



# Library

# The University of Bradford Institutional Repository

http://bradscholars.brad.ac.uk

This work is made available online in accordance with publisher policies. Please refer to the repository record for this item and our Policy Document available from the repository home page for further information.

To see the final version of this work please visit the publisher's website. Available access to the published online version may require a subscription.

Link to original published version: *http://dx.doi.org/10.1002/ajpa.22254* 

Citation: Villa, C., Buckberry, J., Cattaneo, C. and Lynnerup, N. (2013) Technical note: reliability of Suchey-Brooks and Buckberry-Chamberlain methods on 3D visualizations from CT and laser scans. American Journal of Physical Anthropology, 151(1): 158-163.

Copyright statement: © 2013 Wiley. Reproduced in accordance with the publisher's selfarchiving policy. This is the peer reviewed version of the article above, which has been published in final form at *http://dx.doi.org/10.1002/ajpa.22254*. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving



## Technical Note: Reliability of Suchey-Brooks and Buckberry-Chamberlain methods on

## 3D visualizations from CT and laser scans

Chiara Villa<sup>1</sup>, Jo Buckberry<sup>2</sup>, Cristina Cattaneo<sup>3</sup> and Niels Lynnerup<sup>1</sup>

<sup>1</sup>Laboratory of Biological Anthropology, Department of Forensic Medicine, University of

Copenhagen, Denmark;

<sup>2</sup>Biological Anthropology Research Centre, Archaeological Sciences, University of Bradford, UK;

<sup>3</sup>LABANOF, Forensic Anthropology and Odontology Laboratory, Department of Human Morphology, University of Milan, Italy.

Number of text pages (12), plus bibliography (2), number of figures (3), tables (3)

Abbreviated title: Age estimation methods and 3D visualizations

Key words: pubic symphysis, auricular surface, age estimation, 3D visualizations

Chiara Villa, Laboratory of Biological Anthropology, Department of Forensic Medicine, University of Copenhagen, Blegdamsvej 3, 2200 Copenhagen, Denmark, +45 35327229, chiara.villa@forensic.ku.dk

#### Abstract

Previous studies have reported that the ageing method of Suchey- Brooks (pubic bone) and some of the features applied by Lovejoy et al. and Buckberry-Chamberlain (auricular surface) can be confidently performed on 3D visualizations from CT-scans. In this study, seven observers applied the Suchey-Brooks and the Buckberry-Chamberlain methods on 3D visualizations based on CT-scans and, for the first time, on 3D visualizations from laser scans. We examined how the bone features can be evaluated on 3D visualizations and whether the different modalities (direct observations of bones, 3D visualization from CT-scan and from laser scans) are alike to different observers. We found the best inter-observer agreement for the bones versus 3D visualizations, with the highest values for the auricular surface. Between the 3D modalities, less variability was obtained for the 3D laser visualizations. Fair interobserver agreement was obtained in the evaluation of the pubic bone in all modalities. In 3D visualizations of the auricular surfaces, transverse organization and apical changes could be evaluated, although with high inter-observer variability; micro-, macroporosity and surface texture were very difficult to score. In conclusion, these methods were developed for dry bones, where they perform best. The Suchey-Brooks method can be applied on 3D visualizations from CT or laser, but with less accuracy than on dry bone. The Buckberry-Chamberlain method should be modified before application on 3D visualizations. Future investigation should focus on a different approach and different features: 3D laser scans could be analyzed with mathematical approaches and sub-surface features should be explored on CT-scans.

The establishment of 3D digital archives is a growing trend in many laboratories of anthropology, since three-dimensional computer visualizations (CT-scans, laser scans) are becoming more widely used in routine investigations. Indeed, CT-scans of cadavers are regularly performed before autopsy in many forensic departments (Poulsen and Simonsen, 2007); thus an immense amount of data in digital format is collected. Both techniques also play a major role in permanently documenting and recording anthropological material that, because of reburial or repatriation, may no longer be accessible. Furthermore a digital archive may be useful for rapid sharing of material located around the world and to preserve bones from further damage. Notable examples of this are the "Digitized Diseases" project led by the University of Bradford, a digital archive of pathological specimens (<u>http://barc.sls.brad.ac.uk/digitiseddiseases/index.php</u>), and the Smithsonian 3D virtual collection of fossil and primate crania, together with animal bones and artifacts generated both from CT-scans and laser scans (<u>humanorigins.si.edu/evidence/3d-collection</u>).

