



HAL
open science

Proposition d'une démarche commune pour la sélection de variations asymptomatiques osseuses et dentaires répétables et reproductibles en anthropologie biologique

Ariane Ducher, Bérengère Saliba-Serre, Laurine Renaud, Bruno Foti, Émeline Verna

► To cite this version:

Ariane Ducher, Bérengère Saliba-Serre, Laurine Renaud, Bruno Foti, Émeline Verna. Proposition d'une démarche commune pour la sélection de variations asymptomatiques osseuses et dentaires répétables et reproductibles en anthropologie biologique. *Bulletins et Mémoires de la Société d'anthropologie de Paris*, 2023, 35 (1), 10.4000/bmsap.11815 . hal-04082025

HAL Id: hal-04082025

<https://hal.science/hal-04082025v1>

Submitted on 5 Jul 2024

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 - NoDerivatives 4.0 International License

Proposition d'une démarche commune pour la sélection de variations asymptomatiques osseuses et dentaires répétables et reproductibles en anthropologie biologique

Proposal for a common approach for selecting repeatable and reproducible asymptomatic bone and dental variations in biological anthropology

Ariane Ducher, Bérengère Saliba-Serre, Laurine Renaud, Bruno Foti et Émeline Verna



Édition électronique

URL : <https://journals.openedition.org/bmsap/11815>

DOI : [10.4000/bmsap.11815](https://doi.org/10.4000/bmsap.11815)

ISSN : 1777-5469

Éditeur

Société d'Anthropologie de Paris

Référence électronique

Ariane Ducher, Bérengère Saliba-Serre, Laurine Renaud, Bruno Foti et Émeline Verna, « Proposition d'une démarche commune pour la sélection de variations asymptomatiques osseuses et dentaires répétables et reproductibles en anthropologie biologique », *Bulletins et mémoires de la Société d'Anthropologie de Paris* [En ligne], 35 (1) | 2023, mis en ligne le 17 avril 2023, consulté le 18 octobre 2023. URL : <http://journals.openedition.org/bmsap/11815> ; DOI : <https://doi.org/10.4000/bmsap.11815>



Le texte seul est utilisable sous licence CC BY-NC-ND 4.0. Les autres éléments (illustrations, fichiers annexes importés) sont « Tous droits réservés », sauf mention contraire.



BMSAP

Bulletins et Mémoires de la Société d'Anthropologie de Paris

VOLUME 35 # 1

AVRIL 2023

ISSN 0037-8984

Publication fondée en 1859



*Revue soutenue par l'Institut des sciences humaines et sociales (InSHS) du CNRS
2023-2024*

Proposition d'une démarche commune pour la sélection de variations asymptomatiques osseuses et dentaires répétables et reproductibles en anthropologie biologique

Proposal for a common approach for selecting repeatable and reproducible asymptomatic bone and dental variations in biological anthropology

Ariane Ducher ^{1*}, Bérengère Saliba-Serre ¹, Laurine Renaud¹, Bruno Foti ^{1,2}, Émeline Verna ¹

1 UMR 7268 ADES, Aix Marseille Univ, CNRS, EFS, Marseille, France

2 Faculté d'odontologie, Aix-Marseille Université, Marseille, France

* ariane.ducher@univ-amu.fr

Reçu : 27 juin 2022 ; accepté : 07 mars 2023
Bulletins et Mémoires de la Société d'Anthropologie de Paris

Résumé – Les variations asymptomatiques (VA) osseuses et dentaires sont grandement utilisées en anthropologie biologique, notamment pour appréhender les affinités intra- et inter-populationnelles. Cette étude se concentre sur la première étape de toute étude : la répétabilité et la reproductibilité de la méthode mise en place. À partir de l'utilisation du coefficient de Kappa la fiabilité de 198 variables réparties en trois zones anatomiques (crâne, post-crâne et dents) a été analysée. L'échantillon utilisé a été sélectionné parmi les individus les mieux représentés et conservés de la collection ostéo-archéologique Notre-Dame-du-Bourg (Digneles-Bains, périodes médiévale et moderne). Les analyses ont été réalisées, suite à quatre séries de cotations (trois observatrices), et lorsque plus de 20 observations sur des variables polymorphiques ont été obtenues. La qualité de l'accord a été établie selon l'échelle de Fleiss et al. (2003). Bien que les données manquantes aient été un frein dans cette étude, plusieurs propositions sont faites afin de palier ce phénomène. De plus, nous mettons en avant et discutons le fait que l'expérience a un impact sur l'observation des variables qualitatives. Au final, un protocole basé sur la sélection de VA répétables et reproductibles est proposé pour normaliser les comparaisons entre études dans un contexte aussi bien bioarchéologique que médico-légal.

