villotte & Knüsel, 2013). Additional to these pathological categories, any ‘bone-formers’, that is, individuals with an inherent tendency toward bone formation, were similarly excluded from the sample (Mays, 2016; Rogers, Shepstone, & Dieppe, 1997).

Tables S1 and S2 provide descriptive statistics for the CSG properties and EC per sex and age group.

2.4 | Body mass estimation

The body mass of the individuals comprising the Athens Collection is not known; hence, it had to be skeletally estimated. The estimation of body mass from skeletal remains may adopt ‘mechanical’ or ‘morphometric’ approaches (Auerbach & Ruff, 2004). Mechanical approaches are based on the high correlation that exists between body mass and the dimensions of the skeletal elements that support the weight of the body, mostly the femoral head diameter (Auerbach & Ruff, 2004). Morphometric approaches view the human body as a cylinder, with stature equal to the cylinder’s height and iliac breadth equal to its

diameter. As such, these methods use the stature and iliac breadth to estimate body mass (Ruff, 2010). To produce population-specific equations for body mass estimation, a so-called ‘hybrid’ approach may be employed, whereby the morphometric method is used to calculate the body mass, and then regression equations are obtained using femoral head diameter (e.g. Ruff et al., 2012). This approach was adopted by Nikita and Chovalopoulou (2017) based on the Athens Collection using 75 individuals whose skeletal remains were sufficiently well preserved to allow the estimation of stature and the measurement of iliac breadth so that the morphometric method could be applied. The regression equations generated from that study for the estimation of body mass based on femoral head diameter exhibited low random and directional error and were adopted in the current paper to estimate body mass in the 131 skeletons under examination. At this point, we must stress that a rather recent study found that the estimation of body mass from skeletal remains is only accurate and reliable for estimating the population average body mass, not that of individuals (Jeanson et al., 2017). Despite this potential limitation, body mass here is used only as a proxy of body size to examine if larger body size affects the rate of degeneration of the auricular surface and pubic symphysis and not as an absolute measure of the weight of the individuals under study.

2.5 | Statistical analysis

First, we assessed whether there is a significant bilateral asymmetry in the skeletal age marker scores in order to test if we need to examine the association between activity and skeletal ageing in each side separately or not. Wilcoxon signed ranks tests were used to test for the aforementioned bilateral asymmetry. Subsequently, the correlation between skeletal age stages and CSG properties was tested using Spearman correlations. Additionally, we adopted Monte Carlo permutations for the p -value with 1000 iterations. This analysis was followed up with partial Spearman correlations to test again the correlation between skeletal age stages and CSG properties but this time controlling for the effect of chronological age and body mass. Spearman correlations were also used to test the correlation between skeletal age stages and summed EC scores, whereas partial Spearman correlations tested this association controlling for the effect of chronological age and body mass. All (partial) correlations were ran for pooled sexes and separately for males and females to test for any sex bias. Partial Spearman correlations with chronological age and body mass as covariates were rerun for the EC, this time separately for each individual variable, rather than for the summed scores, in order to assess whether there is a significant pattern missed due to the summing process. In addition, generalised linear models (GLM) were used to examine simultaneously the effect of sex, chronological age, body mass, EC sums and CSG properties on the skeletal age stages. In this case, the tests were repeated twice, treating EC sums both as ordinal and as continuous/scale data. Note that EC scores are ordinal variables; however, there are instances in bioarchaeology and forensic anthropology where ordinal variables are treated as continuous when

modelling data (e.g. Klales, Ousley, & Vollner, 2012; Walker, 2008). Such an approach is based on the fact that models based on continuous data are much simpler, whereas when using ordinal variables, all levels of each factor should be present in the dataset, a precondition often not satisfied. For these reasons, in the current study, we treated EC scores in the GLM models both as ordinal and as continuous variables (for more details, see Milella, Belcastro, Mariotti, & Nikita, 2020).

Considering that the accuracy of skeletal ageing methods tends to decrease in older individuals, as an additional means of exploring the effect of activity on skeletal ageing, we divided our sample in individuals younger than or equal to 50 years and older than 50 years. Subsequently, we classified each individual in one of three categories based on the performance of each ageing method: underaged, correctly aged and overaged. This classification was based on whether an individual's chronological age fell within the interval mean age \pm standard deviation provided in the publication of the SB and BC methods and within the age intervals suggested for the L method. The EC scores and the CSG values were then compared among the underaged, correctly aged and overaged individuals using Kruskal–Wallis tests in order to assess if these skeletal activity markers are significantly differentiated in individuals for whom skeletal ageing exhibits differential success.

All correlations were ran using Excel macros developed by Nikita (2017), whereas Wilcoxon signed ranks tests, GLM and Kruskal–Wallis tests were ran in SPSS v. 21.0. The data that support the findings of this study are available from the corresponding author.

3 | RESULTS AND DISCUSSION

No statistically significant difference was found in the bilateral expression of any of the skeletal age stages (SB: Wilcoxon $Z = -0.218$,

p -value = 0.827; L: Wilcoxon $Z = -1.753$, p -value = 0.080; BC: Wilcoxon $Z = -0.865$, p -value = 0.387). Based on these results, we decided to analyse the impact of activity only on the left os coxae.

First, we tested the Spearman correlation between skeletal age stages and CSG properties, additionally adopting Monte Carlo permutations for the p -value with 1000 iterations. The results are given in Table 2. For pooled sexes, it can be seen that this correlation is very small (rho values ranging from -0.225 to 0.149). Nonetheless, it is statistically significant for lx/ly (p -values < 0.05) for all three skeletal age methods. The negative correlation coefficients for lx/ly suggest that as skeletal degeneration increases, the rigidity of the skeleton to bending forces expressed as the distribution of cortical bone along the antero-posterior versus medio-lateral axes of the cross-sectional plane overall decreases, as expected. For TA, rho is also negative, suggesting a decrease in the rigidity of the skeleton to compressive, tensile and shear forces as skeletal degeneration increases; however, this was only significant for the pubic symphysis (p -value = 0.03). In the case of $lmax/lmin$, we obtain positive rho values, which might appear contradictory to the lx/ly results as they suggest an increase in bending rigidity with increasing skeletal degeneration. Nonetheless, $lmax$ and $lmin$ represent orthogonal maximum and minimum bending rigidity along an arbitrary axis on the cross-sectional plane. Taking into account that TA decreases with skeletal age, this merely reflects that the cross-sectional area becomes more elongated suggesting that the femoral diaphysis weakens even further along some particular orientation of bending force. In any case, this pattern was only significant for BC (p -value = 0.04).

The next step was to examine how the above correlations are affected by the chronological age-at-death and body mass of the individuals. For this reason, we repeated the analysis using partial Spearman's correlations, controlling for the effect of chronological age and body mass. The results are also given in Table 2, and it can be seen that for pooled sexes, the only statistically significant correlation

TABLE 2 Spearman correlation coefficients between skeletal age stages and CSG properties and partial Spearman correlation coefficients while controlling for chronological age and body mass

Skeletal ageing method	CSG property	Pooled sexes		Males		Females	
		Spearman's rho	Partial Spearman's rho	Spearman's rho	Partial Spearman's rho	Spearman's rho	Partial Spearman's rho
SB	TA	-0.165^*	-0.075	0.047	-0.136	0.233^*	0.188
L	TA	0	-0.079	0.094	-0.097	0.164	0.119
BC	TA	-0.072	-0.188^*	0.069	-0.173	-0.089	-0.239^*
SB	lx/ly	-0.225^{**}	-0.042	-0.190^*	-0.072	-0.128	-0.010
L	lx/ly	-0.154^*	0.047	-0.099	0.069	-0.124	0.026
BC	lx/ly	-0.137^*	0.041	-0.118	0.033	-0.062	0.070
SB	$lmax/lmin$	0.076	0.011	0.074	-0.043	0.006	0.023
L	$lmax/lmin$	0.081	0.011	0.195^*	0.131	-0.119	-0.158
BC	$lmax/lmin$	0.149^*	0.115	0.306^{**}	0.301^{**}	-0.115	-0.137

Abbreviations: BC, Buckberry and Chamberlain (2002); CSG, cross-sectional geometric; L, Lovejoy et al. (1985); SB, Brooks and Suchey (1990); TA, total subperiosteal area.

* $p < 0.05$. ** $p < 0.01$.

is between TA and BC, whereas rho values cluster even closer to zero. This result supports a very limited association between activity and skeletal degeneration in the pubic symphysis and iliac auricular surface.

When repeating all analyses separately for males and females to test for any sex bias (Table 2), the number of significant correlations between CSG and skeletal age stages decreased notably, whereas when chronological age and body mass were controlled for, merely two significant results were obtained (males BC-*I*_{max}/*I*_{min} rho = 0.301, *p*-value = 0.004; females BC-TA rho = -0.239, *p*-value = 0.014). The smaller number of significant results may be attributed to the smaller sample size as this is now broken down in males and females. In any case, there does not appear to be any clear sex-related pattern in the effect of activity on skeletal age stages, whereas this effect remains minimal, as was the case for pooled data.

Subsequently, we tested the Spearman correlation between the skeletal age stages and summed EC scores. The results are given in Table 3. For pooled sexes, it is seen that the correlation between skeletal degeneration and EC is always positive (rho ranging from 0.156 to 0.642) and statistically significant. This result agrees with earlier studies that have identified a key role of age on EC expression, that is, an increase in EC scores with increasing age (see Section 1). As our study aims at exploring whether heightened activity may accelerate skeletal degeneration in the pubic symphysis and iliac auricular surface, we repeated our analysis using partial Spearman's correlations, controlling for the effect of chronological age and body

mass. The results are again given in Table 3. It can be seen that the correlation between EC and skeletal degeneration in the pubic symphysis and iliac auricular surface ceases being significant and the correlation coefficients drop substantially (-0.145 to 0.165). The only significant results are found for the L method for *iliacus* and *psaos major*, but they are conflicting as *psaos major* shows a positive correlation whereas *iliacus* exhibits a negative correlation with auricular surface degeneration. These results suggest that even though there is a significant correlation between ECs and skeletal age stages, this association is due to the effect of chronological age and to some extent body mass, which simultaneously affect the expression of both variables. Once the effect of age and body mass is removed, there is a lack of association between activity, as captured by means of femoral ECs, and skeletal age stages. In other words, increased activity (greater EC scores) does not appear to accelerate the degeneration of the pubic symphysis or iliac auricular surface.

The examination of an association between ECs and skeletal age stages per sex (Table 3) gave significant results in all cases, except for the *iliacus*, which was significantly correlated only with L for females (rho = 0.232, *p*-value = 0.04). When chronological age-at-death and body mass were included as covariates, only two significant patterns were observed and only among the male sample (*psaos major*-L: rho 0.194, *p*-value = 0.049; *iliacus*-L: rho = -0.201, rho = 0.043). These results agree with those for pooled sexes that activity, as expressed by means of femoral EC, does not appear to affect the rate of skeletal degeneration in the pubic symphysis and iliac auricular surface, and

TABLE 3 Spearman correlation coefficients between skeletal age stages and summed EC scores and partial Spearman correlation coefficients while controlling for chronological age and body mass

Skeletal ageing method	EC	Pooled sexes		Males		Females	
		Spearman's rho	Partial Spearman's rho	Spearman's rho	Partial Spearman's rho	Spearman's rho	Partial Spearman's rho
SB	<i>Gluteus minimus</i>	0.426**	-0.108	0.426**	-0.112	0.367**	-0.114
L	<i>Gluteus minimus</i>	0.509**	-0.064	0.516**	0.019	0.464**	-0.152
BC	<i>Gluteus minimus</i>	0.461**	-0.065	0.488**	-0.047	0.394**	-0.117
SB	<i>Psoas major</i>	0.511**	0.018	0.523**	0.010	0.431**	-0.027
L	<i>Psoas major</i>	0.642**	0.165*	0.634**	0.194*	0.607**	0.123
BC	<i>Psoas major</i>	0.502**	-0.049	0.534**	-0.036	0.420**	-0.109
SB	<i>Iliacus</i>	0.156*	-0.090	0.153	-0.147	0.161	-0.033
L	<i>Iliacus</i>	0.174*	-0.145*	0.138	-0.201*	0.232*	0.021
BC	<i>Iliacus</i>	0.176*	-0.098	0.175	-0.148	0.179	0.003
SB	<i>Gluteus medius</i>	0.367**	-0.050	0.395**	-0.057	0.323**	-0.031
L	<i>Gluteus medius</i>	0.410**	-0.097	0.415**	-0.060	0.382**	-0.071
BC	<i>Gluteus medius</i>	0.345**	-0.137	0.383**	-0.125	0.261*	-0.151

Abbreviations: BC, Buckberry and Chamberlain (2002); EC, enthesal changes; L, Lovejoy et al. (1985); SB, Brooks and Suchey (1990).

p*-value < 0.05. *p*-value < 0.01.

this applies to both sexes. Instead, the strong correlation between EC scores and skeletal age stages is due to the chronological age of the individuals, which affects both simultaneously.

