Inter-observer agreement of the Suchey-Brooks method from a French computed tomography sample

O. Dubourg, R. Toutin, A. Bascou, N. Jousset, N. Telmon, P. Saint-Martin, F. Savall

▶ **To cite this version:**

O. Dubourg, R. Toutin, A. Bascou, N. Jousset, N. Telmon, et al.. Inter-observer agreement of the Suchey-Brooks method from a French computed tomography sample. La Revue de Médecine Légale, 2020, 11, pp.74 - 80. 10.1016/j.medleg.2020.03.002 . hal-03491014

HAL Id: hal-03491014
https://hal.science/hal-03491014

Submitted on 22 Aug 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Version of Record: https://www.sciencedirect.com/science/article/pii/S1878652920300237
Manuscript_e1e37053fe57eddef8a319ed67e7d303

Inter-observer agreement of the Suchey-Brooks method from a French computed tomography sample

Accord inter-observateurs lors de l'application de la méthode de Suchey-Brooks à partir d'un échantillon scanographique

Olivier Dubourg[1,2] M.D. ; Ryan Toutin[3] ; Agathe Bascou[3] ; Nathalie Jousset[2] M.D. ; Norbert Telmon[1,3] M.D., Ph.D. ; Pauline Saint-Martin[1,4] M.D., Ph.D. and Frédéric Savall[1,3] M.D., Ph.D.

Affiliations and addresses

[1] Laboratoire d'Anthropologie Moléculaire et d'Imagerie de Synthèse, Université Toulouse III, AMIS FRE 2960 CNRS,
37 allées Jules Guesde,
31000 Toulouse, France

[2] UNAM, Service de Médecine Légale, Centre Hospitalier Universitaire
4 rue Larrey,
49933 Angers Cedex 9, France

[3] Service de Médecine Légale, Centre Hospitalier Universitaire,
Rangueil, 1 avenue du Professeur Jean Poulhès,
31059 Toulouse Cedex 9, France

[4] Institut Médico-Légal, Centre Hospitalier Régional Universitaire
37044 Tours Cedex 9, France

Corresponding author

Olivier Dubourg, UNAM, Service de Médecine Légale, Centre Hospitalier Universitaire
4 rue Larrey, 49933 Angers Cedex 9, France
✉: olivierdub@hotmail.fr
Tel. 033 (0)2 41 35 59 28
Fax 033 (0)2 41 35 39 48

Conflict of interest

The authors declare that they have no conflicts of interest.

© 2020 published by Elsevier. This manuscript is made available under the CC BY NC user license
https://creativecommons.org/licenses/by-nc/4.0/

Introduction

In forensic anthropology, the estimation of age at death is mainly required to establish a biological profile and facilitate individual identification [1–3]. The pubic symphysis represents one of the most studied skeletal structures due to its late maturation [4]. In this respect, the Suchey-Brooks (SB) system is one of the most commonly used and tested methods of assessment based on the morphological characteristics of the pubic symphysis [5–7].

However, the SB method has certain drawbacks, including frequently inaccurate estimation for individuals over 40 years old [8–11]. The inter-observer reliability is largely described in the literature, highlighting some various results [12,13]. Indeed, we witnessed in the different previous studies that calculation methods would never differentiate the maturation phases from the degeneration ones.

The aim of this work was to study the inter-observer agreement of the SB method among three observers with heterogeneous experience from a computed tomography sample by focusing on the first three phases and the last three phases in the SB method. Moreover, we intended to test the errors of estimation in the three observers by calculating inaccuracy and bias.