Mots clés – caractères morphologiques, variations anatomiques, coefficient de Kappa, répétabilité, reproductibilité

Abstract – Asymptomatic variations (AV) in bones and teeth are widely used in biological anthropology, particularly to understand intra- and inter-population affinities. This study focuses on the first step of any study: the repeatability and reproducibility of the method implemented. Using the Kappa coefficient, the reliability of 198 variables divided into three anatomical areas (cranial, post-cranial and teeth) was analyzed. The sample used was selected among the

best represented and best preserved individuals of the Notre-Dame-du-Bourg osteo-archaeological collection (Digneles-Bains, medieval and modern period). The analyses were performed after four rounds of scoring (three observers), and when more than 20 observations of polymorphic variations were obtained. The quality of agreement was established according to the Fleiss et al. scale (2003). Although missing data were a hindrance in this study, several proposals were made to overcome the problem. We also highlight and discuss the fact that experience has an impact on the observation of qualitative variables. The outcome is a proposal for a protocol based on selecting repeatable and reproducible AVs to normalize comparisons between studies in both osteoarchaeological and forensic contexts.

Keywords – morphological characters, non-metric traits, Kappa coefficient, intra-observer agreement, inter-observer agreement

Introduction

Les caractères morphologiques osseux et dentaires sont étudiés, depuis les débuts de l'anthropologie biologique à la fin du XVIII^e siècle, notamment pour classer les groupes humains ou encore pour distinguer le "normal" de "l'anormal" (Hefner et al., 2016). Depuis les années 1950, leur utilisation connaît un essor important et ces caractères sont alors considérés comme des variations morphologiques. Ces variations, aussi connues sous le terme de traits non-métriques, peuvent être distinguées en deux grands groupes (Pink et al., 2016) : les traits macromorphoscopiques et les variations asymptomatiques. Les traits macromorphoscopiques (MMS), en plus d'être majoritairement utilisés dans des recherches en anthropologie médico-légale, sont d'après Hefner et Linde (2018) "des variations quasi-continues du crâne qui se traduisent par les différences (directes) au

niveau des tissus mous chez les individus vivants”. Les variations asymptomatiques (VA), aussi appelées caractères discrets, discontinus, épigénétiques, ou encore variations anatomiques (Sarfo, 2014 ; Verna et al., 2014), quant à elles ne se limitent pas au crâne, et n’ont pas, ou très rarement, un lien direct avec l’aspect des tissus mous chez les individus vivants.

La littérature sur ces variations morphologiques (MMS et VA) porte sur leur description et leur utilisation dans les différents champs de l’anthropologie biologique. Ces dernières sont ainsi utilisées dans les études intra- et inter-populationnelles portant sur l’estimation de la distance biologique au sens large, afin de permettre d’appréhender l’affinité intra- et inter-populationnelle, pour répondre à un large éventail de questions, depuis l’histoire des migrations des populations, leur évolution, ou encore à une estimation de l’origine biogéographique d’un individu (Saunders et Rainey, 2008 ; Sarfo, 2014 ; Verna, 2014 ; Hefner et al., 2016 ; Pilloud et al., 2016 ; Hefner et Linde, 2018).

Cette étude porte spécifiquement sur les variations asymptomatiques qui sont considérées par les chercheurs comme des proxies génétiques (Berry et Berry, 1972 ; Verna, 2014). Cette hypothèse remonte aux travaux de Grüneberg en 1963. Ce dernier démontre que ces variations, dont la manifestation suit la règle du “tout ou rien”, sont héritées comme des variants continus (Berry et Berry, 1972). L’héritage de ces variants est dû à l’action d’un grand nombre de gènes ayant chacun de petits effets additifs (Sjøvold, 1973). L’addition de ces facteurs génétiques et leurs interactions avec des facteurs environnementaux, jusqu’à un certain seuil, entraînent leur apparition phénotypique. Actuellement les rôles joués par le génome et l’environnement restent encore assez mal compris (Saunders et Rainey, 2008 ; Pink et al., 2016).

Les mécaniques génétiques précises de formation de ces variations sont mal connues. En revanche certaines composantes qui contrôlent la formation des os et des dents, et de certaines particularités ont été découvertes (Murshed, 2018) et peuvent influencer la présence de certaines VA osseuses hyperostotiques notamment (Eyal et al., 2019). L’os pouvant se modifier tout au long de la vie de l’individu (remodelage, etc.) il est alors plus difficile d’appréhender la part de l’influence génétique dans la présence des VA. Concernant les dents, l’impact aura lieu lors de sa formation (Jernvall et Jung, 2000 ; Kimura et al., 2015). Une fois que l’éruption dentaire est terminée, la dent en tant que telle ne se modifiera plus, et l’influence génétique s’arrête. Les seules modifications que subira ensuite la dent seront liées à l’environnement, telle que l’usure (Jernvall et Thesleff, 2012).

Les variations asymptomatiques sont largement utilisées dans les domaines de l’anthropologie biologique. Depuis le début des années 2000, la littérature s’oriente notamment sur un objectif de standardisation de l’étude des variations morphologiques dans un but comparatif (MMS avec Hefner et Linde (2018) ; les variations dentaires et le système ASU-DAS avec Turner et al. (1991) ou le système de l’Université de Freiburg avec Alt et al. (données non

publiées) ; les variations crâniennes avec les quatre articles de Hanihara et Ishida (2001a-d) ; et les variations post-crâniennes avec Verna (2014).