As stated above, the fact that we have summed the ordinal scores of the different variables per enthesis may have ‘masked’ some significant effects. Many of the Coimbra method variables systematically exhibited zero values in our dataset, whereas others exhibited diversity in their scores; thus, by summing all scores per

enthesis, we may be losing the effect of the variables that are driving variation in EC expression in our sample. To examine whether this is the case, we repeated the partial correlations with chronological age and body mass as covariates, separately for each of the EC variables. The results are given in Table 4, where it can be seen that 31 correlations were significant out of the 288 (~11%). Most of the significant correlations (21 out of the 31) were identified in Zone 1, but no other pattern emerged with

TABLE 4 Partial Spearman correlation coefficients between skeletal age stages and ECs controlling for chronological age and body mass

EC	Variable	SB			L			BC		
		Pooled sexes	Males	Females	Pooled sexes	Males	Females	Pooled sexes	Males	Females
<i>Gluteus minimus</i>	BF Z1	-0.129	-0.282**	-0.033	-0.078	-0.130	-0.020	-0.152*	-0.218*	-0.152
	ER Z1	-0.132	-0.236*	0.005	-0.218**	-0.189*	-0.265*	-0.115	-0.137	-0.113
	BF Z2	-0.087	-0.114	-0.082	-0.078	0.022	-0.130	-0.059	0.014	-0.098
	ER Z2	0.113	0.154	0.068	-0.059	0.110	-0.234*	-0.064	0.006	-0.138
	TC	-0.035	-0.105	-0.021	-0.134	-0.154	-0.144	0.046	-0.018	0.061
	FPO	-0.170*	-0.174	-0.181	-0.110	-0.064	-0.130	-0.005	-0.030	0.054
	MPO	-0.060	0.018	-0.133	0.056	0.161	-0.090	-0.064	0.060	-0.227
	CA	0.032	0.094	0.035	-0.013	0.009	-0.005	-0.075	-0.065	0.008
<i>Psoas major</i>	BF Z1	-0.086	0.095	-0.218	0.158*	0.207*	0.216	-0.017	-0.084	0.165
	ER Z1	-0.088	-0.024	-0.205	-0.156*	-0.251*	0.097	-0.156*	-0.195*	-0.060
	BF Z2	0.109	0.054	0.104	0.112	0.175	0.059	-0.039	-0.024	-0.077
	ER Z2	-0.009	-0.103	0.061	-0.037	-0.052	0.005	-0.105	-0.098	-0.165
	TC	0.073	0.130	-0.010	0.101	0.140	0.017	0.072	0.122	-0.025
	FPO	0.026	-0.101	0.129	0.139	0.299**	-0.129	0.134	0.139	0.064
	MPO	0.026	0.093	-0.083	-0.099	0.074	-0.339**	-0.050	-0.024	-0.145
	CA	0.006	-0.077	0.088	0.173*	0.142	0.164	-0.141	-0.067	-0.303*
<i>Iliacus</i>	BF Z1	-0.244**	-0.266*	-0.152	-0.217**	-0.212*	-0.146	-0.188*	-0.319**	0.010
	ER Z1	-0.103	-0.105	0.035	-0.035	-0.023	-0.005	-0.072	-0.068	0.009
	BF Z2	0.090	0.060	0.129	0.069	0.053	0.103	0.032	0.133	-0.123
	ER Z2	-0.111	-0.200*	0.066	0.046	0.022	0.141	0.049	-0.051	0.179
	TC	-0.009	0.022	0.035	-0.007	0.008	-0.005	-0.021	0.027	0.009
	FPO	0.020	-0.021	-0.028	-0.061	-0.154	0.117	-0.008	0.009	0.005
	MPO	0.055	-0.044	0.161	-0.100	-0.114	-0.080	-0.025	0.008	-0.090
	CA	-0.009	0.022	0.035	-0.007	0.008	-0.005	-0.021	0.027	0.008
<i>Gluteus medius</i>	BF Z1	0.033	-0.012	0.067	-0.049	-0.066	0.045	-0.059	-0.016	-0.109
	ER Z1	-0.039	-0.075	0.023	-0.129	-0.044	-0.182	-0.195*	-0.088	-0.315*
	BF Z2	-0.008	0.072	-0.083	-0.071	-0.070	0.032	-0.099	-0.081	-0.053
	ER Z2	-0.067	-0.097	-0.024	-0.166*	-0.113	-0.204	-0.176*	-0.204*	-0.135
	TC	-0.024	0.022	-0.009	0.010	0.008	0.017	-0.024	0.027	-0.025
	FPO	-0.134	-0.148	-0.177	-0.083	-0.036	-0.166	-0.099	-0.142	-0.098
	MPO	-0.001	-0.011	0.048	0.002	0.130	-0.216	-0.051	-0.001	-0.089
	CA	0.050	0.051	0.021	0.049	0.054	-0.012	0.005	-0.023	-0.017

Abbreviations: BC, Buckberry and Chamberlain (2002); BF Z1, bone formation in zone 1; BF Z2, bone formation in zone 2; CA, cavitation; EC, enthesal changes; ER Z1, erosion in Z1; ER Z2, erosion in Z2; FPO, fine porosity; L, Lovejoy et al. (1985); MPO, macro-porosity; SB, Brooks and Suchey (1990); TC, textural change.

*p-value < 0.05. **p-value < 0.01.

regard to the enthesis, the specific EC or the skeletal ageing markers most commonly exhibiting a significant effect.

In order to examine simultaneously the impact of sex, body mass, chronological age, ECs and CSG on skeletal age stages, we employed GLM. Because the analysis of each individual EC variable (Table 4) did not reveal any particular pattern, and rather the general conclusion that there is no systematic effect of EC on skeletal ageing agreed with the results obtained from the summed scores, we decided to use summed EC scores in the GLM. An additional reason for our decision was the fact that the inclusion of too many variables in GLM would have resulted in unnecessary complex models. The results given in Table 5 show that chronological age is the only factor that systematically has a significant effect on the expression of skeletal age stages, as expected. Sex is significant in the pubic symphysis, suggesting that there are differences between males and females in the rate of degeneration of this joint, which are not linked to activity, body mass or chronological ageing and must be attributed to other factors, such as hormonal differences (see Mays, 2015). CSG properties never exhibit a significant effect on skeletal degeneration in the pelvic joints. Finally, among the ECs, only the *gluteus minimus* and *iliacus* exhibit some significant effect on the iliac auricular surface morphology. Overall, the GLM results support a strong, as expected, effect of chronological age, significant sexual dimorphism for the pubic symphysis, and limited effect of activity on the iliac auricular surface, as captured by a few of the femoral ECs.

The comparison of EC scores and CSG values among underaged, correctly aged and overaged individuals (Tables 6 and 7) largely corroborates the above findings and offers additional interesting insights. It can be seen that among the CSG properties (Table 6), only TA is very rarely significantly different among these groups, supporting a lack of impact of activity, as deduced through CSG properties, on the degeneration of the pubic symphysis and the iliac auricular surface. With regard to EC (Table 7), when there is no separation of the dataset in younger and older individuals, these are almost always significantly different among underaged, correctly aged and overaged individuals. However, upon closer inspection of the results, it is observed that the significant difference is seen almost exclusively in individuals over 50 years old; hence, it must be linked to the well-documented effect that chronological age has on the expression of EC rather than on the effect that the latter have on skeletal degeneration in the pelvic joints. It is interesting to note that the above pattern was not visible in the L method, where no significant results were obtained when dividing the samples in age groups, contrary to the BC

one, where it is very clear. Even though both methods focus on the same anatomical area, the iliac auricular surface, they adopt distinct ways of recording morphological changes in this area, and they associate different age intervals with different morphological changes (notably, the L method places all individuals over 60 years in a single age group). It is also interesting to observe that the significant difference in EC expression between underaged, correctly aged and overaged individuals when employing the SB method is only seen in females over 50 years old but not in males. This sex dimorphic pattern is most likely linked to hormonal changes related to menopause, which affects skeletal degeneration in older females differently than males (Ji & Yu, 2015). However, other sex-related physiological and functional differences such as parturition might explain why this differentiation between males and females is only observed in the pubic symphysis and not in the auricular surface (Becker, Woodley, & Stringer, 2010).

Our results overall suggest that there is often a significant association between skeletal activity markers and skeletal age stages. This association is identified particularly in *Ix/Iy* and TA among the CSG properties, as well as in all EC. Nonetheless, once the effect of chronological age is controlled for, this association diminishes, suggesting that it is directly attributed to the chronological age of the individuals, which simultaneously affects the expression of skeletal activity markers and skeletal age stages. Similarly, differences in CSG scores are very rarely present among underaged, correctly aged and overaged individuals, while such differences are found rather systematically only in individuals older than 50 years for ECs, supporting the key role of chronological age in the expression of morphological changes both on the pelvic joints and the entheses. Taken together, the results of our analyses support that in our dataset, there does not appear to be a direct effect of activity on skeletal degeneration in the pubic symphysis and iliac auricular surface. These results agree with those of Campanacho et al. (2012), who investigated pubic symphyseal morphology in males in a recent skeletal collection from Portugal and did not find an association with physically demanding occupations. A similar study, although focused on the acetabulum, employing material from Christ Church Spitalfields, London, was conducted by Mays (2012) and also did not find any association between manual occupations and more rapid age changes.

A point worth highlighting is the possible impact of body mass. One could argue that heavier individuals would be expected to exhibit higher skeletal age marker scores at an earlier age than normal in the sense that the increased mechanical loading of the body mass on the pelvis should have resulted in a faster degeneration of the pelvic joints

TABLE 5 Generalised linear model results (*p*-values)

Dependent variable	Predictors									
	Sex	BM	<i>Gluteus minimus</i>	<i>Psoas major</i>	<i>Iliacus</i>	<i>Gluteus medius</i>	Age	TA	<i>Ix/Iy</i>	<i>I_{max}/I_{min}</i>
SB	0.017	0.973	1.000	0.955	0.372	1.000	0.000	0.438	0.254	0.929
L	0.180	0.824	0.937	–	0.002	0.659	0.000	0.163	0.332	0.942
BC	0.053	0.132	0.009	0.097	0.107	0.511	0.000	0.351	0.176	0.635

Abbreviations: BC, Buckberry and Chamberlain (2002); BM, body mass; L, Lovejoy et al. (1985); SB, Brooks and Suchey (1990).

TABLE 6 Kruskal–Wallis test results for the comparison of CSG scores among underaged, correctly aged and overaged individuals

Ageing method	Groups	TA				Ix/Iy				I _{max} /I _{min}			
		Mean ranks				Mean ranks				Mean ranks			
		1	2	3	p-value	1	2	3	p-value	1	2	3	p-value
SB	All	73.7	66.5	40.6	0.019	60.0	67.33	76.5	0.332	66.1	68.0	54.7	0.483
	Males	45.5	35.6	32.2	0.166	34.4	40.1	43.4	0.521	41.2	37.9	31.0	0.615
	Females	29.4	31.0	15.6	0.037	25.3	27.2	35.7	0.270	25.4	31.2	22.6	0.260
	≤50 years	29.7	29.8	19.3	0.093	14.7	28.6	25.6	0.299	12.7	29.7	23.2	0.105
	>50 years	42.5	36.7	–	0.254	41.1	38.0	–	0.549	40.2	38.8	–	0.780
	M ≤ 50 years	18	17.1	16.2	0.972	10.5	17.7	16	0.580	8.5	18.2	14.4	0.320
	M > 50 years	24.5	19.4	–	0.181	21.7	22.3	–	0.884	23.7	20.2	–	0.356
	F ≤ 50 years	13.0	11.3	9.3	0.701	2.0	11.3	10.6	0.325	5.0	12.8	8.6	0.188
	F > 50 years	16.8	19.0	–	0.529	19.2	17.0	–	0.529	16.8	19.1	–	0.508
L	All	73.7	61.8	66.6	0.340	59.6	63.8	76.4	0.158	57.1	71.9	63.6	0.174
	Males	40.4	36.9	39.0	0.843	31.3	39.4	44.4	0.136	30.9	44.4	37.5	0.084
	Females	29.2	29.7	22.5	0.403	26.0	27.1	32.2	0.583	28.5	28.4	26.6	0.942
	≤50 years	26.5	24.0	28.5	0.698	22.0	28.3	28.6	0.443	18.7	30.1	29.2	0.103
	>50 years	47.8	35.7	44.0	0.107	38.9	40.2	34.4	0.851	37.9	42.1	19.6	0.099
	M ≤ 50 years	18.3	12.0	18.7	0.243	11.8	14.6	19.6	0.165	11.5	17.6	18.5	0.299
	M > 50 years	22.0	21.8	24.0	0.958	20.9	23.9	12.7	0.310	18.2	25.6	13.3	0.086
	F ≤ 50 years	11.7	11.3	9.5	0.747	10.3	13.5	9.4	0.502	7.2	13.8	11.2	0.197
	F > 50 years	20.4	17.2	23.5	0.597	12.4	18.7	22.5	0.368	25.8	17.5	5.5	0.051
BC	All	85.5	69.5	59.5	0.149	59.2	61.0	73.2	0.185	86.2	63.1	67.5	0.337
	Males	67.3	39.2	34.8	0.049	34.0	35.2	43.5	0.267	62.0	35.6	40.2	0.115
	Females	37.0	30.8	23.6	0.162	27.0	24.9	31.7	0.312	28.0	28.4	27.6	0.985
	≤50 years	–	34.1	25.4	0.107	–	25.6	27.3	0.750	–	20.4	28.5	0.134
	>50 years	49.8	38.4	40.1	0.494	41.2	39.7	37.5	0.936	49.7	38.5	39.6	0.514
	M ≤ 50 years	–	18.0	16.6	0.716	–	14.0	18.1	0.275	–	12.9	18.5	0.135
	M > 50 years	37.0	20.8	21.3	0.100	21.0	22.1	22.1	0.990	33.0	20.8	23.0	0.265
	F ≤ 50 years	–	15.0	10.3	0.435	–	15.0	10.3	0.435	–	10.0	10.5	0.931
	F > 50 years	22.3	17.4	18.4	0.732	20.7	17.9	16.8	0.872	18.0	18.3	16.2	0.913

Abbreviations: 1 = underaged; 2 = correctly aged; 3 = overaged; BC, Buckberry and Chamberlain (2002); L, Lovejoy et al. (1985); SB, Brooks and Suchey (1990), TA, total subperiosteal area.

in bigger individuals. Nonetheless, no significant body mass effect was found in our data in the GLM models. This finding is also in agreement with earlier studies. In modern US collections, Wescott and Drew (2015) found that obese individuals had aged more rapidly in the auricular surface but not in the pubic symphysis, whereas Merritt (2015) found that individuals with low body size actually tended to age earlier according to the degeneration of the pubic symphysis, auricular surface and sternal rib end.