Materials and Methods

Sample

We undertook a retrospective study of pubic bones from male and female adults undergoing clinical multi-slice computed tomography (MSCT) in a hospital in Toulouse, France, between November 2017 and April 2018. MSCT was mainly requested in the clinical context of general condition alteration or suspicion of abdominal diseases. Patients with a known history of bone disease or pelvic trauma as well as patients with materials causing artifacts, such as hip replacements, were excluded. We hypothesized that the SB method could be applied for both dry bones and CT images without any differences. In order to constitute the sample, a confirmed experimenter preliminarily performed phase allocation of many symphyseal faces of pubic bones using morphological criteria of the SB system [5]. As described in this method, the phases are defined as follows: *Phase I* : symphyseal face has a billowing surface with ridges and furrows which usually extends to include the pubic tubercle, the horizontal ridges are well-marked and ventral beveling may be commencing with a lack of delimitation

of either extremity ; *Phase II* : symphyseal face may still show ridge development, the face has commencing delimitation of lower and/or upper extremities occurring with or without ossific nodules, and the ventral rampart may be in beginning phases as an extension of the bony activity at either or both extremities ; *Phase III* : symphyseal face is smooth and shows lower extremity and ventral rampart in process of completion, dorsal plateau is complete and there is no lipping of symphyseal dorsal margin and bony ligamentous outgrowths ; *Phase IV* : symphyseal face is generally fine grained and usually the oval outline is complete at this stage, but a hiatus can occur in upper ventral rim ; *Phase V* : symphyseal face is completely rimmed with some slight depression of the face itself, relative to the rim, moderate lipping is usually found on the dorsal border with more prominent ligamentous outgrowths on the ventral border ; *Phase VI* : symphyseal face may show ongoing depression as rim erodes, ventral ligamentous attachments are marked, crenulations may occur and the shape of the face is often irregular at this stage [5]. Within each of the six usual phases (I to VI), the early pattern and the advanced pattern were separated [5]. This step enabled the creation of a stage-balanced sample that consisted of a total of 50 males (mean age 40.1 years; median age 33 years; minimum age 15 years; maximum age 84 years; SD 22) and 50 females (mean age 46.6 years; median age 42 years; minimum age 15 years; maximum age 98 years; SD 25.2) (Fig. 1).

According to French law, as well as the patient's non-objection, the results of medical imaging examinations may be used retrospectively without their consent when these examinations have been conducted for clinical purposes and when they have been recorded anonymously (Article 40-1, Law 94-548 of 1 July 1994).

MSCT

MSCT images were obtained through a Picture Archiving and Communication System (PACS, McKesson Medical Imaging Group, Richmond, BC, Canada) used by the hospital. Examinations were performed on an Optima 64-detector row CT scanner (General Electric Medical Systems). The image matrix was 512 x 512 pixels. A soft tissue filter was used and axial reconstructions were performed every 1.25 mm with a 1.25 mm slice thickness.

Scans were saved as digital imaging and communications in medicine (DICOM) files.

Morphological 3D analysis and phase allocation

The right pubis was chosen arbitrarily. Reconstruction of the surface was performed using the *Isosurface* function of Amira® software (Amira 5.4.2, Mercury Computer System, Inc.,

Chelmsford, MA, USA). The right pubis was isolated and edited with the *Extract Surface* and *Surface Editor* functions. A phase was allocated to each pubis using morphological criteria of the SB system with the early and the advanced patterns [5]. Phase allocation was consensually performed by three observers: one experienced anthropologist (observer 1), one forensic pathologist (observer 2) and one student in forensics (observer 3). All images were analyzed without knowledge of the age of the subject and in an arbitrary order. Observers could rotate the bone in three dimensions. Six symphyseal faces of pubic bones representative of each stage are illustrated in Fig. 2.

General statistical analysis

All statistical analyses and the graphical representations were conducted with R 3.0.2 software (R Development Core Team, http://www.R-project.org).
To assess inter-observer reliabilities between the three observers, squared weighted Cohen's kappas were calculated, with and without the allocation of the early and advanced patterns. Following the 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, 0.6-0.8 substantial, and more than 0.8 near complete agreement.

Errors of estimation in the three observers

The errors of estimation in the three observers were tested by the calculation of inaccuracy and bias. The assigned age of individuals in the sample was defined as the mean age of the obtained phase. Mean errors between the estimated and real ages were calculated at 10-year age intervals. Inaccuracy was calculated as: $\sum|estimated\ age - real\ age|/n$, showing the average magnitude of the absolute error, for each observer. Moreover comparison of the inaccuracy for each pair of observers was carried out using the paired Student's t-test. Bias was calculated as: $\sum(estimated\ age - real\ age)/n$, expressing the tendency for either over- or underestimation of age.