The benefits of 3D applications in anthropological and forensic sciences are highlighted in many publications and a number of previous studies investigated to what extent the existing methods could be applied to 3D visualizations of the bones. Sholts et al. (2010) demonstrated that the morphological features used for sex assessment could be easily evaluated on 3D visualizations of cranium and pelvis generated from a laser scanner; Dedouit et al. (2007 a, b), Grabherr et al. (2009), Decker et al. (2011) had similar results using 3D visualizations from CT-scans. Telmon et al. (2005) tested how two observers could apply the Suchey-Brooks method on 3D visualizations from CT-scans of pubic bone versus the direct observation of bones and reported almost total concordance between the modalities. Dedouit et al. (2007 a, b) applied the same approach in two forensic cases without reporting any difficulty. 3D visualizations of the auricular surface were used by Barrier et al. (2009), but

they couldn't completely apply the methods of Lovejoy et al. (1985) or of Buckberry and Chamberlain (2002), as they were unable to properly evaluate microporosity and the texture. They considered transverse organization, macroporosity and apical change, and introduced new criteria. Barrier et al. (2009), Pasquier et al. (1999) and recently Wade et al. (2011) focused on the trabecular bone structures. Ferrant et al. (2009) used a new approach that considered quantitative criteria such as angles, lengths and heights both on the pubic bone and the auricular surface. The existing methods for age estimation were never applied on 3D visualizations from laser scans; only a preliminary test using a quantitative numerical approach was conducted by Tocheri et al. (2002) on the reference casts of the os pubis (Suchey et al., 1988).

We decided to perform new tests: seven observers from three different laboratories tested the methods of Suchey-Brooks (Brooks and Suchey, 1990) and Buckberry and Chamberlain (2002) on 3D CT-scan visualizations. Furthermore, for the first time, we tried the methods on 3D visualizations from laser scans. We examined how the bone features essential for the application of the existing methods can be evaluated on 3D visualizations and whether the different modalities (direct observations of bones, 3D visualizations from CT and laser scans) are alike to different observers, taking no notice of the accuracy of age estimations at this stage.

### MATERIALS AND METHODS

The sample consisted of 12 right male pubic bones from cadavers and 8 auricular surfaces (4 left and 4 right non-paired) from an archeological collection. The pubic bones were obtained from autopsies at the Department of Human Morphology, University of Milan, Italy. The auricular surfaces were randomly chosen from a small "recording kit" of bones assembled by Buckberry and Chamberlain as representative of the different scores of the

features described in their paper using unstratified bones from the Black Gate cemetery in Newcastle (Buckberry and Chamberlain, 2002). The bones were scanned at 120kV and 110 mAs using a Siemens Somatom Sensation 4 Multislice spiral scanner. The reconstructions were performed using 0.75 mm slice thickness, 0.3 mm slice increment interval. The field of view was 234 mm and the pixel size 0.46 mm. We used Mimics software to create 3D visualizations: to generate a 3D visualization, we created a mask with attenuation from -300 to 3071 HU for the auricular surface and from -600 to 3071 HU for the pubic symphysis. In this way we could include the whole bone. Surface scans were performed using a 650nm red line laser (optical output power 5mW) and a high resolution camera (CCD monochrome camera with 1024 x 768 resolution, sensitivity of 0.5 lx at 1/15s, gain 20 dB, 30 frame rate with focal length of the lens of 6 mm). The resolution of the 3D laser visualization was 0.1 mm. The 3D visualizations of the bones were exported as STL format and visualized with MeshLab software (http://meshlab.sourceforge.net); the observers could view and rotate the bones from any perspective.

We assigned a phase (from I to VI) to each pubic bone applying the Suchey-Brooks scoring system (Brooks and Suchey, 1990) and a stage (from I to VII) at each auricular surface based on the composite score obtained using the method of Buckberry and Chamberlain (2002). All observers used the Suchey-Brooks casts (Suchey et al., 1988) while scoring the pubic bones, except one (JB), who is very familiar with their appearance. Three tests were performed, all involving both pubic bone and auricular surface:

- Test 1: intra-observer agreement pertaining to the direct observation of bones, 3D CTscan visualizations and 3D laser scan visualizations by the most experienced person working with data from laser and CT-scans.
- Test 2: inter-observer agreement pertaining to all of the above-mentioned materials. Three observers with 5-10 years of anthropological experience participated.