There are many factors that may explain the overall lack of significant association between activity and earlier onset of skeletal degeneration in the pelvis. First of all, the skeletal activity markers we employed have inherent limitations. As mentioned in Section 1, numerous studies have criticised the role of EC as skeletal activity markers, stressing that the expression of these changes is largely controlled by the chronological age of the individuals and their body mass

(Benjamin, Toumi, Suzuki, Hayashi, & McGonagle, 2009; Cardoso & Henderson, 2010; Mariotti, Facchini, & Belcastro, 2004, 2007; Michopoulou et al., 2015; Milella et al., 2012; Niinimäki, 2011; Robb, 1998; Villotte, 2009; Weiss, 2003, 2004, 2007; Wilczak, 1998). In addition, even though CSG properties are considered accurate skeletal markers of activity, they express resistance to mechanical loads applied on the long bones along specific planes, and these may not correspond to the same mechanical loads applied on the pubic symphysis or the iliac auricular surface.

A final factor is the possible impact of parameters that could not be taken into account in our study, such as diet and disease. The quality of diet affects bone formation and loss, bone mass and mineral density, as well as bone remodelling rates (Sahni, Zoltick, McLean, & Hannan, 2010). In our material, just like in most archaeological and modern skeletal collections, it is not possible to know the dietary quality

TABLE 7 Kruskal–Wallis test results for the comparison of EC summed scores among underaged, correctly aged and overaged individuals

Ageing method	Groups	<i>Gluteus minimus</i>					<i>Psoas major</i>					<i>Iliacus</i>					<i>Gluteus medius</i>				
		Mean ranks			p-value	Mean ranks			p-value	Mean ranks			p-value	Mean ranks			p-value	Mean ranks			
		1	2	3		1	2	3		1	2	3		1	2	3					
SB	All	91.8	57.4	37.0	0.000	87.7	57.7	37.0	0.000	80.2	59.2	56.4	0.006	85.8	59.1	35.5	0.000				
	Males	53.4	32.4	24.5	0.000	49.5	32.6	24.3	0.002	46.1	33.7	38.4	0.059	49.5	33.9	21.1	0.002				
	Females	40.5	25.7	11.9	0.000	40.0	26.1	11.6	0.000	34.7	26.2	21.4	0.061	37.1	25.9	14.3	0.001				
	≤50 years	21.5	27.4	27.3	0.780	22.7	25.9	28.8	0.722	32.5	25.9	28.6	0.649	22.7	28.2	23.1	0.438				
	>50 years	46.9	32.5	–	0.004	44.3	33.8	–	0.034	43.6	34.5	–	0.060	44.6	33.5	–	0.025				
	M ≤ 50 years	17.8	16.9	17.5	0.979	14.3	16.4	17.7	0.890	19.3	16.3	19.8	0.673	17.3	17.5	14.1	0.710				
	M > 50 years	25.5	18.4	–	0.057	22.9	19.9	–	0.417	23.9	18.9	–	0.164	23.2	19.6	–	0.321				
	F ≤ 50 years	4.0	12.2	9.4	0.251	9.0	11.2	9.9	0.833	14.0	10.1	10.6	0.778	6.0	11.6	8.9	0.379				
	F > 50 years	23.0	13.8	–	0.007	22.9	13.8	–	0.007	20.3	16.1	–	0.195	22.2	14.5	–	0.024				
	All	71.7	75.7	42.3	0.000	66.2	79.4	37.6	0.000	71.5	72.5	46.7	0.001	67.9	74.4	44.9	0.001				
L	Males	46.8	40.4	27.4	0.009	41.4	44.8	23.6	0.001	45.8	39.9	27.6	0.011	42.5	43.3	26.0	0.006				
	Females	26.4	33.3	15.4	0.003	25.5	33.8	14.9	0.001	24.3	32.4	19.8	0.029	26.4	30.8	19.6	0.092				
	≤50 years	28.5	21.8	28.5	0.337	32.1	28.1	23.6	0.193	26.3	33.5	24.6	0.177	29.8	25.5	25.7	0.655				
	>50 years	43.8	39.4	22.8	0.169	33.5	41.7	34.5	0.315	44.3	38.4	23.2	0.124	37.1	40.1	36.2	0.828				
	M ≤ 50 years	17.8	11.9	18.9	0.143	18.3	17.5	15.6	0.758	20.2	19.2	15.1	0.344	21.1	15.3	16.4	0.411				
	M > 50 years	25.1	21.4	10.0	0.139	19.3	22.9	22.0	0.654	24.8	20.2	14.2	0.248	20.0	23.0	18.0	0.648				
	F ≤ 50 years	10.7	11.6	10.0	0.868	13.2	10.9	8.8	0.278	8.0	15.4	10.1	0.096	10.1	10.6	9.7	0.951				
	F > 50 years	24.8	17.2	12.8	0.225	16.7	18.7	11.5	0.583	20.0	18.3	9.3	0.383	19.6	17.6	19.0	0.913				
	All	112.8	78.4	45.1	0.000	113.3	79.4	40.5	0.000	98.0	70.1	55.9	0.006	118.3	76.4	44.6	0.000				
	BC	Males	63.5	44.5	28.0	0.001	64.5	44.9	24.1	0.000	66.3	39.8	32.6	0.018	70.0	44.7	26.1	0.000			
Females		50.8	35.1	16.8	0.000	49.7	36.0	15.9	0.000	34.5	30.9	23.8	0.173	49.0	32.4	18.7	0.000				
≤50 years		–	18.3	29.0	0.028	–	24.9	26.9	0.689	–	24.5	27.6	0.534	–	29.9	25.7	0.375				
>50 years		60.6	40.6	23.3	0.003	61.8	40.3	19.6	0.000	53.1	38.6	33.7	0.184	66.6	38.7	26.6	0.001				
M ≤ 50 years		–	13.4	18.3	0.138	–	16.5	16.5	1.000	–	16.1	17.3	0.725	–	18.9	16.3	0.422				
M > 50 years		31.2	23.0	13.2	0.065	32.8	22.5	10.3	0.017	35.5	21.1	16.6	0.066	37.0	21.4	15.3	0.029				
F ≤ 50 years		–	4.0	10.8	0.217	–	16.5	10.2	0.252	–	5.0	10.8	0.292	–	13.5	9.8	0.473				
F > 50 years		30.8	18.0	10.4	0.022	29.8	18.3	9.4	0.018	19.2	17.9	17.7	0.975	30.2	17.8	11.9	0.043				

Abbreviations: 1 = underaged; 2 = correctly aged; 3 = overaged; BC, Buckberry and Chamberlain (2002); EC, enthesal changes; L, Lovejoy et al. (1985); SB, Brooks and Suchey (1990).

of the individuals under study. With regard to pathological conditions, we excluded from the dataset individuals with visible signs of conditions that may affect the rate of degeneration of the pelvic joints, such as ankylosing spondylitis, Reiter's syndrome, rheumatoid arthritis, tuberculosis and trauma (see Campanacho et al., 2012, for similar approach). Nonetheless, we cannot exclude the possible effect of other, less easily identifiable pathological conditions on the os coxae.

Despite the above possible limitations, the current study highlights that the effect of activity on the degeneration of the pelvic joints is not as pronounced as we would anticipate, which is in agreement to a number of earlier studies, even though our materials and methods differ from those of these earlier studies. Considering the difficulty in identifying past occupations skeletally (Jurmain, Alves Cardoso, Henderson, & Villotte, 2012), our results are important as they support that this often missing parameter in past osteobiographies does not significantly affect skeletal age. However, it should be stressed that these results are specific to the Athens Collection, and there could be differences in other populations making it worthwhile duplicating this work on different samples. Moreover, our results suggest that taking into account the EC scores when evaluating skeletal age may provide further insights in age-at-death estimation in older individuals in the sense that older individuals are often underaged using skeletal age markers. In such cases, an observation of pronounced EC may point to a higher than estimated age-at-death for the individual in question, and there have in fact been recent attempts to use EC as age-at-death estimators by means of regression analysis (Milella et al., 2020). Of course, this requires additional work on a larger and broader sample to test its validity and, most importantly, to calculate a statistically appropriate bias correction model.

ACKNOWLEDGEMENTS

The authors would like to thank two anonymous reviewers for their valuable comments, which greatly improved this manuscript. Efthymia Nikita's contribution to this research was supported by the European Union's Horizon 2020 research and innovation programme under grant agreement No 811068 and the European Regional Development Fund and the Republic of Cyprus through the Research and Innovation Foundation (People in Motion project: EXCELLENCE/1216/0023).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ORCID

Andreas Bertatos  <https://orcid.org/0000-0001-8734-5236>

Maria-Eleni Chovalopoulou  <https://orcid.org/0000-0001-8187-8031>

Efthymia Nikita  <https://orcid.org/0000-0003-2094-5047>

REFERENCES

Auerbach, B. M., & Ruff, C. B. (2004). Human body mass estimation: A comparison of "morphometric" and "mechanical" methods. *American*

Journal of Physical Anthropology, 125, 331–342. <https://doi.org/10.1002/ajpa.20032>

Becker, I., Woodley, S. J., & Stringer, M. D. (2010). The adult human pubic symphysis: A systematic review. *Journal of Anatomy*, 217, 475–487. <https://doi.org/10.1111/j.1469-7580.2010.01300.x>

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. <https://doi.org/10.1136/ard.2008.091074>

Berg, G. E. (2008). Pubic bone age estimation in adult women. *Journal of Forensic Sciences*, 53, 569–577. <https://doi.org/10.1111/j.1556-4029.2008.00712.x>

Bertsatos, A., & Chovalopoulou, M. E. (2019). A novel method for analyzing long bone diaphyseal cross-sectional geometry. A GNU octave CSG toolkit. *Forensic Science International*, 297, 65–71. <https://doi.org/10.1016/j.forsciint.2019.01.041>

Bertsatos, A., Gkaniatsou, E., Papageorgopoulou, C., & Chovalopoulou, M. (2019). "What and how should we share?" An inter-method inter-observer comparison of measurement error with landmark-based craniometric datasets. *Anthropologischer Anzeiger*, 77, 109–120. <https://doi.org/10.1127/anthranz/2019/1047>

Bocquet-Appel, J. P., & Masset, C. (1982). Farewell to paleodemography. *Journal of Human Evolution*, 11, 321–333. [https://doi.org/10.1016/S0047-2484\(82\)80023-7](https://doi.org/10.1016/S0047-2484(82)80023-7)

Boldsen, J. L., Milner, G. R., Konigsberg, L. W., & Wood, J. W. (2002). Transition analysis: A new method for estimating age from skeletons. In R. D. Hoppa, & J. Vaupel (Eds.), *Paleodemography: Age distributions from skeletal samples* (pp. 73–106). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511542428.005>

Brooks, S., & Suchey, J. M. (1990). Skeletal age determination based on the os pubis: A comparison of the Acsádi-Nemeskéri and Suchey-Brooks methods. *Human Evolution*, 5, 227–238. <https://doi.org/10.1007/BF02437238>

Buckberry, J. L., & Chamberlain, A. T. (2002). Age estimation from the auricular surface of the ilium: A revised method. *American Journal of Physical Anthropology*, 119, 231–239. <https://doi.org/10.1002/ajpa.10130>

Budak, M. J., & Oliver, T. B. (2013). There's a hole in my symphysis—A review of disorders causing widening, erosion, and destruction of the symphysis pubis. *Clinical Radiology*, 68, 173–180. <https://doi.org/10.1016/j.crad.2012.03.021>

Campanacho, V., Santos, A. L., & Cardoso, H. F. (2012). Assessing the influence of occupational and physical activity on the rate of degenerative change of the pubic symphysis in Portuguese males from the 19th to 20th century. *American Journal of Physical Anthropology*, 148, 371–378. <https://doi.org/10.1002/ajpa.22059>

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. <https://doi.org/10.1002/oa.2285>

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. <https://doi.org/10.1002/ajpa.21171>

Chen, X., Zhang, Z., & Tao, L. (2008). Determination of male age at death in Chinese Han population: Using quantitative variables statistical analysis from pubic bones. *Forensic Science International*, 175, 36–43. <https://doi.org/10.1016/j.forsciint.2007.04.231>

Christensen, A. M., Passalacqua, N. V., & Bartelink, E. J. (2014). *Forensic anthropology: Current methods and practice*. San Diego: Academic Press.

Cowie, R., Bekvalac, J., & Kausmally, T. (2008). Late 17th- to 19th-century burial and earlier occupation at All Saints, Chelsea old church, Royal Borough of Kensington and Chelsea. In *MoLAS archaeology studies series 18*. London: Museum of London Archaeology Service.