Results

General statistical analysis

The distribution by allocated phases of the population showed a stage-balanced sample for each observer (Fig. 3). Regarding the entire study sample, the inter-observer reliabilities were excellent for both sexes between all the observers (with and without early and advanced patterns) (Tables 1 and 2).

By focusing on the first three phases (I–III) versus the last three phases (IV–VI) for each sex, the inter-observer reliabilities were moderate to excellent for both sexes between observers 1 and 2 and moderate to substantial for observer 3 (Tables 3 and 4).

Errors of estimation in the three observers

Tables 5 and 6 show the inaccuracy and bias in 10-year intervals for each observer when the mean ages of the SB series were applied to our sample.

Discussion

In the SB system, the changes are divided into six phases for males and females, each with an age range, mean and an early and late stage (not considered for age estimation in the method) [5,7]. Some previous studies have highlighted how most estimation errors occur for the advanced stages (IV, V and VI), whereas reproducibility seems to be considerable in the initial stages (I, II and III) [4,12–14]. Most of them reported inter-observer tests based on only two observers.

Regarding our study, we focused on the inter-observer agreement in the SB method among three observers with heterogeneous experience from a French computed tomography sample. With and without early and advanced patterns, we highlighted a very moderate "learning effect" with excellent inter-observer reliabilities for both sexes between the most experienced observers (kappa values between 0.91 and 0.95) and very strong inter-observer agreements by considering the least experienced observer (kappa values between 0.8 and 0.94). The topic of the inter-observer reliability is largely described in the literature. Lottering et al. reported a kappa value of 0.87, whereas Villa et al. reported a kappa value of 0.3 [12,13]. Unlike our study using a stage- and sex-balanced sample, most of the previous studies used a sample with an overrepresentation of older individuals [12] and males [15].

As the comparison was between the inter-observer reliabilities of the first three phases versus the last three phases, we could have expected that reproducibility was better for young individuals (phases I to III) as compared to older individuals (phases IV to VI). Indeed, it seems to be more difficult to assign a phase for older individuals – with continuous morphological changes specific to degeneration, requiring a major experience – than for younger individuals where the phase allocation is linked to discrete morphological criteria specific to maturation. However, our results revealed there was no significant difference between the inter-observer reliabilities for these two groups (kappa values between 0.44 and 0.85). Besides, the inter-observer agreements were overall similar, ranging from moderate to excellent between observers 1 and 2 and moderate to substantial for observer 3.

Moreover, we noticed that the inaccuracy rose with age, especially from 60 years old and beyond, with a growing underestimation for both sexes. We reported a 20-years underestimation over 60 years old and a 30-years underestimation over 80 years old for both sexes by the three observers. Besides, we highlighted a moderate overestimation for younger individuals. These errors of estimation were slightly dependent on the experience of the observer. Despite some different methodological analyses and various populations, these results generally correspond with the values in the literature [9,12,16]. This high degree of inaccuracy is absolutely not acceptable, especially regarding death investigations in the forensic context. The results should be interpreted with great caution and many reservations when the SB series are considered to estimate the age at death of an individual. The underestimation of the older individuals is a prominent issue still debated in the literature. Hoppa noticed that this underestimation was not simply a product of having only six stage means [8]. Within this framework, it is interesting to note that more and more authors have suggested to add a later stage, stage VII, to specify age estimation for the last decades [17–19].

Furthermore, the use of a reference sample built from a forensic population, such as in the SB sample, in which young males are overrepresented likely increases this lack of precision, particularly by the effect of “attraction to the mean” [20]. In the field of paleodemography, Bocquet-Appel and Masset were the first researchers to highlight this effect and showed that the age structure of the reference sample was reflected in the estimated age distribution of the target sample [21,22].