5

• Test 3: inter-observer agreement pertaining to just the 3D visualizations (from CT-scan and from laser scan). Seven observers with varying degrees of anthropological experience (from minimum 3 to maximum 15 years) participated.

The datasets were scored independently and all tests were performed with at least two weeks interval to avoid bias.

#### **Statistical Analysis**

Statistical analysis was performed using SPSS. As Telmon et al. (2005), we calculated the weighted kappa (Cohen's kappa) to evaluate intra- and inter-observer agreement, as well as for agreement between datasets. Following guidelines of Landis and Koch (1977), a kappa of <0.2 was considered poor agreement, 0.21-0.4 fair, 0.41-0.6 moderate, 0.61-0.8 substantial and more than 0.8 near complete agreement.

#### RESULTS

### Test 1

The intra-observer results for bones and 3D visualizations (from CT and laser scans) are shown in Table 1. For the pubic bone, substantial intra-observer agreement was found for bones and 3D CT-scan visualizations (0.61 and 0.66, respectively) and moderate agreement was obtained for laser scans (k=0.47). The agreement values for the auricular surfaces were higher than those achieved for pubic bones: intra-agreement for all datasets was excellent (k $\geq$  0.81).

## Test 2

The best inter-observer agreements (3 observers) were obtained for the direct observation of bones rather than for the 3D visualizations, with kappa highest for the auricular surface. For 3D visualizations, the best agreements were reached using the laser scans.

## *a) Pubic bone*

Fair agreement was found between observers, in a very similar degree for all datasets (k values between 0.23 and 0.30). Similar agreement was obtained between bones and 3D visualizations (agreement between datasets). We observed a better agreement between bones and 3D CT-scan visualizations by the observer most experienced with CT data. Indeed, observer #1 obtained a kappa value of 0.54, as compared with the values of 0.17 and 0.07 for the observers #2 and #3.

#### *b)* Auricular surface

The inter-observer agreement was substantial for bones (k=0.77), moderate for laser scans (k=0.60), but a value of 0.19 (poor agreement) was achieved for CT-scans. The comparison of the scores between bones and 3D visualizations revealed a substantial agreement with 3D laser scan visualizations (k=0.66) and only fair agreement with 3D CT scan visualizations (k=0.36).

## Test 3

3D visualizations generated from CT-scans and laser scans were scored by seven observers (Table 2). As for test 2, higher results of agreement were achieved in the 3D visualizations from laser scans compared with CT scans, and the auricular surface showed higher kappa values than the pubic bone.

#### *a) Pubic bone*

The inter-observer agreement was fair: very similar values of k were detected within three (test 2) or seven (test 3) observers. An increase in the k values could be observed for laser scans, but they remained in the fair agreement interval. A moderate agreement was achieved

between 3D visualizations. No clear trend was detected among the datasets and it seemed that the variability depended on the observers but was not attributable to experience.

Scoring results for two particularly problematic pubic bones are shown in Figures 1-3 to illustrate some issues affecting visualization of features using the different modalities. For sample #1 (young man of 24 years) we attributed seven different combinations of phases and only in two observations there was agreement between CT and laser visualizations (Fig.1a). 3D visualizations from laser scans produced lower phase scores than for CT scan visualizations by the same observer in three cases, whereas laser visualizations were higher than CT visualizations in two cases. For the 3D CT-scan visualization (Fig. 2b), the billowing was less marked than on the dry bone and on the laser scan visualization; this could explain the large disagreement in attributed phases. For sample #9 (a 55 year-old man) similar results were obtained: there was residual billowing in the ventral portion (Fig.3a), but this feature was difficult to identify on 3D visualizations, especially in the laser scans; the ventral margin, closer to completion, could be interpreted differently, as an evidence of breakdown – probably for that reason observer #2 scored the pubic bone as phase VI (Fig.1b).

## b) Auricular surface

In comparison with the results obtained from 3 observers, we noted an improvement in the inter-observer agreement both in 3D CT-scan and 3D laser scan visualizations, from poor agreement between three observers to moderate agreement between seven observers (CT) and from moderate to substantial agreement (laser) (Table 2).