- Cunningham, P. M., Brennan, D., O'Connell, M., MacMahon, P., O'Neill, P., & Eustace, S. (2007). Patterns of bone and soft-tissue injury at the symphysis pubis in soccer players: Observations at MRI. *American Journal of Roentgenology*, 188, W291–W296. <https://doi.org/10.2214/AJR.06.0051>
- Damen, L., Buyruk, H. M., Güler-Uysal, F., Lotgering, F. K., Snijders, C. J., & Stam, H. J. (2001). Pelvic pain during pregnancy is associated with asymmetric laxity of the sacroiliac joints. *Acta Obstetrica et Gynecologica Scandinavica*, 80, 1019–1024. <https://doi.org/10.1034/j.1600-0412.2001.801109.x>
- Dehghan, F., Haerian, B. S., Muniandy, S., Yusof, A., Dragoo, J. L., & Salleh, N. (2014). The effect of relaxin on the musculoskeletal system. *Scandinavian Journal of Medicine & Science in Sports*, 24, e220–e229. <https://doi.org/10.1111/sms.12149>
- Eliopoulos, C., Lagia, A., & Manolis, S. (2007). A modern, documented human skeletal collection from Greece. *Homo*, 58, 221–228. <https://doi.org/10.1016/j.jchb.2006.10.003>
- Falys, C. G., Schutkowski, H., & Weston, D. A. (2006). Auricular surface aging: Worse than expected? A test of the revised method on a documented historic skeletal assemblage. *American Journal of Physical Anthropology*, 130, 508–513. <https://doi.org/10.1002/ajpa.20382>
- Gardner, W. U. (1936). Sexual dimorphism of the pelvis of the mouse, the effect of estrogenic hormones upon the pelvis and upon the development of scrotal hernias. *American Journal of Anatomy*, 59, 459–483. <https://doi.org/10.1002/aja.1000590307>
- Gilbert, B. M., & McKern, T. W. (1973). A method for aging the female os pubis. *American Journal of Physical Anthropology*, 38, 31–38. <https://doi.org/10.1002/ajpa.1330380109>
- Gilleard, W., & Smith, T. (2007). Effect of obesity on posture and hip joint moments during a standing task, and trunk forward flexion motion. *International Journal of Obesity*, 31, 267–271. <https://doi.org/10.1038/sj.ijo.0803430>
- Hall, K. (1950). The effect of oestrone and progesterone on the histological structure of the symphysis pubis of the castrated female mouse. *Journal of Endocrinology*, 7, 54–63. <https://doi.org/10.1677/joe.0.0070054>
- Hanihara, K., & Suzuki, T. (1978). Estimation of age from the pubic symphysis by means of multiple regression analysis. *American Journal of Physical Anthropology*, 48, 233–239. <https://doi.org/10.1002/ajpa.1330480218>
- Hartnett, K. M. (2010). Analysis of age-at-death estimation using data from a new, modern autopsy sample—Part I: Pubic bone. *Journal of Forensic Sciences*, 55, 1145–1151. <https://doi.org/10.1111/j.1556-4029.2010.01399.x>
- 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 enthesal changes. *International Journal of Osteoarchaeology*, 26, 925–932. <https://doi.org/10.1002/oa.2477>
- Igarashi, Y., Uesu, K., Wakebe, T., & Kanazawa, E. (2005). New method for estimation of adult skeletal age at death from the morphology of the auricular surface of the ilium. *American Journal of Physical Anthropology*, 128, 324–339. <https://doi.org/10.1002/ajpa.20081>
- Jeanson, A. L., Santos, F., Villa, C., Dupej, J., Lynnerup, N., & Brůžek, J. (2017). Body mass estimation from the skeleton: An evaluation of 11 methods. *Forensic Science International*, 281, 183.e1–183.e8. <https://doi.org/10.1016/j.forsciint.2017.10.026>
- Ji, M. X., & Yu, Q. (2015). Primary osteoporosis in postmenopausal women. *Chronic Diseases and Translational Medicine*, 1, 9–13. <https://doi.org/10.1016/j.cdtm.2015.02.006>
- Jurmain, R., Alves Cardoso, F., Henderson, C., & Villotte, S. (2012). Bioarchaeology's Holy Grail: the reconstruction of activity. In A. Grauer (Ed.), *A companion to paleopathology* (pp. 531–552). New York: Blackwell Publishing. <https://doi.org/10.1002/9781444345940.ch29>
- Karakostis, F. A., Jeffery, N., & Harvati, K. (2019). Experimental proof that multivariate patterns among muscle attachments (entheses) can reflect repetitive muscle use. *Scientific Reports*, 9, 1–9. <https://doi.org/10.1038/s41598-019-53021-8>
- Karakostis FA, Wallace IJ, Konow N, Harvati K. 2019. Experimental evidence that physical activity affects the multivariate associations among muscle attachments (entheses). *Journal of Experimental Biology* 222: jeb.213058. <https://doi.org/10.1242/jeb.213058>
- Katz, D., & Suchey, J. M. (1986). Age determination of the male os pubis. *American Journal of Physical Anthropology*, 69, 427–435. <https://doi.org/10.1002/ajpa.1330690402>
- Keriakos, R., Bhatta, S. C., Morris, F., Mason, S., & Buckley, S. (2011). Pelvic girdle pain during pregnancy and puerperium. *Journal of Obstetrics and Gynaecology*, 31, 572–580. <https://doi.org/10.3109/01443615.2011.598970>
- Kimmerle, E. H., Konigsberg, L. W., Jantz, R. L., & Baraybar, J. P. (2008). Analysis of age-at-death estimation through the use of pubic symphyseal data. *Journal of Forensic Sciences*, 53, 558–568. <https://doi.org/10.1111/j.1556-4029.2008.00711.x>
- Klales, A. R., Ousley, S. D., & Vollner, J. M. (2012). A revised method of sexing the human innominate using Phenice's nonmetric traits and statistical methods. *American Journal of Physical Anthropology*, 149, 104–114. <https://doi.org/10.1002/ajpa.22102>
- Konigsberg, L. W., & Frankenberg, S. R. (1992). Estimation of age structure in anthropological demography. *American Journal of Physical Anthropology*, 89, 235–256. <https://doi.org/10.1002/ajpa.1330890208>
- Larsen, C. S. (2015). *Bioarchaeology: Interpreting behavior from the human skeleton*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139020398>
- Lentz, S. S. (1995). Osteitis pubis: A review. *Obstetrical & Gynecological Survey*, 50, 310–315. <https://doi.org/10.1097/00006254-199504000-00026>
- Lottering, N., Reynolds, M. S., MacGregor, D. M., Meredith, M., & Gregory, L. S. (2014). Morphometric modelling of ageing in the human pubic symphysis: Sexual dimorphism in an Australian population. *Forensic Science International*, 236, 195.e1–e11. <https://doi.org/10.1016/j.forsciint.2013.12.041>
- Lovejoy, C. O., Meindl, R. S., Pryzbeck, T. R., & Mensforth, R. P. (1985). Chronological metamorphosis of the auricular surface of the ilium: A new method for the determination of adult skeletal age at death. *American Journal of Physical Anthropology*, 68, 15–28. <https://doi.org/10.1002/ajpa.1330680103>
- Mariotti, V., Facchini, F., & Belcastro, M. (2004). Enthesopathies—Proposal of a standardized scoring method and applications. *Collegium Antropologicum*, 28, 145–159.
- 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, 291–313.
- Marnach, M. L., Ramin, K. D., Ramsey, P. S., Song, S. W., Stensland, J. J., & An, K. N. (2003). Characterization of the relationship between joint laxity and maternal hormones in pregnancy. *Obstetrics & Gynecology*, 101, 331–335. [https://doi.org/10.1016/s0029-7844\(02\)02447-x](https://doi.org/10.1016/s0029-7844(02)02447-x)
- Mays, S. (2012). An investigation of age-related changes at the acetabulum in 18th–19th century AD adult skeletons from Christ Church Spitalfields, London. *American Journal of Physical Anthropology*, 149, 485–492. <https://doi.org/10.1002/ajpa.22146>
- Mays, S. (2015). The effect of factors other than age upon skeletal age indicators in the adult. *Annals of Human Biology*, 42, 332–341. <https://doi.org/10.3109/03014460.2015.1044470>
- Mays, S. (2016). Bone-formers and bone-losers in an archaeological population. *American Journal of Physical Anthropology*, 159, 577–584. <https://doi.org/10.1002/ajpa.22912>
- McKern, T. W., & Stewart, T. D. (1957). *Skeletal age changes in young American males: Analysed from the standpoint of age identification (technical*

- report EP-45). Quartermaster Research and Development Command: Natick, MA.
- Meindl, R. S., Lovejoy, C. O., Mensforth, R. P., & Walker, R. A. (1985). A revised method of age determination using the os pubis, with a review and tests of accuracy of other current methods of pubic symphyseal aging. *American Journal of Physical Anthropology*, *68*, 29–45. <https://doi.org/10.1002/ajpa.1330680104>
- Merritt, C. E. (2015). The influence of body size on adult skeletal age estimation methods. *American Journal of Physical Anthropology*, *156*, 35–57. <https://doi.org/10.1002/ajpa.22626>
- Michopoulou, E., Nikita, E., & Valakos, E. D. (2015). Evaluating the efficiency of different recording protocols for enthesal changes in regards to expressing activity patterns using archival data and cross-sectional geometric properties. *American Journal of Physical Anthropology*, *158*, 557–568. <https://doi.org/10.1002/ajpa.22822>
- Milella, M., Belcastro, G. M., Zollikofer, C. P., & Mariotti, V. (2012). The effect of age, sex, and physical activity on enthesal morphology in a contemporary Italian skeletal collection. *American Journal of Physical Anthropology*, *148*, 379–388. <https://doi.org/10.1002/ajpa.22060>
- Milella, M., Belcastro, M. G., Mariotti, V., & Nikita, E. (2020). Estimation of adult age-at-death from enthesal robusticity: A test using an identified Italian skeletal collection. *American Journal of Physical Anthropology*, *173*, 190–199. <https://doi.org/10.1002/ajpa.24083>
- Mitra, R. (2010). Osteitis condensans ilii. *Rheumatology International*, *30*, 293–296. <https://doi.org/10.1007/s00296-009-1100-7>
- Nawrocki, S. P. (2010). The nature and sources of error in the estimation of age at death from the skeleton. In K. E. Latham, & M. Finnegan (Eds.), *Age estimation of the human skeleton* (pp. 79–101). Springfield: Charles C Thomas.
- Niinimäki, S. (2011). What do muscle marker ruggedness scores actually tell us? *International Journal of Osteoarchaeology*, *21*, 292–299. <https://doi.org/10.1002/oa.1134>
- Nikita, E. (2017). *Osteoarchaeology: A guide to the macroscopic study of human skeletal remains*. San Diego: Academic Press.
- Nikita, E. (2020). Documented skeletal collections in Greece: Composition, research, and future prospects. *American Journal of Physical Anthropology*. <https://doi.org/10.1002/ajpa.24050>
- 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. <https://doi.org/10.1016/j.jchb.2017.11.002>
- Nikita, E., & Nikitas, P. (2019). Skeletal age-at-death estimation: Bayesian versus regression methods. *Forensic Science International*, *297*, 56–64. <https://doi.org/10.1016/j.forsciint.2019.01.033>
- Nikita, E., Xanthopoulou, P., Bertatos, A., Chovalopoulou, M. E., & Hafez, I. (2019). A three-dimensional digital microscopic investigation of enthesal changes as skeletal activity markers. *American Journal of Physical Anthropology*, *169*, 704–713. <https://doi.org/10.1002/ajpa.23850>
- Osborne, D. L., Simmons, T. L., & Nawrocki, S. P. (2004). Reconsidering the auricular surface as an indicator of age at death. *Journal of Forensic Sciences*, *49*, 905–911.
- Pottie, P., Presle, N., Terlain, B., Netter, P., Mainard, D., & Berenbaum, F. (2006). Obesity and osteoarthritis: More complex than predicted! *Annals of the Rheumatic Diseases*, *65*, 1403–1405. <https://doi.org/10.1136/ard.2006.061994>
- Prince, D. A., Kimmerle, E. H., & Konigsberg, L. W. (2008). A Bayesian approach to estimate skeletal age-at-death utilizing dental wear. *Journal of Forensic Sciences*, *53*, 588–593. <https://doi.org/10.1111/j.1556-4029.2008.00714.x>
- Richmond, S. A., Fukuchi, R. K., Ezzat, A., Schneider, K., Schneider, G., & Emery, C. A. (2013). Are joint injury, sport activity, physical activity, obesity, or occupational activities predictors for osteoarthritis? A systematic review. *Journal of Orthopaedic & Sports Physical Therapy*, *43*, 515–519. <https://doi.org/10.2519/jospt.2013.4796>
- Rissech, C., Wilson, J., Winburn, A. P., Turbón, D., & Steadman, D. (2012). A comparison of three established age estimation methods on an adult Spanish sample. *International Journal of Legal Medicine*, *126*, 145–155. <https://doi.org/10.1007/s00414-011-0586-1>
- Robb, J. E. (1998). The interpretation of skeletal muscle sites: A statistical approach. *International Journal of Osteoarchaeology*, *8*, 363–377. [https://doi.org/10.1002/\(SICI\)1099-1212\(1998090\)8:5<363::AID-OA438>3.0.CO;2-K](https://doi.org/10.1002/(SICI)1099-1212(1998090)8:5<363::AID-OA438>3.0.CO;2-K), [https://doi.org/10.1002/\(SICI\)1099-1212\(1998090\)8:5<363::AID-OA438>3.0.CO;2-K](https://doi.org/10.1002/(SICI)1099-1212(1998090)8:5<363::AID-OA438>3.0.CO;2-K)
- Rogers, J., Shepstone, L., & Dieppe, P. (1997). Bone formers: Osteophyte and enthesophyte formation are positively associated. *Annals of the Rheumatic Diseases*, *56*, 85–90. <https://doi.org/10.1136/ard.56.2.85>
- Ruff, C. (1987). Sexual dimorphism in human lower limb bone structure: Relationship to subsistence strategy and sexual division of labor. *Journal of Human Evolution*, *16*, 391–416. [https://doi.org/10.1016/0047-2484\(87\)90069-8](https://doi.org/10.1016/0047-2484(87)90069-8)
- Ruff, C., Holt, B. M., Niskanen, M., Sladěk, V., Berner, M., Garofalo, E., ... Tompkins, D. (2012). Stature and body mass estimation from skeletal remains in the European Holocene. *American Journal of Physical Anthropology*, *148*, 601–617. <https://doi.org/10.1002/ajpa.22087>
- Ruff, C. B. (2008). Biomechanical analyses of archaeological human skeletons. In M. A. Katzenberg, & S. R. Saunders (Eds.), *Biological anthropology of the human skeleton* (pp. 183–206). New York: Wiley Liss. <https://doi.org/10.1002/9780470245842.ch6>
- Ruff, C. B. (2010). Body size and body shape in early hominins: Implications of the Gona pelvis. *Journal of Human Evolution*, *58*, 166–178. <https://doi.org/10.1016/j.jhevol.2009.10.003>
- Ruff, C. B., & Hayes, W. C. (1983). Cross-sectional geometry of Pecos Pueblo femora and tibiae—A biomechanical investigation: II. Sex, age, and side differences. *American Journal of Physical Anthropology*, *60*, 383–400. <https://doi.org/10.1002/ajpa.1330600309>
- Sahni, S., Zoltick, E. S., McLean, R. R., & Hannan, M. T. (2010). Non-D vitamins and bone health in adults. *IBMS BoneKey*, *7*, 431–446. <https://doi.org/10.3390/nu2070693>
- Shibata, Y., Shirai, Y., & Miyamoto, M. (2002). The aging process in the sacroiliac joint: Helical computed tomography analysis. *Journal of Orthopaedic Science*, *7*, 12–18. <https://doi.org/10.1007/s776-002-8407-1>
- Sowers, M. (2001). Epidemiology of risk factors for osteoarthritis: Systemic factors. *Current Opinion in Rheumatology*, *13*, 447–451. <https://doi.org/10.1097/00002281-200109000-00018>
- Sparacello, V. S., & Pearson, O. M. (2010). The importance of accounting for the area of the medullary cavity in cross-sectional geometry: A test based on the femoral midshaft. *American Journal of Physical Anthropology*, *143*, 612–624. <https://doi.org/10.1002/ajpa.21361>
- Stock, J. T., & Shaw, C. N. (2007). Which measures of diaphyseal robusticity are robust? A comparison of external methods of quantifying the strength of long bone diaphyses to cross-sectional geometric properties. *American Journal of Physical Anthropology*, *134*, 412–423. <https://doi.org/10.1002/ajpa.20686>
- Tague, R. G. (1988). Bone resorption of the pubis and preauricular area in humans and nonhuman mammals. *American Journal of Physical Anthropology*, *76*, 251–267. <https://doi.org/10.1002/ajpa.1330760214>
- Todd, T. W. (1920). Age changes in the pubic bone. I. The male white pubis. *American Journal of Physical Anthropology*, *3*, 285–334. <https://doi.org/10.1002/ajpa.1330030301>
- Todd, T. W. (1921a). Age changes in the pubic bone. II. The pubis of the male negro-white hybrid, iii. the pubis of the white female, iv. the pubis of the female negro-white hybrid. *American Journal of Physical Anthropology*, *4*, 1–70. <https://doi.org/10.1002/ajpa.1330040102>
- Todd, T. W. (1921b). Age changes in the pubic bone. VI. The Interpretation of Variations in the Symphyseal Area. *American Journal of Physical Anthropology*, *4*, 407–429. <https://doi.org/10.1002/ajpa.1330040403>