As our sample regards, individuals underwent a scan in the clinical context of general condition alteration or suspicion of abdominal diseases, which was a selection bias. Yet, they were not necessarily affected by a chronic disease or any organic pathology. Thus, this

population was quite naturally diversified and may in that sense be representative of the general population.

Finally, this original study including three observers enables to conclude according to a threefold perspective: firstly, on the opposite of our expectations, we did not highlight some differences between the first three phases and the last three ones regarding the inter-observer agreement; and secondly, the errors of estimation were slightly dependent on the experience of the observer.

Conflict of interest

The authors declare that they have no conflicts of interest.

References

1. Cunha E, Baccino E, Martrille L et al (2009) The problem of aging human remains and living individuals: a review. Forensic Sci Int 193:1-13

2. Fanton L, Gustin MP, Paultre U et al (2010) Critical study of observation of the sternal end of the right 4th rib. J Forensic Sci 55:467-472

3. Konigsberg LW, Hermann NP, Wescott DJ et al (2008) Estimation and evidence in forensic anthropology: age-at-death. J Forensic Sci 53:541-557

4. Savall F, Rérolle C, Hérin F et al (2016). Reliability of the Suchey-Brooks method for a French contemporary population. Forensic Sci Int 266:586.e1-586.e5

5. Brooks S, Suchey JM (1990) Skeletal age determination based on the os pubis: a comparison of the Acsadi-Nemeskeri and Suchey-Brooks methods. Hum Evol 5:227-238

6. Todd TW (1920) Age changes in the pubic bones, I: the white male pubis. Am J Phys Anthropol 3:285-334

7. McKern TW, Stewart TD (1957) Skeletal age changes in young American males. Quartermaster Research and Development Command Technical Report, Natick MA.

8. Hoppa RD (2000) Population variation in osteological aging criteria: an example from the pubic symphysis. Am J Phys Anthropol 111:185-191

9. Schmitt A (2004) Age-at-death assessment using the os pubis and the auricular surface of the ilium: a test on an identified Asian sample. Int J Osteoarchaeol 14:1-6

10. Baccino E, Tavernier JC, Lamendin H et al (1991) Recherche d'une méthode multifactorielle simple pour la détermination de l'âge des cadavres adultes. Journal de Médecine Légale et Droit Médical 34:27-33

11. Lovejoy CO, Meindl RS, Tague RG et al (1995) The senescent biology of the hominoid pelvis. Riv Antropol 73:31-49.

12. Lottering N, MacGregor DM, Meredith M et al (2013) Evaluation of the Suchey-Brooks method of age estimation in an Australian subpopulation using computed tomography of the pubic symphyseal surface. Am J Phys Anthropol 150:386-399

13. Villa C, Buckberry J, Cattaneo C et al (2013) Technical note: reliability of Suchey-Brooks and Buckberry-Chamberlain methods on 3D visualizations from CT and laser scans: age estimation methods and 3D visualizations. Am J Phys Anthropol 151:158-163

14. Telmon N, Gaston A, Chemla P et al (2005) Application of the Suchey-Brooks method to three-dimensional imaging of the pubic symphysis. J Forensic Sci 50:507-512

15. Acsadi G, Nemeskeri J (1970) History of human life span and mortality. Akademiai Kiado, Budapest.

16. San Millán M, Rissech C, Turbon D (2013) A test of Suchey-Brooks (pubic symphysis) and Buckberry-Chamberlain (auricular surface) methods on an identified Spanish sample: paleodemographic implications. J Archaeol Sci 40:1743-1751

17. Berg GE (2008) Pubic bone age estimation in adult women. J Forensic Sci 53:569-577

18. Hartnett KM (2010) Analysis of age-at-death estimating using data from a new, modern autopsy sample – part I: pubic bone. J Forensic Sci 55:1145-1151

19. Dubourg O, Faruch-Bilfeld M, Telmon N et al (2019) Correlation between pubic bone mineral density and age from a computed tomography sample. Forensic Sci Int 298:345-350

20. Bocquet-Appel JP, Masset C (1996) Paleodemography: expectancy and false hope. Am J Phys Anthropol 99:571-583

21. Bocquet-Appel JP, Masset C (1982) Farewell to paleodemography. J Hum Evol 11:321-333

22. Usher B (2002) Reference samples: the first step in linking biology and age in the human skeleton. In: Hoppa RD, Vaupel JW (eds) Paleodemography: age distributions from skeletal samples. Cambridge University Press, New York, pp 29-47

Tables

Table 1: Inter-observer reliabilities for males (with and without early and advanced patterns).