It seems reasonable to expect a better agreement between stages, which summarize the score of all morphological features, then between each single one of them. The agreements between scores for each one of the five features of the auricular surface, following by Buckberry-Chamberlain method, are shown in Table 3. K< 0, i.e. agreement is less than expected by chance, was obtained between the bones and the 3D visualizations from CT-

scans for transverse organization, surface texture and microporosity. No better result was achieved for macroporosity (k=0.09); the only feature with moderate agreement was apical change. The transverse organization could be scored by all observers but the proportion of the surface covered was extremely difficult to evaluate. Several observers could not see any surface texture and microporosity; in many cases, one point of score was attributed to most visualizations, corresponding to "90% or more of surface is finely granular" and "no microporosity is present". Macroporosity was also not easy to evaluate; in many cases it was difficult to establish whether larger holes were macroporosity, cortical defects or taphonomic damage. The scoring of apical change was slightly easier. Similar results and similar problems were noticed in the scoring of the 3D visualizations from laser scans. Looking at the agreement between bones and laser and between the 3D visualizations, kappa values are slightly higher, but not so different: the surface texture, the micro- and macroporosity created doubts during scoring (poor agreement). Transverse organization seemed easier to score for laser scan visualizations, but again the degree of expression was conditioned by the observers. The best agreement for macroporosity was obtained between the 3D visualizations (k=0.55). For apical changes, moderate agreement was achieved between bones and laser scans and between 3D visualizations.

#### DISCUSSION

Previous studies (Telmon et al., 2005; Barrier et al., 2009) reported that the method of Suchey- Brooks and some of the features applied in the Lovejoy and Buckberry-Chamberlain methods can be confidently performed on 3D pelvic bone visualizations from CT-scans, however their inter-observer tests were based only on two observers. In this study, seven observers, from different laboratories, applied the Suchey-Brooks method and performed the

9

entire scoring system of the Buckberry and Chamberlain method on 3D visualizations from CT-scans and from laser scans, and we could not agree with the previous studies.

The CT- scanner settings determine the spatial resolution of the 3D visualization. The voxel size in the 3D visualization depends on pixel size and slice thickness. The x and y (pixel dimensions) of each slice in the transverse plane are determined by the matrix size and the fields of view. The third voxel dimension (z) is perpendicular to the transverse plane and corresponds to the slice thickness (Kalender, 2011). In this study the pixel size was 0.46 mm and the slice thickness 0.75 mm for both the auricular surfaces and the pubic bones. The voxel dimensions in Telmon et al. (2005) for the pubic bones were coarser, corresponding to 0.98 mm in x/y-axes and 0.8 mm in z-axis. In Barrier et al. (2009), the pixel size ranged from 0.12 to 0.16 with a slice thickness of 0.6, with a better spatial resolution than our CT-scans (Fabrice Dedouit and Pierre Barrier, personal communication, 2013).

The best agreement values were found for the direct observation of bones. This is not surprising, since the tested methods were developed and calibrated on dry bones. 3D visualizations generated from a laser scanner reproduced the morphological features better than CT-scans due to the better resolution, resulting in better agreements. We found only fair inter-observer agreement (k= 0.3) applying Suchey-Brooks method on dry bones, but same results were reported by Galera et al. (1995). As we expected, slightly less agreement was found in 3D visualizations. Our inter-observer agreements for the dry auricular surface are slightly higher than those reported by Buckberry and Chamberlain (2002) (kappa of 0.77 rather than 0.66), perhaps because we chose the bones from a "recording kit" which comprises bone specimens specifically selected to be the best representatives of the different features. The agreements in auricular surface were better than those in the pubic bone, because the Buckberry-Chamberlain method allows separate evaluation of the morphological features. Looking at the intra-agreement results, we deduced that the same person evaluated