- Villotte, S. (2009). *Enthesopathies et activités des homes préhistoriques- recherche méthodologique et application aux fossils Européens du Paléolithique Supérieur et du Mésolithique*. Oxford: Archaeopress.
- Villotte, S., & Knüsel, C. J. (2013). Understanding enthesal changes: Definition and life course changes. *International Journal of Osteoarchaeology*, 23, 135–146. <https://doi.org/10.1002/oa.2289>
- Vix, V., & Ryu, C. Y. (1971). The adult symphysis pubis: Normal and abnormal. *American Journal of Roentgenology*, 112, 517–525. <https://doi.org/10.2214/ajr.112.3.517>
- Walker, P. L. (2008). Sexing skulls using discriminant function analysis of visually assessed traits. *American Journal of Physical Anthropology*, 136, 39–50. <https://doi.org/10.1002/ajpa.20776>
- Weiss, E. (2003). Understanding muscle markers: Aggregation and construct validity. *American Journal of Physical Anthropology*, 121, 230–240. <https://doi.org/10.1002/ajpa.10226>
- Weiss, E. (2004). Understanding muscle markers: Lower limbs. *American Journal of Physical Anthropology*, 125, 232–238. <https://doi.org/10.1002/ajpa.10397>
- 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. <https://doi.org/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. <https://doi.org/10.1002/oa.1183>
- Wescott, D. J., & Drew, J. L. (2015). Effect of obesity on the reliability of age-at-death indicators of the pelvis. *American Journal of Physical Anthropology*, 156, 595–605. <https://doi.org/10.1002/ajpa.22674>
- 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. [https://doi.org/10.1002/\(SICI\)1099-1212\(199809\)8:5<311::AID-OA443>3.0.CO;2-E](https://doi.org/10.1002/(SICI)1099-1212(199809)8:5<311::AID-OA443>3.0.CO;2-E), [https://doi.org/10.1002/\(SICI\)1099-1212\(199809\)8:5<311::AID-OA443>3.0.CO;2-E](https://doi.org/10.1002/(SICI)1099-1212(199809)8:5<311::AID-OA443>3.0.CO;2-E)
- Xanthopoulou, P., Valakos, E., Youlatos, D., & Nikita, E. (2018). Assessing the accuracy of cranial and pelvic ageing methods on human skeletal remains from a modern Greek assemblage. *Forensic Science International*, 286, 266.e1–266.e8. <https://doi.org/10.1016/j.forsciint.2018.03.005>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Bertsatos A, Chovalopoulou M-E, Boskovits N-M, Garoufi N, Nikita E. The impact of activity on pelvic age-at-death estimation. *Int J Osteoarchaeol*. 2020; 1–14. <https://doi.org/10.1002/oa.2941>



Advanced procedures for skull sex estimation using sexually dimorphic morphometric features

Andreas Bertsatos¹ · Maria-Eleni Chovalopoulou² · Jaroslav Brůžek³ · Šárka Bejdová³

Received: 28 February 2020 / Accepted: 29 May 2020
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

Abstract

This paper introduces an automated method for estimating sex from cranial sex diagnostic traits by extracting and evaluating specialized morphometric features from the glabella, the supraorbital ridge, the occipital protuberance, and the mastoid process. The proposed method was developed and evaluated using two European population samples, a Czech sample comprising 170 crania reconstructed from anonymized CT scans and a Greek sample of 156 crania from the Athens Collection. It is based on a fully automatic algorithm applied on 3D models for extracting sex diagnostic morphometric features which are further processed by computer vision and machine learning algorithms. Classification accuracy was evaluated in a population specific and a population generic 2-way cross-validation scheme. Population-specific accuracy for individual morphometric features ranged from 78.5 to 96.7%, whereas population generic correct classification ranged from 71.7 to 90.8%. Combining all sex diagnostic traits in multi-feature sex estimation yielded correct classification performance in excess of 91% for the entire sample, whereas the sex of about three fourths of the sample could be determined with 100% accuracy according to posterior probability estimates. The proposed method provides an efficient and reliable way to estimate sex from cranial remains, and it offers significant advantages over existing methods. The proposed method can be readily implemented with the *skullanalyzer* computer program and the *estimate_sex.m* GNU Octave function, which are freely available under a suitable license.

Keywords Morphometric feature extraction · Cranial sex diagnostic traits · Supervised machine learning · Sex discriminant models · European populations

Introduction

Sex estimation is a primary task when dealing with human skeletal remains. The most reliable skeletal element for sex estimation is considered the hip bone since its morphological

adaptation to facilitate childbirth allows for the highest accuracy levels along with low population specificity [1]. The cranium also exhibits significant sexual dimorphism, which can facilitate accurate sex estimation [2, 3]. In contrast to the pelvic bone, the main drawback of the cranium is that its expression of sexual dimorphism is population specific [4, 5]. Hence, some debate as to whether the cranium should be considered the second best bone structure for estimating sex as opposed to other postcranial bones [6, 7]. Nevertheless, the cranium is more resistant to damage and usually better maintained than the postcranial skeleton [7]. Thus, it is common that sex identification may rely on the cranium morphology if the pelvic bones are missing or highly deteriorated. This is reflected not only by the vast literature in assessing cranial sexual dimorphism and cranial sex estimation techniques but also by the continuous effort to improve upon existing methods and further increase their accuracy and reliability.

Cranial sex assessment is usually performed either by visual observation or a variety of morphometric methods. Visual observation relies on the subjective evaluation of sexual

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00414-020-02334-9>) contains supplementary material, which is available to authorized users.

✉ Andreas Bertsatos
abertsatos@biol.uoa.gr

¹ Department of Animal and Human Physiology, Faculty of Biology, School of Sciences, University of Athens, Panepistimiopolis, GR 157 01 Athens, Greece

² Science and Technology in Archaeology and Culture Research Center, The Cyprus Institute, 2121 Aglantzia, Nicosia, Cyprus

³ Department of Anthropology and Human Genetics, Faculty of Science, Charles University, Viničná 7, 128 44 Prague 2, Czech Republic

dimorphic areas of the cranium, such as the glabella or the mastoid process, which are empirically scored according to standardized categorical typologies and thereafter used to provide an assessment [4, 8–11]. On the other hand, morphometric methods may involve a number of measurement techniques such as standard caliper-based measurements [12–14], landmark-based geometric morphometrics [15–19], or even surface analysis of 3D mesh models focusing either on particular sexually dimorphic traits [20–23] or on the entire cranium [24, 25]. Considering the benefits and shortcomings of each method summarizes to that visual observation although easier and quicker to apply tends to rely on the observer's experience and has been criticized for high subjectivity and poor reproducibility [8, 26, 27]. On the contrary, morphometric methods are more reproducible and often more accurate [28, 29], but they are more elaborate and often fail to adequately capture subtle sexually dimorphic variations [4]. Moreover, their resulted discriminant functions are inherently prone to population specificity [30, 31].

Merging the best aspects of both worlds seems intuitive. That is devising new sex estimation methods that bear the accuracy and reproducibility of the morphometric methods while maintaining the agility of a highly trained observer to visually assess the subtle morphological variations between sexes. As a matter of fact, over the past decade, several studies have proposed customized morphometric variables for specific sex diagnostic traits such as the mastoid process [32, 33] and the browridge [34]. More recently, Nikita and Michopoulou (2017) proposed a set of specialized discriminant variables (DVs) to quantify the shape of the glabella-frontal bone, the mastoid process, and the external occipital protuberance as outlines seen from a lateral image of the cranium. The proposed technique, which partially mimics visual assessment, yielded a correct sex classification accuracy of almost 94% on a contemporary Greek skeletal sample followed by the development of an R script that quantifies the shape of selected cranial traits and automates sex estimation [35], which considerably simplified the initially proposed method. Newer methods have set the bar even higher by adopting computer-aided measurements from 3D cranial models and elaborate machine learning algorithms. More specifically, Yang and colleagues [36], who studied a sample from the Uighur ethnic group in the north of China, they applied an improved back-propagation neural network on specific geometric variables measured on 3D cranial models derived from CT scans, and they reported 96.7% testing accuracy. In another work, Yang and colleagues [37] demonstrated the use of wavelet transform and Fourier transform on cranial sex estimation.

Following the recent innovations in cranial sex estimation and in light of the increasing availability of 3D skeletal samples and the maturity of 3D scanning technologies in osteological work, the present study proposes a novel method for

sex estimation from the cranium, which relies on a fully automated and precise morphometric feature extraction from 3D models. This is achieved with minimal user effort and integrated optimization routines ensure that measurement error is practically bound only by the accuracy of the 3D models themselves. The proposed method further introduces the concept of height map images as morphometric features, which may be regarded as an invertible transformation of the 3D cranial surface of specific sex diagnostic cranial areas to 2D grayscale images. The latter are subsequently processed by computer vision and machine learning algorithms. Additionally, the present study elaborately utilizes two distinct European modern population samples in a repeated mixed-population 2-way cross-validation approach to mitigate against sample over-fitting and population specificity in order to ensure that the reported sex estimation accuracy reflects the true capacity of the proposed method in unidentified real cases.

Materials and methods

Population sample

The sample utilized in the present study consists of 324 cranial 3D models representing two contemporary European populations. The Greek population sample consists of 154 models of dry crania from the Athens Collection, which is housed in the Department of Animal and Human Physiology (Faculty of Biology, National & Kapodistrian University of Athens). These crania belong to Greek adult individuals of known sex (79 males and 75 females), age at death (19–99 years old), cause of death as well as place and year of birth (born between 1879 and 1965) derived from death records [38]. Only individuals without any evidence of pathological deformation and absence of craniotomy were included. The Czech population sample consists of 170 models, which have been generated from previously anonymized DICOM files containing cranial CT scans of patients (102 males and 68 females, aged from 21 to 87 years old) from the Na Homolce Hospital in Prague, Czech Republic. The CT scans were captured with a CT Somatron Definition Flash 2 × 128 machine (Siemens, Erlangen Germany) with the following parameters: cut distance 0.6 mm; 0.4 kernel bone tissue H60 (reconstructive computational algorithm for bone tissue). All individuals gave their informed consent with the CT imaging and the use of their data for research purposes. The DICOM data used in the present paper were collected and anonymized by an experienced radiologist, Dr. Martin Horák, upon the approval of the ethics committee of the hospital and they are stored in the Laboratory of 3D Visualization and Analytical Methods of the Department of Anthropology and Human Genetics (Faculty of Science, Charles University, Prague).

3D model reconstruction

The 3D models of the Greek population sample had been previously constructed by means of 3D photogrammetry (Photoscan Pro v1.4, Agisoft LLC, Russia). These models, which also contain texture information, comprise 500,000 faces each and have a rated accuracy of ~ 0.24 mm [39]. The 3D models of the Czech population sample were initially converted from the respective DICOM files into Wavefront Alias OBJ format with the `dicom2mesh` (Adrian Schneider, AOT AG, Switzerland) command line program using an ISO threshold of 350 and a reduction coefficient of 0.8 along with centered and smoothing flags enabled. All converted models were post-processed in Meshlab [40] to remove all unnecessary elements (including the mandible and the cervical vertebrae) apart from the cranium. In order to eliminate holes and other artifacts due to noisy CT scans and produce a watertight model of the ectocranial surface, the cleaned cranial 3D models went through a custom-built algorithm by Carmelo Mineo, which was based on novel algorithms for 3D surface point cloud boundary detection and edge reconstruction [41–43]. The resulted triangular mesh 3D models, which lacked any texture information, comprised a variable number of faces ranging from 192k to 1820k. A quality assessment of the DICOM-derived 3D models was performed by repeated measurements on 20 randomly chosen samples. The basion-nasion length and the basion-bregma height were repeatedly measured on the iso-surface (ISO threshold of 350) generated from the DICOM files with Amira 5.3.3 (ZIB, Germany) and on the final post-processed watertight models with Meshlab. The mean absolute error of those repeated measurements was calculated to be 0.38 mm, which implies that the conversion of the DICOM files to 3D models and their subsequent post-processing is fit for craniometric purposes since it has negligible effect on the overall geometry of the cranium.