Pairs of observers	Weighted Cohen's kappa (phases **without** early and advanced patterns)	Weighted Cohen's kappa (phases **with** early and advanced patterns)
1-2	0.91	0.94
2-3	0.8	0.86
1-3	0.83	0.87

Table 2: Inter-observer reliabilities for females (with and without early and advanced patterns).

Pairs of observers	Weighted Cohen's kappa (for phases **without** early and advanced patterns)	Weighted Cohen's kappa (for phases **with** early and advanced patterns)
1-2	0.93	0.95
2-3	0.94	0.94
1-3	0.86	0.87

Table 3: Inter-observer reliabilities for males (with and without early and advanced patterns).

Pairs of observers	Weighted Cohen's kappa (phases **without** early and advanced patterns)		Weighted Cohen's kappa (phases **with** early and advanced patterns)	
	Phases I➔III	Phases IV➔VI	Phases I➔III	Phases IV➔VI
1-2	0.72	0.77	0.84	0.83
2-3	0.61	0.44	0.74	0.49
1-3	0.45	0.59	0.62	0.59

Table 4: Inter-observer reliabilities for females (with and without early and advanced patterns).

Pairs of observers	Weighted Cohen's kappa (phases **without** early and advanced patterns)		Weighted Cohen's kappa (phases **with** early and advanced patterns)	
	Phases I➔III	Phases IV➔VI	Phases I➔III	Phases IV➔VI
1-2	0.83	0.71	0.83	0.85
2-3	0.83	0.82	0.83	0.82
1-3	0.56	0.56	0.59	0.63

Table 5: Inaccuracy and bias of estimated ages relative to the mean ages of the SB series for males.

	Observer 1		Observer 2		Observer 3	
Age range	Inaccuracy (years)	Bias (years)	Inaccuracy (years)	Bias (years)	Inaccuracy (years)	Bias (years)
10-19	4.46	4.8	3.21	3.13	6.65	6.56
20-29	3.77	2.62	4.02	4.02	3.77	0.88
30-39	4.52	2.6	2.76	-0.47	6.3	5.55
40-49	10.2	-10.2	9.55	-9.55	6.95	-6.95
50-59	15.6	-14.72	14.3	-13.42	18.9	-18.9
60-69	19.95	-19.95	25.47	-25.47	21.25	-21.25
70-79	30.73	-30.73	30.73	-30.73	34.2	-34.2
80-89	29.35	-29.35	29.35	-29.35	29.35	-29.35
Entire range	10.86	-5.98	10.36	-7.12	12.05	-5.62

Table 6: Inaccuracy and bias of estimated ages relative to the mean ages of the SB series for females.

	Observer 1		Observer 2		Observer 3	
Age range	Inaccuracy (years)	Bias (years)	Inaccuracy (years)	Bias (years)	Inaccuracy (years)	Bias (years)
10-19	11	11	5.28	5.28	3.18	3.18
20-29	3.63	2.38	3.63	2.38	3.4	1.13
30-39	5.1	5	1.8	1.7	3.93	-0.5
40-49	4.58	-1.51	5.86	-3.16	6.95	-4.25
50-59	11.66	-10.11	11.82	-10.48	12.45	-11.12
60-69	19.77	-19.77	20.87	-20.87	23.35	-23.35
70-79	24.4	-24.4	29.35	-29.35	34.3	-34.3
80-89	28	-28	28	-28	28	-28
90-99	33.66	-33.66	35.65	-35.65	35.65	-35.65
Entire range	12.83	-6.37	12.23	-8.47	12.74	-9.93