the features in a similar way between the datasets, and that the experience in working with 3D visualizations increased the reliability of the test, and not just the experience in anthropology, as reported in Telmon et al. (2005). The same authors found excellent agreement both between direct observation of bones and 3D CT-scan visualization and between observers in the use of the Suchey-Brooks method (Telmon et al., 2005). Even though our 3D visualizations had less voxel dimensions than those used by Telmon et al. (2005), we found only fair agreement in all tests. Some observers found difficulties in the evaluation of the ventral rampart and the depth of the depression. For instance, it was difficult to establish if the ventral rampart was almost complete, corresponding to phase IV, or if signs of erosion were present corresponding to early phase VI. Telmon et al. (2005) reported similar difficulties. Testing all features following the Buckberry - Chamberlain method, we found that microporosity and surface texture are impossible to score in a reliable way on 3D visualizations, as reported in Barrier et al. (2009), who excluded these features after preliminary analysis as they were not informative. However, our results disagree with those of Barrier et al. (2009) because we could not consider macroporosity and transverse organization as features that observers can consistently evaluate in the same way as on dry bones. Transverse organization could be identified but the degree of expression was difficult to evaluate and it was difficult to establish whether larger holes were macroporosity, cortical defects or taphonomic damage. We found moderate agreement only for apical changes. The 3D visualizations from laser scans seemed to better express auricular surface features, with moderate agreement in apical change and transverse organization, but micro- and macroporosity showed only fair agreement. Again surface texture was difficult to distinguish. Barrier et al. (2009) performed scans with a thinner slice thickness and smaller pixel size compared to our CT-scans and these settings could partially explain their better results. We

did not find so high agreement, even in 3D laser visualizations that had a similar or better resolution than the 3D CT-scans visualizations obtained from Barrier et al. (2009).

Both laser scans and CT techniques are useful instruments to record and document anthropological and archeological bones. The major advantage of 3D visualizations from CTscans is the possibility to make virtual anthropological examinations even on soft tissuecovered bone, thus avoiding time spent to macerating bones. However, we think that the features useful for age estimation must be evaluated with a different approach, since the observers found several problems in scoring them. CT-images allow exploration of the subsurface features of the bones and they should be further investigated: Pasquier et al. (1999) and Wade et al. (2001) assessed the potential of trabecular bone structure for age estimation. The 3D visualizations from laser scans have the important benefit of significantly better resolution than 3D CT-scan visualizations; the morphological features of the bone are better replicated than by CT-scan, but one loses the tactile perception that may also yield some information never can be felt in the virtual bone. The use of the texture, i.e. photograph, on 3D laser visualizations may add detail and improve the age estimation if dry bones cannot be examined (e.g. after reburial or repatriation). We believe that these 3D models should be analyzed in a quantitative way, e.g. curvature analysis as suggested by Tocheri et al. (2002), so that the expression of the surface features can be related to age and perhaps reduce the impact of the operator's subjectivity.

#### CONCLUSIONS

The method of Suchey-Brooks, developed for direct observation of bones, can still be applied on 3D visualizations from CT-scan or from laser scans, but with reduced accuracy. The method of Buckberry and Chamberlain is difficult to use with 3D visualizations: the transverse organization and apical changes can be evaluated, but with high inter-observer

variability, however micro-, macroporosity and surface texture are very difficult to score. In both cases, we recommend training with images from CT and laser scanner before using the existing methods on 3D visualizations.

Future investigation should perhaps look more closely at the evaluation of different morphological features; we should not use methods developed for dry bones on 3D visualizations. Digital data have different potential: 3D laser models could be analyzed with mathematical approaches, for example analyzing the shape and undulation of the surfaces. CT scans should be used to explore the sub-surface features and not limit investigation to the exterior of the bones.

## Acknowledgements

The authors thank Charlotte Primeau from Laboratory of Biological Anthropology (University of Copenhagen), Daniel Gaudio, Daniele Gibelli and Debora Mazzarelli from LABANOF (University of Milan) for their contribution to the inter-observers variability test. We are also grateful to the Editor-in-Chief, the Associate Editor, and one anonymous reviewer for their helpful suggestions and comments.