Morphometric data acquisition

All 324 models were analyzed with the “skullanalyzer” software [44], a command line program that automatically extracts morphometric features of the most sex diagnostic traits from cranial 3D models and is freely available at <https://github.com/pr0m1th3as/skullanalyzer> under the GNU General Public License v3. The skullanalyzer extracts height map images (HMIs) from the occipital protuberance, the supraorbital ridge, and the mastoid processes as well as the 2D projection of the *nasion-bregma* ectocranial segment for which the corresponding elliptic Fourier descriptors (EFDs) are automatically calculated. Figure 1 illustrates the correspondence between the HMI of the lateral view of the left mastoid process and the respective cranial 3D geometry, which has been manually aligned in Meshlab to approximate the same viewpoint. Detailed descriptions of all

mentioned morphometric features, and the underlying extraction algorithm are provided in the skullanalyzer’s user manual [45]. The tri-dimensional coordinates of the landmarks, which are required by the skullanalyzer to perform initial alignment and subsequent morphometric feature extraction, were captured from the 3D models with Meshlab. In cases of missing or damaged cranial traits (e.g., mastoid process) the respective landmarks were ignored and the respective features were not extracted. The skullanalyzer’s sequential analysis of the 324 cranial models from the two population samples produced 2557 files, which constitute the raw dataset of the morphometric features analyzed in the present study. More specifically, 1909 height map images (saved in .csv files) were extracted. Yet, another 324 polylines (also saved in .csv format) were generated, whose corresponding EFDs were saved along with the existing height map images in specialized data text format (saved in .mat files) compliant with the GNU Octave programming language [46]. The complete dataset of morphometric features from this study is available at Zenodo repository [47].

Data analysis

The design of the data analysis aims to evaluate the aforementioned sex diagnostic cranial traits in a population generic approach. The entire analysis, which includes sample manipulation, image analysis, supervised machine learning, results’ plotting, and statistical modeling for applied sex estimation, was performed with custom built functions in Python and GNU Octave programming languages. The entire code base for the data analysis implemented in this work is available at <https://github.com/pr0m1th3as/cranial-sex-estimation-code>. Note that the results presented in this paper might not be exactly reproducible due to the random permutations incorporated in the sample splitting. However, the present results are provided along with the code.

Data analysis—sample splitting

Each of the two population samples was split into a major and a minor subgroup, which are summarized in Table 1. The divisions were based on existing subdivisions of the available population samples. In the Greek sample, the WLH group concerns the first part of the Athens Collection, which was initially prepared in the Malcolm H. Wiener Laboratory for Archaeological Sciences of the American School of Classical Studies at Athens and later donated to the Department of Animal and Human Physiology of the University of Athens. The ABH group concerns the rest and biggest part of the skeletal collection, which was later prepared in the Department of Animal and Human Physiology. Accordingly, the Czech sample was subdivided into the MOTOL and HOMOLKA groups after the names of two

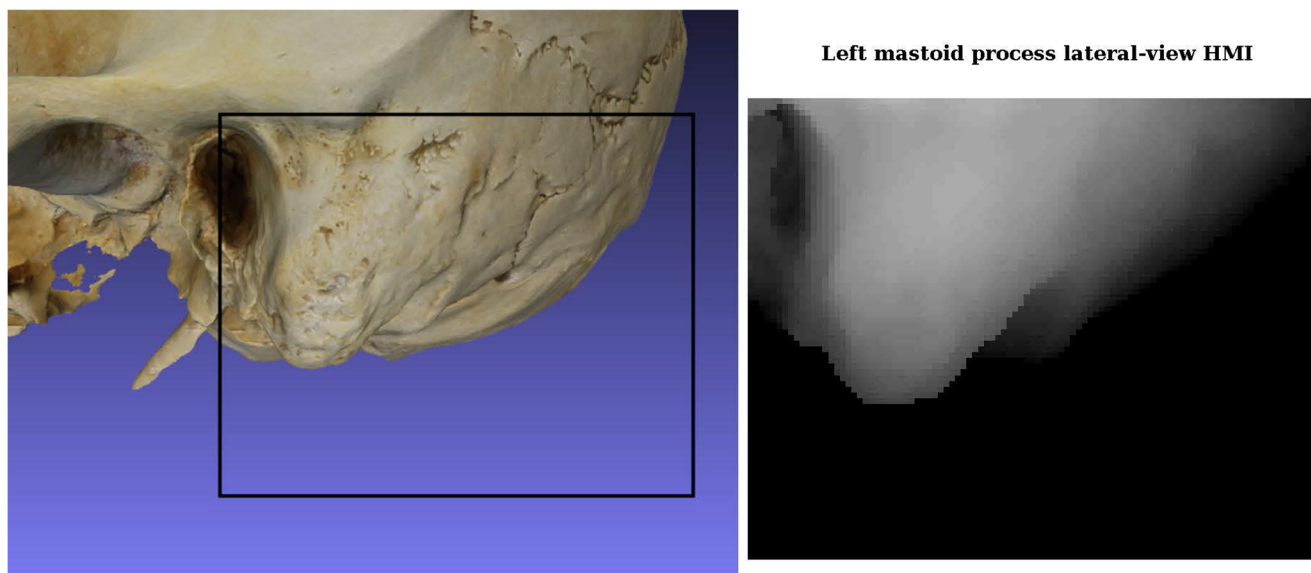


Fig. 1 Correspondence example between an HMI and its respective cranial 3D geometry

hospitals in Prague, Czech Republic (Na Homolce Hospital and Motol University Hospital). These hospitals, immediately adjoin to each, share the Department of Radiology, which is localized at Na Homolce Hospital and where the CT scans were captured. Individuals from both subgroups come partially from the Na Homolce Hospital and the Motol University Hospital. The HOMOLKA group was collected in years 2009–2011 and the MOTOL group in year 2017. These subdivisions were adopted for creating a large pooled sample for iteratively training and testing sex estimation statistical models, which would comprise primarily one population and the minor group of the second population, whereas the major group of the second population would be used for an additional cross-validation step for measuring the correct sex classification of each statistical model. This split approach was deemed necessary since the training and testing samples were selected with random permutation from the large pooled sample over 100 consecutive iterations, whereas the cross-

validation sample was kept the same but each time it was tested with a different statistical model. This iterative approach was performed twice, one with the Greek population primarily for training and the Czech major group exclusively for cross-validation and another with the two population samples inverted. Table 1 also summarizes the subgroup divisions in each iterative approach.

Data analysis—machine learning

Evaluation of the sex diagnostic capacity of each morphometric feature was performed with supervised machine learning, which utilized the training, testing, and cross-validation samples. The six distinct types of height map images, which correspond to different morphometric features, were initially processed with the histogram of oriented gradients (HOGs) feature extraction algorithm and subsequently fed to a support vector machine (SVM) classifier utilizing a radial basis

Table 1 Population samples' subdivision and subgroup pooling for iterative statistical modeling

Population sample	Group subdivision	Sample size	Males	Females
Greek (154) ^a	WLH	53	32	21
	ABH	101	47	54
Czech (170) ^a	MOTOL	67	44	23
	HOMOLKA	103	58	45
Sample splitting ^b	Training-testing pooled groups	Sample size	Cross-validation group	Sample size
#1	ABH, WLH, MOTOL	221	HOMOLKA	103
#2	HOMOLKA, WLH, MOTOL	223	ABH	101

^a Total sample size per population

^b Processed with 100 iterations

Table 2 Optimal parameters for supervised machine learning of each morphometric feature

Morphometric features	HOG parameters			RBF kernel parameters		Training sample size	Testing sample size ^a	Cross-validation sample size ^a
	orientations	Pixels per cell	Cells per block	C	γ			
Occipital protuberance HMI	8	(10,15)	(1,1)	20	1/432	161	60 / 62	103 / 101
Supraorbital ridge HMI	8	(10,12)	(1,1)	20	1/200	161	51 / 56	103 / 98
Left mastoid process inferior HMI	8	(10,10)	(1,1)	20	1/288	151	65 / 68	102 / 99
Left mastoid process lateral HMI	8	(10,10)	(1,1)	20	1/640	151	65 / 68	102 / 99
Right mastoid process inferior HMI	8	(10,10)	(1,1)	20	1/288	151	64 / 68	102 / 98
Right mastoid process lateral HMI	8	(10,10)	(1,1)	20	1/640	151	64 / 68	102 / 98
Nasion-bregma EFDs ^b	Supervised learning with LDA kernel					160	61 / 63	103 / 101

^a Testing and cross-validation sample sizes are given for both sample splits (#1/#2)

^b Utilized coefficients: a1, a2, a3, a4, a7, c2, c3

function (RBF) kernel ($C = 20$, $\gamma = 1/\text{number of HOG features}$). The optimized parameters used for the HOGs of each morphometric feature are given in Table 2 along with the respective sizes for training and testing samples. The final parameters for both the RBF kernel and the HOGs were optimized by trial and error. The EFDs of the *nasion-bregma* segment were initially analyzed by linear discriminant analysis (LDA), and only the seven most discriminatory coefficients (a1, a2, a3, a4, a7, c2, c3) were fed into the final LDA classifier.

Data analysis—plotting results

All seven classifiers were trained and tested on both sample splits (see Table 1) for 100 iterations. Each classifier was evaluated independently on the randomly chosen available testing samples as well as on the fixed cross-validation samples by comparing each sample's score of the respective classifier's decision function to each sample's referenced sex. Since RBF kernel-based SVM classifiers do not provide direct estimates of posterior probabilities, the correct classification was assessed at the nominal cut-off point (set to zero) of each classifier as well as with respect to different decision threshold ranges incremented by ± 0.1 up to ± 1.4 . In each case, the sample scores outside the respective decision threshold range were deemed as valid for classification, whereas the samples with scores inside the decision threshold range were rejected as unidentified. Correct classification was calculated as the percentage ratio of the correctly classified samples to those valid for classification. For every morphometric feature and each of the two sample splits, the correct classification and the sample rejection rates over the 100 iterations were summarized in box plots illustrating the relevant descriptive statistics for each decision threshold range separately.

Data analysis—combined models

After evaluating the performance of each trained classifier with the aforementioned iterative process, the model parameters of the best classifiers were further utilized in multi-feature sex estimation. All available morphometric features from each cranial sample (324 individuals) were evaluated with the respective classifiers, and subsequently, each sample was classified according to the predicted sex by a weighted sum of the applied decision functions on the available features. Once again, correct classification was evaluated following the same template of different decision threshold ranges. Initially, the testing-training pooled group and the cross-validation group from each split (Table 1) were evaluated separately by the respective best performing classifiers derived from each sample split, and their classification and rejection rates were recorded separately for each group. Subsequently, all 14 classifiers (2 for each feature, 7 from each sample split training group) were collectively applied to entire sample, and their respective weighted sum prediction was recorded for each population subgroup.

Results

Due to the abundance of classification results for each morphometric feature, the descriptive statistics regarding correct sex classification have been comprehensively summarized in the supplementary material (S1) as clusters of box plots. Four box plots illustrating classification accuracy and their counterparts showing the respective rejection rates at every decision threshold range were collated for every morphometric feature. Each pair corresponds to the distinct testing and cross-validation groups as well as the two sample splits utilized in the present study. Representative classification results are

selectively summarized in Tables 3 and 4 regarding the first and second sample split, respectively. These selected results concern the median and maximum correct classification outcomes over the 100 iterations of supervised learning on each sample split and their corresponding evaluation on testing and cross-validation groups at the nominal cutoff point (zero) as well as the ± 0.2 decision threshold range. Despite the box plots in S1 providing a more comprehensive visual assessment of the emerging sex discriminant patterns among the morphometric features, the results in Tables 3 and 4 are adequate to evaluate the sex discriminant capacity of each sexually dimorphic trait as well as the observed differences between the two population groups.

The least sex-discriminant morphometric features are the inferior HMIs of the mastoid processes followed by the occipital protuberance and the supraorbital ridge HMIs with the median classification scores at the nominal cutoff point ranging from 66.1 to 80.6%. On the other hand, the *nasion-bregma* EFDs constitute the best single morphometric feature sex estimator, with median scores at the nominal cutoff point between 82.5 and 85.2%, followed closely by the lateral HMIs of the mastoid processes with respective median scores ranging from 78.4 up to 86.7%. When considering the ± 0.2 offset in decision threshold boundaries, the correct classification scores are boosted by 2 up to 10 percentage units at the expense of rejection rates between 11 and 24% of the evaluated samples. The maximum classification scores, which correspond to the best trained classifiers, have yielded an eminent increase in the classification accuracy of every morphometric feature examined. The most notable classification score concerns the *nasion-bregma* EFDs, which reached a stunning 96.7% (at nominal cutoff point) on the primarily Greek

population testing sample (sample split #1) and 93.7% respectively for the primarily Czech population testing sample (sample split #2). Just as notable is the dramatic increase in classification accuracy observed for the lateral HMIs of the mastoid processes as well as the occipital protuberance and supraorbital ridge HMIs, which in most cases exceeded 10 percentage units compared to the median scores.

Apart from the notable differences in classification scores among the evaluated morphometric traits, pairwise comparisons between testing and cross-validation groups also reveal population related differences in classification accuracy. As expected, the classification scores for both cross-validation groups are generally lower as opposed to the respective testing groups, since the former are merely represented as a population group in the training sample of each applied classifier. Nevertheless, there are certain exceptions to this general trend. The most notable exception concerns the lateral HMI of the right mastoid process, which in sample split #1 the classification scores are lower in the cross-validation group (median 78.4%–max 82.4%) of the Czech subgroup (HOMOLKA) compared to the testing group (82.8–92.2%, respectively), whereas in sample split #2, the exact opposite pattern is observed. The Greek subgroup (ABH) utilized for cross-validation yielded much higher results (median 86.7%–max 90.9%) as opposed to the respective testing group (76.5–85.3%). Most other exceptions concern moderate fluctuations of the median classification scores between testing and cross-validation groups in both sample splits, whereas the maximum classification scores are systematically lower in the cross-validation groups.

The classification results for multi-feature sex estimation are presented cumulatively in the two plots of Fig. 2.