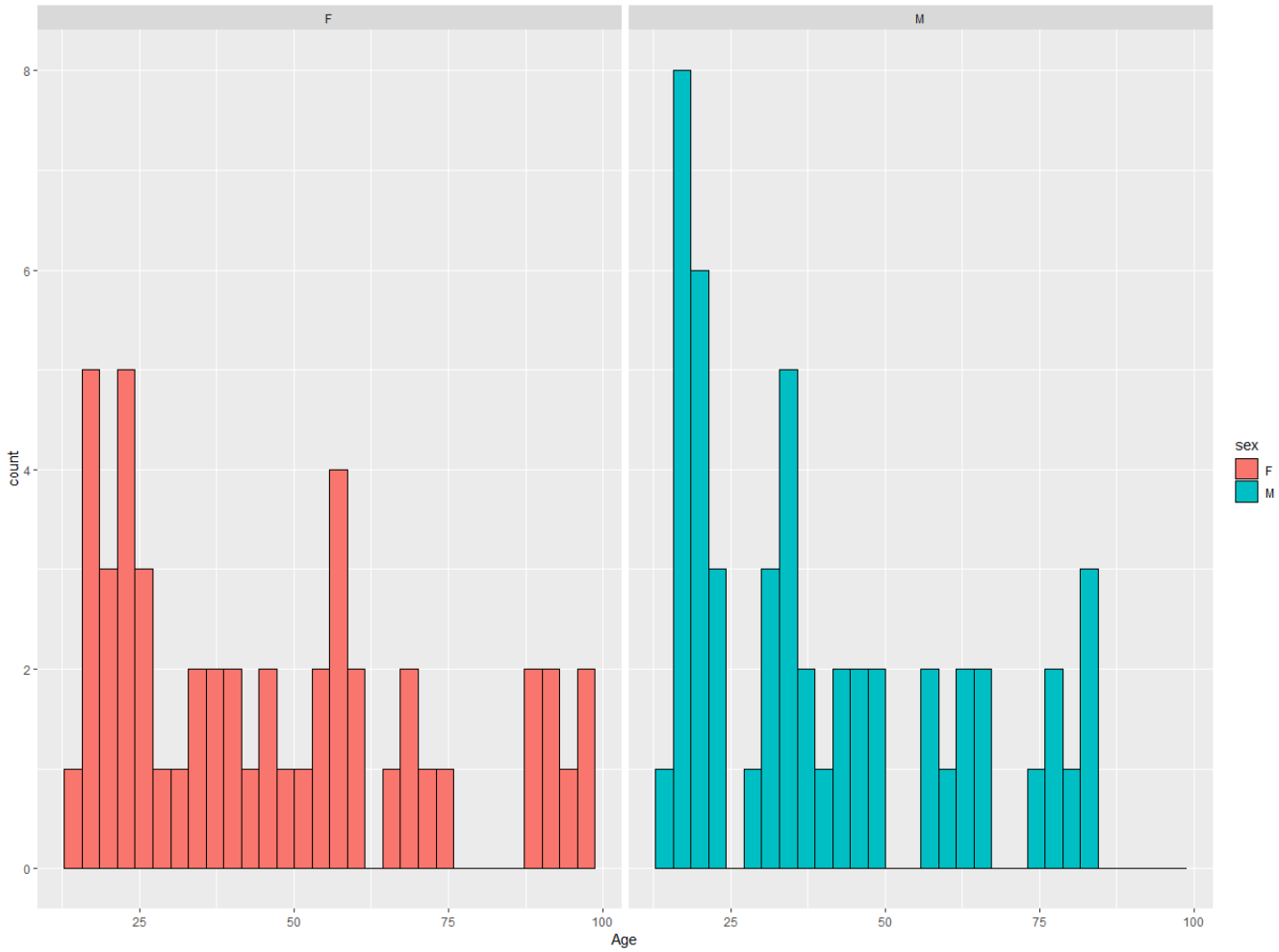


Fig. 1. Histogram showing the distribution by age of the study sample for males and females by 2-year intervals.