# Literature cited

- Barrier P, Dedouit F, Braga J, Joffre F, Rouge D, Rousseau H, Telmon N. 2009. Age at death estimation using multislice computed tomography reconstructions of the posterior pelvis. J Forensic Sci 54:773-778.
- Brooks S, Suchey J. 1990. Skeletal age determination based on the Os Pubis: A Comparison of the Acsadi-Nemeskeri and Suchey-Brooks Methods. J Hum Evol 5.
- Buckberry JL, Chamberlain AT. 2002. Age estimation from the auricular surface of the ilium: A revised method. Am J Phys Anthropol 119:231-239.
- Decker SJ, Davy-Jow SL, Ford JM, Hilbelink DR. 2011. Virtual determination of sex: metric and nonmetric traits of the adult pelvis from 3D computed tomography models. J Forensic Sci 56:1107-1114.
- Dedouit F, Telmon N, Costagliola R, Otal P, Joffre F, Rouge D. 2007a. Virtual anthropology and forensic identification: report of one case. Forensic Sci Int 173:182-187.
- Dedouit F, Telmon N, Guilbeau-Frugier C, Gainza D, Otal P, Joffre F, Rouge D. 2007b. Virtual autopsy and forensic identification-practical application: a report of one case. J Forensic Sci 52:960-964.
- Ferrant O, Rouge-Maillart C, Guittet L, Papin F, Clin B, Fau G, Telmon N. 2009. Age at death estimation of adult males using coxal bone and CT scan: a preliminary study. Forensic Sci Int 186:14-21.
- Galera V, Ubelaker DH, Hayek LC. 1995. Interobserver error in macroscopic methods of estimating age at death from the human skeleton. International Journal of Anthropology 10:229-239.
- Grabherr S, Cooper C, Ulrich-Bochsler S, Uldin T, Ross S, Oesterhelweg L, Bolliger S, Christe A, Schnyder P, Mangin P, Thali, MJ. 2009. Estimation of sex and age of "virtual skeletons"--a feasibility study. Eur Radiol 19:419-429.
- Kalender WA. 2011. Computed Tomography: Fundamentals, System Technology, Image Quality, Applications. Germany: Publicis, Erlangen.
- Landis JR, Koch GG. 1977. Application of Hierarchical Kappa-Type Statistics in Assessment of Majority Agreement among Multiple Observers. Biometrics 33:363-374
- Lovejoy CO, Meindl RS, Pryzbeck TR, Mensforth RP. 1985. Chronological metamorphosis of the auricular surface of the ilium: a new method for the determination of adult skeletal age at death. Am J Phys Anthropol 68:15-28.
- Pasquier E, De Saint Martin Pernot L, Burdin V, Mounayer C, Le Rest C, Colin D, Mottier D, Roux C, Baccino E. 1999. Determination of age at death: assessment of an algorithm of age prediction using numerical three-dimensional CT data from pubic bones. Am J Phys Anthropol 108:261-268.
- Poulsen K,Simonsen J. 2007. CT-scanning as a routine examination before Medico-Legal autopsy. Ugeskr Laeger 169:3298-3299.
- Sholts SB, Warmlander SK, Flores LM, Miller KW, Walker PL. 2010. Variation in the measurement of cranial volume and surface area using 3D laser scanning technology. J Forensic Sci 55:871-876.
- Suchey JM, Brooks ST, Katz D. 1988. Instructions for use of the Suchey-Brooks system for age determination of the female os pubis. Instructional materials accompanying female pubic symphysial models of the Suchey-Brooks system. Fort Collins, Colorado: France Casting.
- Telmon N, Gaston A, Chemla P, Blanc A, Joffre F, Rouge D. 2005. Application of the Suchey-Brooks method to three-dimensional imaging of the pubic symphysis. J Forensic Sci 50:507-512.

- Tocheri MW, Razdan A, Dupras TL, Bae M, Liu D. Three Dimensional Quantitative Analyses of Human Pubic Symphyseal Morphology: Can Current Limitations of Skeletal Aging Methods Be Resolved. Am J Phys Anthropol (Suppl34) 155.
- Wade A, Nelson A, Garvin G, Holdsworth DW. 2011. Preliminary radiological assessment of age-related change in the trabecular structure of the human os pubis. J Forensic Sci 56:312-319.



Figure 1: scatter plots for cases 1 and 9 of pubic bones evaluated in 3D visualizations from CT-scans (x-axis) and laser scans (y-axis), divided by observers



Figure 2: photograph (a), 3D visualization from CT scan (b) and 3D visualization from laser scan(c) of sample 1. Scale bar corresponds to 1 cm



Figure 3: photograph (a), 3D visualization from CT scan (b) and 3D visualization from laser scan(c) of sample 9. Scale bar corresponds to 1 cm