Table 3 Classification results for testing and cross-validation groups of sample split #1

Evaluation groups	Testing (ABH, WLH, MOTOL)				Cross-validation (HOMOLKA)			
	Median		Maximum		Median		Maximum	
Classification statistics ^a	Zero ^c		Zero ^c		Zero ^c		Zero ^c	
Cutoff points ^b	$\pm 0.2^d$		$\pm 0.2^d$		$\pm 0.2^d$		$\pm 0.2^d$	
Occipital protuberance HMI	71.7	75.4/20.7	83.3	82.5/14.2	67.0	70.5/14.6	73.8	76.5/17.5
Supraorbital ridge HMI	76.5	80.0/13.9	86.3	89.5/14.4	80.6	86.5/15.2	85.4	89.9/16.5
Left mastoid process inferior HMI	66.2	68.4/19.7	78.5	82.2/30.8	70.6	74.6/20.0	77.5	78.6/19.6
Left mastoid process lateral HMI	80.0	84.3/12.7	90.8	91.4/10.8	79.4	84.1/15.8	84.3	86.7/18.6
Right mastoid process inferior HMI	67.2	72.2/19.5	79.7	85.4/25.0	70.6	74.3/14.6	77.5	80.7/13.7
Right mastoid process lateral HMI	82.8	86.6/12.7	92.2	93.3/7.0	78.4	80.6/11.2	82.4	82.2/9.3
Nasion-bregma EFDs	85.2	89.8/18.6	96.7	98.1/14.8	82.5	92.7/23.4	85.4	92.9/26.9

^a Classification statistics are given for the median and maximum scores over 100 iterations

^b The cutoff points for evaluating the output of each classifier's decision function. The full range of decision thresholds and their respective classification scores are shown in the box plots figures provided as supplementary material

^c At single cutoff point, the classification results (shown in percentage %) apply to the entire sample size

^d At decision threshold with lower and upper cutoff points both correct classification scores and rejection rates are given as percentages (classification/rejection). The rejection rate concerns the portion of the sample which cannot be determined with the given threshold range

Table 4 Classification results for testing and cross-validation groups of sample split #2

Evaluation groups	Testing (HOMOLKA, WLH, MOTOL)				Cross-validation (ABH)			
	Median		Maximum		Median		Maximum	
Cutoff points ^b	Zero ^c	± 0.2 ^d	Zero ^c	± 0.2 ^d	Zero ^c	± 0.2 ^d	Zero ^c	± 0.2 ^d
Occipital protuberance HMI	66.1	68.4/19.4	80.6	80.0/11.3	72.3	78.6/24.0	80.2	84.0/25.7
Supraorbital ridge HMI	78.6	82.1/11.2	89.3	90.8/9.5	74.5	78.1/15.5	80.6	85.5/15.3
Left mastoid process inferior HMI	67.6	69.9/17.7	83.8	84.9 / 22.1	62.6	65.7/21.2	71.7	72.0/27.8
Left mastoid process lateral HMI	80.9	83.7/14.4	89.7	91.1/17.6	79.8	84.2/16.8	86.9	87.5/11.1
Right mastoid process inferior HMI	69.1	73.6/20.0	79.4	84.0/23.5	70.4	73.6/19.9	80.6	79.7/19.4
Right mastoid process lateral HMI	76.5	80.6/15.8	85.3	87.7/12.3	86.7	91.2/12.9	90.8	92.7/11.9
Nasion-bregma EFDs	84.1	90.9/22.2	93.7	96.2/17.5	83.2	86.8/16.1	86.1	88.2/15.8

^a Classification statistics are given for the median and maximum scores over 100 iterations

^b The cutoff points for evaluating the output of each classifier's decision function. The full range of decision thresholds and their respective classification scores are shown in the box plots figures provided as supplementary material

^c At single cutoff point, the classification results (shown in percentage %) apply to the entire sample size

^d At decision threshold with lower and upper cut-off points both correct classification scores and rejection rates are given as percentages (classification/rejection). The rejection rate concerns the portion of the sample which cannot be determined with the given threshold range

The left plot shows the correct classification scores and their corresponding rejection rates of the combined best-performing classifiers from each sample split (weighted sum of 7 best-performing classifiers) applied to the respective testing and cross-validation groups of each sample split respectively. The plot on the right hand side of Fig. 2 shows the classification results of the multi-feature multi-sample combined classifiers (weighted sum of all 14 best-performing classifiers) applied to each population subgroup independently. The lines at the top of each plot correspond to the correct classification scores and their same color counterparts at the bottom show the respective rejection rates with the decision thresholds being offset accordingly. Combining the best-performing classifiers per sample split yields very high classification accuracy for the testing sample. More specifically, the accuracy is above 94% for the entire testing group (cutoff point at zero) and it reaches full separability between identified sexes just before the rejection rates reach 20%. The classification scores for the cross-validation groups are lower (at about 87% and 83% for splits #1 and #2, respectively), and they also rise at a lower rate with respect to the rejection rates increment. Utilizing all classifiers from both sample splits further improves the correct classification performance of the combined classifier, which is in excess of 91% at the nominal cutoff point for all subgroups under evaluation and rises faster with respect to rejection rates. It should be noted that in both cases of multi-feature combinations, the worst performing subgroup is the HOMOLKA sample. Its deficit regarding the classification accuracy is also apparent when considering the full

results on the individual morphometric features (see supplementary S1).

The multi-feature classifiers have been incorporated in a separate GNU Octave function, named "estimate_sex.m," which can be used to estimate sex from the morphometric features stored in the relevant .mat file generated with the skullanalyzer. The function, which is also available at <https://github.com/pr0m1th3as/cranial-sex-estimation-code>, can utilize any of the two sample-split multi-feature classifiers, which is advised when an unidentified case is expected to be of Greek or Czech population affinity, or the multi-sample multi-feature combined classifiers, when no population affinity is known. The function will output the estimated sex along with a posterior probability inferred from the cumulative decision function score and the rejection rates across different decision threshold ranges.

Discussion

The explicit purpose of this study has been to produce an efficient yet reliable method for estimating sex from unidentified crania. The aim for efficiency has been achieved through software implementation, namely the *skullanalyzer* computer program and the *estimate_sex.m* GNU Octave function, which collectively can analyze a cranium and estimate sex in a matter of seconds. Overall, given a cranial 3D model, the entire sex estimation procedure including initial landmark acquisition in Meshlab can be performed on any modern computer in a couple of minutes. Of course, this presumes the availability of a 3D model, but with the rapid advancements in 3D

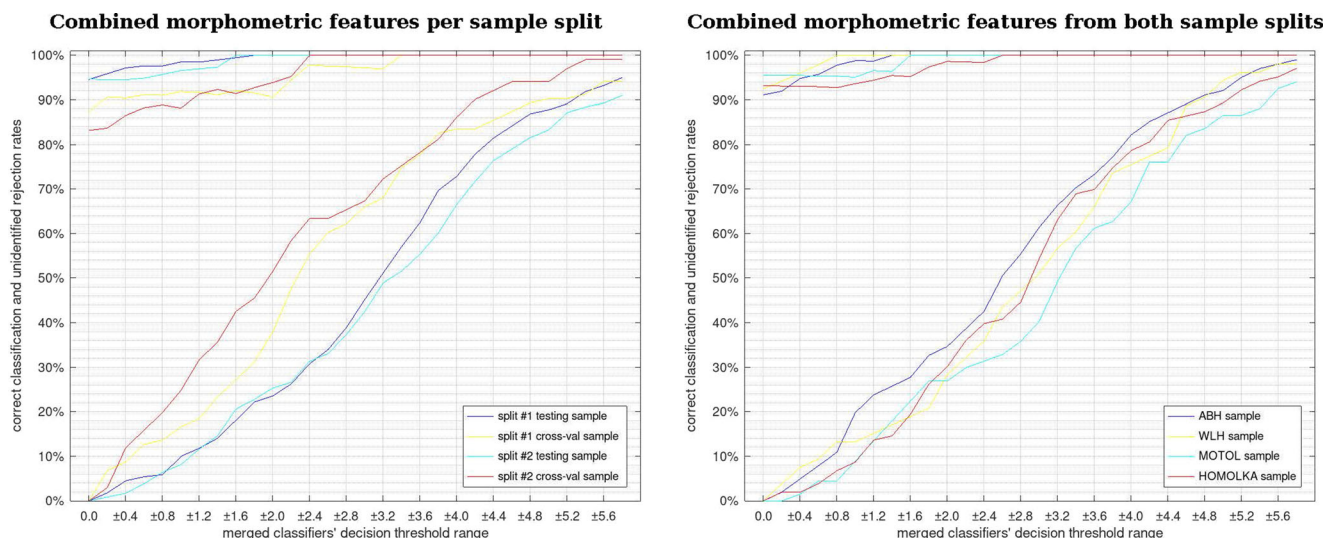


Fig. 2 Multi-feature classification performance

technologies and especially structured light scanners, this is a fairly accessible process, which is bound only to become cheaper and more efficient. Reliability, which is perhaps even more important, has been aimed and largely achieved in two ways. On one hand, incorporating landmark optimization algorithms to minimize observer error renders the extracted morphometric features, hence the final sex estimation, highly reproducible. On the other hand, applying an elaborate scheme of statistical analysis not only ensures that the reported accuracy is not overrated by the limitations of the samples but it also provides a posterior probability estimate that make sex estimation of an unidentified case even more reliable. To this end, the mixed-population sample-split, the repeated 2-way cross-validation approach, and the variable threshold range for classification played fundamental role.

Sex estimation heavily relies on statistical inference from population samples regardless the evaluation method or specific statistical approach. This, of course, emphasizes the importance of skeletal reference collections for such studies but further reflects the imposed limitations of these population samples regarding their representativeness. Previous work indicated that the Athens Collection is fairly representative of the Greek population. This evidence came not only from stature estimates against biometric records of living people of the same time period [48] but also from evaluating population specific performance of mandibular sex-discriminant functions against a Cretan sample from southern Greece [30]. Nevertheless, this representativeness does not necessarily interpret that the observed cranial variation reflects the overall variation present in the population. This applies to the Czech sample to an even greater extent since it comprises a particular group of healthcare recipients, and it is also unbalanced with many more males than female individuals. These inherent limitations regarding the representativeness of the available population samples were the major factor for adopting the mixed population sample-split strategy in this study with the minor

being more appropriate sample sizes for the repeated 2-way cross-validation. Of course, adding a relatively small group of individuals from a different population to a given population sample does not make the latter more representative. However, this practice artificially adds to the existing variation of the sample, which presumably results in a more generalized model fitting during supervised machine learning. Nevertheless, the trained models continue to exhibit population specificity in favor of the major population present in the training sample as shown by the reduced correct classification scores of the cross-validation group in both sample-splits. It should be noted that the present study is not designed to explore the differences in the expression of sexual dimorphism between the two populations. Hence, the differences in classification scores between testing and cross-validation groups do not reflect the magnitude of differential expression across individual sex diagnostic traits. These differences only testify the presence of population specificity related to cranial sex estimation.

The 2-way cross-validation scheme in evaluating individual sex diagnostic traits aimed to overcome the limitations of the population samples' size. Utilizing SVM classification can easily lead to overfitting when dealing with small samples [49]. Model overfitting is apparent when the classification accuracy between training and testing datasets differs significantly. Such an example is demonstrated in a recent work on sex estimation using the external morphology of the frontal bone and the frontal sinuses in the same contemporary Czech population used in the present study, which utilized SVM and leave-one-out cross-validation (LOOCV) and reported a total success rate of 98.05%, which decreased to 84.46% after cross-validation [20]. For this reason, the present work adopted the 2-way cross-validation scheme by evaluating the SVM classification accuracy on two distinct subgroups, which are different from the training sample in every iteration. Thus, the reported differences between testing and cross-validation

groups (Tables 3 and 4) are not because of overfitting but rather due to population specificity. Furthermore, the iterative approach aimed to identify the best SVM classifiers for each individual sex diagnostic trait. Nevertheless, sample size restrictions prevented a stratified repetitive approach which would ensure that sex is equally represented between training and testing datasets in every iteration. This resulted in a certain amount of the 100 iterations being heavily biased in sex representation, hence responsible for the low classification values observed in the descriptive statistics shown in supplementary S1. Consequently, the median classification scores reported in Tables 3 and 4 may be considered to underestimate the true sex-discriminant capacity of the respective sex diagnostic traits.

Sex-related skeletal variation is spread along a continuum from generally smaller, more gracile females to larger, more robust males [50] and despite morphometric methods being very capable of quantifying this variation, the plain dichotomous sex estimation usually applied on the measured variables reduces their effectiveness when applied to individual unidentified cases. This happens because every individual estimate is as reliable as the reported accuracy of its discriminant function although in practice the continuum of the expression of sexual dimorphism dictates that the majority of individual cases should be estimated with much higher reliability. Recent work on the validation and reliability of the DSP2 software (Diagnose Sexuelle Probabiliste v2) by Brůžek and his colleagues [51] has shown that taking into account posterior probabilities can dramatically increase the reliability of sex assessment for the vast majority of individually tested cases. Of course, the statistical properties of SVM classification with RBF kernel do not allow a straightforward calculation of posterior probabilities as performed by DSP2. In the present study, this is achieved through an indirect estimate of the posterior probability by utilizing a variable threshold range. Our results corroborate this approach since excluding the evaluated samples near the nominal cutoff point as undetermined yields much higher classification accuracy for the remaining valid ones. Finally, the combined multi-feature classifier (see right plot in Fig. 2) shows that although the overall classification accuracy ranges from 91 to 95% according to the which subsample is being tested, approximately 80% of the tested cases from ABH, WLH, and MOTOL groups can be determined with 100% accuracy, whereas the same also applies for the 60% of the HOMOLKA group.