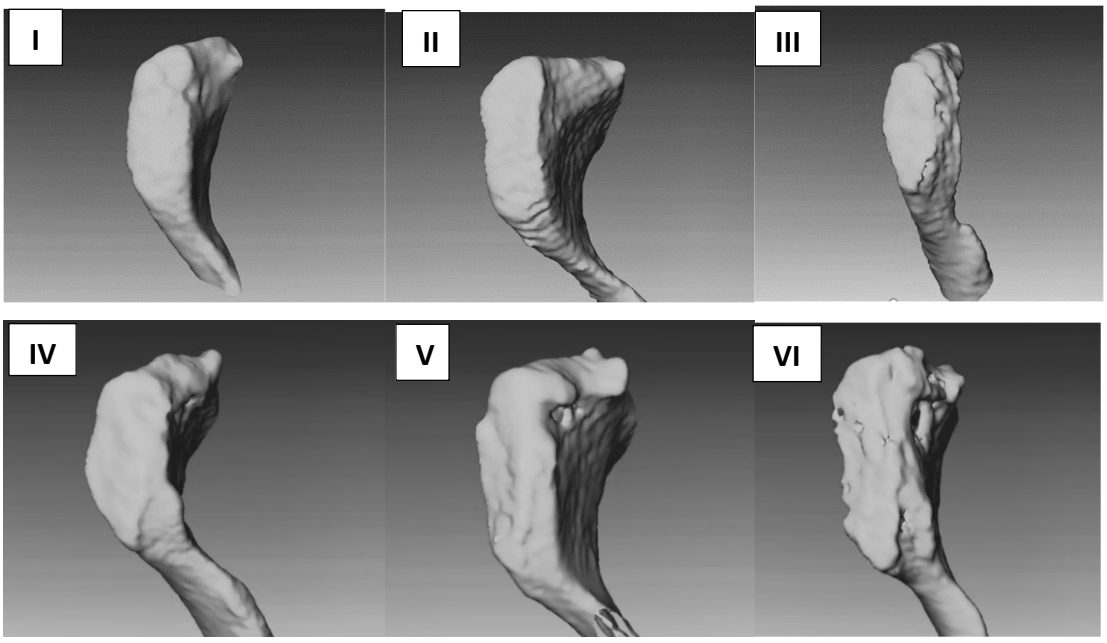
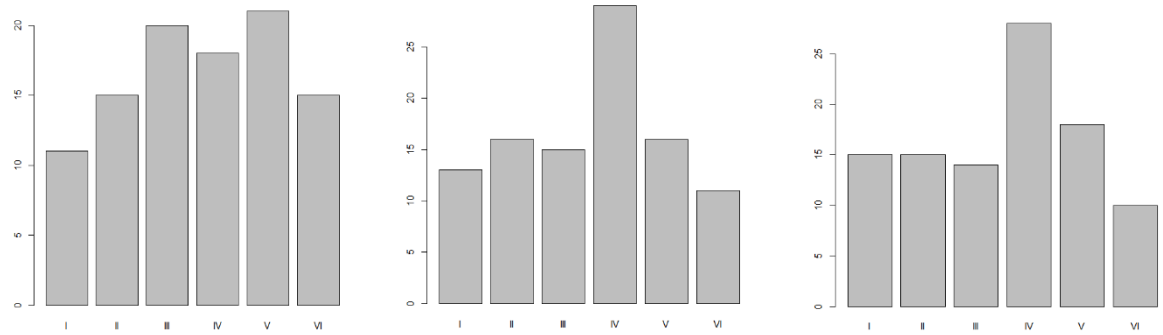


Fig. 2. Symphyseal faces of pubic bones representative of each stage using the morphological features of the SB method.



Observer 1 Observer 2 Observer 3

Fig. 3. Charts showing the distribution by allocated stages of the study sample for each observer.