The present work focused on efficiency and reliability in cranial sex estimation rather than merely reporting the highest possible classification accuracy. The present results regarding the classification accuracy of the proposed method may not be optimal due to the adopted sampling strategy and related limitations explained earlier, but they are still at the same levels with recent work featuring other novel approaches [35, 36]. Although the latter were assessed on single population

samples, which may have boosted their classification results. One of the key aspects of the newly proposed method is that it allows negligible space for observer error. This offers a significant advantage over the requirement for carefully taken photographs required by Nikita's [35] sex estimation R script or the need for user operated computer-aided measurements of specific geometric variables to be assessed by the improved back-propagation neural network proposed by Yang and his colleagues [36]. The presented sex estimation method bears a number of novelties into its implementation with, perhaps, the most important being the introduction of HMIs, which allow for the 3D cranial geometry to be evaluated by computer vision algorithms targeted to 2D images. Further work with larger and more diverse population samples is required to improve upon the existing performance of this method. To this end, the entire code base of the present work including the *skullanalyzer* software and the morphometric feature dataset has been made publicly available under a suitable license to facilitate future work and creative collaboration.

Acknowledgments The authors are grateful to Dr. Carmelo Mineo from the Strathclyde University for providing a customized algorithm to extract the ectocranial surface of the DICOM-derived 3D models, which enabled the inclusion of the Czech population sample in this study, and to Mr. Christos Garoufis from the National Technical University of Athens for his help with the Python scripting language. The authors wish to thank Dr. Martin Horák for managing the CT imaging in Na Homolce Hospital and are also thankful to Dr. Jana Velemínská from the Charles University and Prof. Efstratios Valakos from the National and Kapodistrian University of Athens for permission to study the Czech and Greek population samples respectively.

Funding information This work was partially supported by an Erasmus+ internship program (7026/2018) issued by the State Scholarships Foundation (Greece) granted to Andreas Bertsatos.

Availability of data and material All data used in the present study are publicly available at Zenodo repository. See manuscript for more details.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethics approval The patients' DICOM data have been used upon the approval of the ethics committee of the respective hospital. See manuscript for more details.

Code availability All code of custom built software is licensed under the GNU GPL v3 open source license and is available at GitHub. See manuscript for more details.

References

1. Bruzek J (2002) A method for visual determination of sex, using the human hip bone. *Am J Phys Anthropol* 117:157–168. <https://doi.org/10.1002/ajpa.10012>

2. Nikita E, Michopoulou E (2018) A quantitative approach for sex estimation based on cranial morphology. *Am J Phys Anthropol* 165:507–517. <https://doi.org/10.1002/ajpa.23376>
3. Bertsatos A, Papageorgopoulou C, Valakos E, Chovalopoulou M-E (2018) Investigating the sex-related geometric variation of the human cranium. *Int J Legal Med* 132:1505–1514. <https://doi.org/10.1007/s00414-018-1790-z>
4. Walker PL (2008) Sexing skulls using discriminant function analysis of visually assessed traits. *Am J Phys Anthropol* 136:39–50. <https://doi.org/10.1002/ajpa.20776>
5. Oikonomopoulou E-K, Valakos E, Nikita E (2017) Population-specificity of sexual dimorphism in cranial and pelvic traits: evaluation of existing and proposal of new functions for sex assessment in a Greek assemblage. *Int J Legal Med* 131:1731–1738. <https://doi.org/10.1007/s00414-017-1655-x>
6. Bruzek J, Murail P (2006) Methodology and reliability of sex determination from the skeleton. In: Schmitt A, Cunha E, Pinheiro J (eds) *Forensic anthropology and medicine*. Humana Press, Totowa, NJ, pp 225–242
7. Spradley MK, Jantz RL (2011) Sex estimation in forensic anthropology: skull versus postcranial elements. *J Forensic Sci* 56:289–296. <https://doi.org/10.1111/j.1556-4029.2010.01635.x>
8. Williams BA, Rogers TL (2006) Evaluating the accuracy and precision of cranial morphological traits for sex determination. *J Forensic Sci* 51:729–735. <https://doi.org/10.1111/j.1556-4029.2006.00177.x>
9. Krüger GC, L'Abbé EN, Stull KE, Kenyhercz MW (2015) Sexual dimorphism in cranial morphology among modern South Africans. *Int J Legal Med* 129:869–875. <https://doi.org/10.1007/s00414-014-1111-0>
10. Langley NR, Dudzik B, Cloutier A (2018) A decision tree for non-metric sex assessment from the skull. *J Forensic Sci* 63:31–37. <https://doi.org/10.1111/1556-4029.13534>
11. Godde K, Thompson MM, Hens SM (2018) Sex estimation from cranial morphological traits: use of the methods across American Indians, modern North Americans, and ancient Egyptians. *HOMO* 69:237–247. <https://doi.org/10.1016/j.jchb.2018.09.003>
12. Kranioti EF, İşcan MY, Michalodimitrakis M (2008) Craniometric analysis of the modern Cretan population. *Forensic Sci Int* 180:110.e1–110.e5. <https://doi.org/10.1016/j.forsciint.2008.06.018>
13. Franklin D, Cardini A, Flavel A, Kuliukas A (2013) Estimation of sex from cranial measurements in a Western Australian population. *Forensic Sci Int* 229:158.e1–158.e8. <https://doi.org/10.1016/j.forsciint.2013.03.005>
14. Casado AM (2017) Quantifying sexual dimorphism in the human cranium: a preliminary analysis of a novel method. *J Forensic Sci* 62:1259–1265. <https://doi.org/10.1111/1556-4029.13441>
15. Bigoni L, Velemínská J, Brůžek J (2010) Three-dimensional geometric morphometric analysis of cranio-facial sexual dimorphism in a Central European sample of known sex. *HOMO* 61:16–32. <https://doi.org/10.1016/j.jchb.2009.09.004>
16. Chovalopoulou M-E, Valakos ED, Manolis SK (2013) Sex determination by three-dimensional geometric morphometrics of the palate and cranial base. *Anthropol Anz* 70:407–425. <https://doi.org/10.1127/0003-5548/2013/0363>
17. Chovalopoulou M-E, Valakos ED, Manolis SK (2016) Sex determination by three-dimensional geometric morphometrics of the vault and midsagittal curve of the neurocranium in a modern Greek population sample. *HOMO* 67:173–187. <https://doi.org/10.1016/j.jchb.2015.09.007>
18. Chovalopoulou M-E, Valakos ED, Manolis SK (2016) Sex determination by three-dimensional geometric morphometrics of cranio-facial form. *Anthropol Anz* 73:195–206. <https://doi.org/10.1127/anthranz/2016/0470>
19. Gao H, Geng G, Yang W (2018) Sex determination of 3D skull based on a novel unsupervised learning method. *Comput Math Methods Med* 2018:1–10. <https://doi.org/10.1155/2018/4567267>
20. Čechová M, Dupej J, Brůžek J, Bejdová Š, Horák M, Velemínská J (2019) Sex estimation using external morphology of the frontal bone and frontal sinuses in a contemporary Czech population. *Int J Legal Med* 133:1285–1294. <https://doi.org/10.1007/s00414-019-02063-8>
21. Garvin HM, Ruff CB (2012) Sexual dimorphism in skeletal brow-ridge and chin morphologies determined using a new quantitative method. *Am J Phys Anthropol* 147:661–670. <https://doi.org/10.1002/ajpa.22036>
22. Schlager S, Rüdell A (2017) Sexual dimorphism and population affinity in the human zygomatic structure-comparing surface to outline data: analyzing zygomatic shape. *Anat Rec* 300:226–237. <https://doi.org/10.1002/ar.23450>
23. Bejdová Š, Dupej J, Krajiček V, Velemínská J, Velemínský P (2018) Stability of upper face sexual dimorphism in central European populations (Czech Republic) during the modern age. *Int J Legal Med* 132:321–330. <https://doi.org/10.1007/s00414-017-1625-3>
24. Luo L, Wang M, Tian Y, Duan F, Wu Z, Zhou M, Rozenholc Y (2013) Automatic sex determination of skulls based on a statistical shape model. *Comput Math Methods Med* 2013:1–6. <https://doi.org/10.1155/2013/251628>
25. Arigbabu OA, Liao IY, Abdullah N, Mohamad Noor MH (2017) Computer vision methods for cranial sex estimation. *IPSP Trans Comput Vis Appl* 9:19. <https://doi.org/10.1186/s41074-017-0031-6>
26. Walrath DE, Turner P, Bruzek J (2004) Reliability test of the visual assessment of cranial traits for sex determination. *Am J Phys Anthropol* 125:132–137. <https://doi.org/10.1002/ajpa.10373>
27. Lewis CJ, Garvin HM (2016) Reliability of the Walker cranial nonmetric method and implications for sex estimation. *J Forensic Sci* 61:743–751. <https://doi.org/10.1111/1556-4029.13013>
28. Fruciano C (2016) Measurement error in geometric morphometrics. *Dev Genes Evol* 226:139–158. <https://doi.org/10.1007/s00427-016-0537-4>
29. Shearer BM, Cooke SB, Halenar LB, Reber SL, Plummer JE, Delson E, Tallman M (2017) Evaluating causes of error in landmark-based data collection using scanners. *PLoS One* 12: e0187452. <https://doi.org/10.1371/journal.pone.0187452>
30. Bertsatos A, Athanasopoulou K, Chovalopoulou M-E (2019) Estimating sex using discriminant analysis of mandibular measurements from a modern Greek sample. *Egypt J Forensic Sci* 9:25. <https://doi.org/10.1186/s41935-019-0133-7>
31. Noble J, Flavel A, Aviv R, Franklin D (2019) Forensic anthropological standards for cranial sex estimation in Canada: preliminary results. *Aust J Forensic Sci* 51:S1–S4. <https://doi.org/10.1080/00450618.2019.1569720>
32. Kanchan T, Gupta A, Krishan K (2013) Estimation of sex from mastoid triangle – a craniometric analysis. *J Forensic Legal Med* 20:855–860. <https://doi.org/10.1016/j.jflm.2013.06.016>
33. Petaros A, Sholts SB, Slaus M, Bosnar A, Wärmländer SKTS (2015) Evaluating sexual dimorphism in the human mastoid process: a viewpoint on the methodology. *Clin Anat* 28:593–601. <https://doi.org/10.1002/ca.22545>
34. Shearer BM, Sholts SB, Garvin HM, Wärmländer SKTS (2012) Sexual dimorphism in human browridge volume measured from 3D models of dry crania: a new digital morphometrics approach. *Forensic Sci Int* 222:400.e1–400.e5. <https://doi.org/10.1016/j.forsciint.2012.06.013>
35. Nikita E (2019) Quantitative sex estimation based on cranial traits using R functions. *J Forensic Sci* 64:175–180. <https://doi.org/10.1111/1556-4029.13833>

36. Yang W, Liu X, Wang K, Hu J, Geng G, Feng J (2019) Sex determination of three-dimensional skull based on improved backpropagation neural network. *Comput Math Methods Med* 2019:1–8. <https://doi.org/10.1155/2019/9163547>
37. Yang W, Zhou M, Zhang P, Geng G, Liu X, Zhang H (2020) Skull sex estimation based on wavelet transform and Fourier transform. *Biomed Res Int* 2020:1–10. <https://doi.org/10.1155/2020/8608209>
38. Eliopoulos C, Lagia A, Manolis S (2007) A modern, documented human skeletal collection from Greece. *HOMO* 58:221–228. <https://doi.org/10.1016/j.jchb.2006.10.003>
39. Bertsatos A, Gkaniatsou E, Papageorgopoulou C, Chovalopoulou M-E (2019) “What and how should we share?” An inter-method inter-observer comparison of measurement error with landmark-based craniometric datasets. *Anthropol Anz* 92489. <https://doi.org/10.1127/anthranz/2019/1047>
40. Cignoni P, Callieri M, Corsini M, et al (2008) MeshLab: an open-source mesh processing tool. *Eurographics Ital chapter Conf* 8 pages. <https://doi.org/10.2312/LOCALCHAPTEREVENTS/ITALCHAP/ITALIANCHAPCONF2008/129-136>
41. Mineo C, Pierce SG, Nicholson PI, Cooper I (2016) Robotic path planning for non-destructive testing – a custom MATLAB toolbox approach. *Robot Comput Integr Manuf* 37:1–12. <https://doi.org/10.1016/j.rcim.2015.05.003>
42. Mineo C, Pierce SG, Nicholson PI, Cooper I (2017) Introducing a novel mesh following technique for approximation-free robotic tool path trajectories. *J Comput Des Eng* 4:192–202. <https://doi.org/10.1016/j.jcde.2017.01.002>
43. Mineo C, Pierce SG, Summan R (2019) Novel algorithms for 3D surface point cloud boundary detection and edge reconstruction. *J Comput Des Eng* 6:81–91. <https://doi.org/10.1016/j.jcde.2018.02.001>
44. Bertsatos A (2019) Skullanalyzer: a concrete way of extracting cranial morphometric features. Zenodo. <https://doi.org/10.5281/zenodo.3594565>
45. Bertsatos A (2019) User Manual & Algorithm Description Document “skullanalyzer” v1.0. Zenodo. <https://doi.org/10.5281/zenodo.3519248>
46. Eaton JW, Bateman D, Hauberg S, Wehbring R (2019) GNU Octave version 5.1.0 manual: a high-level interactive language for numerical computations
47. Bertsatos A, Chovalopoulou M-E, Brůžek J, Bejdová S (2020) Cranial morphometric features from Greek and Czech population samples. Zenodo. <https://doi.org/10.5281/zenodo.3632180>
48. Bertsatos A, Chovalopoulou M-E (2018) Secular change in adult stature of modern Greeks. *Am J Hum Biol* 30:e23077. <https://doi.org/10.1002/ajhb.23077>
49. Han H, Jiang X (2014) Overcome support vector machine diagnosis overfitting. *Cancer Inform* 13s1:CIN.S13875. <https://doi.org/10.4137/CIN.S13875>
50. Nieves JW, Formica C, Ruffing J, Zion M, Garrett P, Lindsay R, Cosman F (2004) Males have larger skeletal size and bone mass than females, despite comparable body size. *J Bone Miner Res* 20: 529–535. <https://doi.org/10.1359/JBMR.041005>
51. Brůžek J, Santos F, Dutailly B, Murail P, Cunha E (2017) Validation and reliability of the sex estimation of the human os coxae using freely available DSP2 software for bioarchaeology and forensic anthropology. *Am J Phys Anthropol* 164:440–449. <https://doi.org/10.1002/ajpa.23282>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.