Les études de répétabilité et de reproductibilité sont une étape indispensable pour valider la méthode de collecte des données qui sera utilisée ultérieurement dans une recherche scientifique (Desquilbet, 2012). Pour la répétabilité, ce sera sous la définition de Desquilbet (2012) qu’elle sera utilisée, où l’opérateur, les conditions et la méthode de prise de données sont les mêmes, pour deux périodes de temps distincts. Pour la reproductibilité, la définition de Desquilbet (2012) de reproductibilité inter-observateurs sera retenue : les conditions et la méthode de prise de données sont les mêmes, mais réalisées par différents opérateurs.

Le coefficient de Kappa (κ), créé par Cohen en 1960, et son adaptation par Conger, en 1980, (Cohen, 1960 ; Cohen, 1968 ; Conger, 1980 ; Fleiss et al., 2003) ont été choisis comme méthode pour calculer l’accord intra- et inter-observateurs, à la suite d’une revue de la littérature (Hefner, 2009 ; Pilloud et al., 2019 ; Corron et al., 2021 ; Corron et al., 2022). Malgré ses limites bien connues (Pontius et Millones, 2011), il reste ici un choix pertinent, en particulier parce qu’il n’existe pas encore d’alternative réellement pertinente pour l’étude des variables ordinales. L’intérêt de l’utilisation du coefficient de Kappa sera discuté par rapport à d’autres coefficients utilisables pour l’évaluation de l’accord intra- et inter-observateurs. Ce travail propose un protocole permettant de sélectionner les VA les plus pertinentes en fonction de leur observabilité et de leur représentation dans un premier temps, et de leur répétabilité et reproductibilité (R&R) dans un second temps. Cette démarche basée sur une succession de critères spécifiques, pourra ainsi faciliter la comparaison d’études ultérieures utilisant des VA.

Matériel

Répartition des variations asymptomatiques sur le squelette

Afin de faciliter l’analyse des données, les VA ont été réparties selon trois zones anatomiques distinctes : crâne, post-crâne et dents.

La fiabilité de la cotation de 94 VA paires et impaires a été évaluée. Un tableau récapitulatif des 94 variations étudiées est disponible dans le matériel supplémentaire (tableau S1). Pour chaque variation, l’os ou la dent sur lequel elle est observée, est renseigné, ainsi que son abréviation, son type de cotation, son caractère pair ou impair et une référence bibliographique. De plus, les VA paires ont été cotées pour les côtés gauche et droit. Pour certaines VA présentes sur le rachis, elles ont été cotées soit par segment (cervical, thoracique et lombaire) soit pour chaque vertèbre prise séparément. Au total, 198 variables ont été cotées.

Quatre cas de figure sont présentés afin d’explicitier la prise de données :

- une variation impaire est située sur la ligne médiane du squelette. Sa cotation donnera lieu à une seule variable ;

- une variation paire est cotable à droite et à gauche. La cotation de cette variation donnera lieu à deux variables ;
- une variation peut être présente sur toutes les vertèbres cervicales (C1 à C7) de manière bilatérale. La cotation de cette variation donnera deux variables par vertèbre cervicale, soit 14 variables au total ;
- une variation impaire peut être présente sur l'ensemble du rachis. Dans cette étude, la cotation de cette variation donnera lieu à trois variables, une pour chaque segment (cervical, thoracique et lombaire).

Les données issues d'une collection ostéo-archéologique

Les données proviennent de l'étude de la collection ostéo-archéologique de la nécropole de la cathédrale de Notre-Dame-du-Bourg (NDB) à Digne-les-Bains (04) (Démians d'Archimbaud et al., 2010 ; Lantéri, 2016). Cette collection est composée de 729 individus ; lors de l'étude de répétabilité et de reproductibilité 469 individus avaient été étudiés. En fonction de la représentation et de la conservation différentielle des zones anatomiques étudiées, entre 33 et 35 individus ont été sélectionnées par zone. Au total, 77 individus composent l'échantillon utilisé pour l'étude R&R (tableau 1). Les estimations du sexe et de l'âge des individus adultes ont été réalisées à partir de l'os coxal (Schmitt, 2005 ; Brůžek et al., 2017). L'estimation de l'âge des individus immatures a été réalisée à partir de la maturation osseuse et dentaire (AlQahtani et al., 2010 ; Coqueugniot et al., 2010).

Méthodes

Protocole de prise de données

Quatre sessions de prise de données par trois observatrices différentes (O1, O2, O3) sur les 77 individus ont été réalisées.

Pour étudier la répétabilité, l'observatrice O1 a réalisé les deux prises de données à trois mois d'intervalle, ces sessions seront notées dans la suite de l'article O1.1 et O1.2.

Pour la reproductibilité, la prise de données O1.1 a été utilisée avec celles réalisées par les deux autres observatrices O2 et O3.

Pour l'observation des variations osseuses, du crâne et du post-crâne, les observatrices O1 et O2 sont expérimentées et l'observatrice O3 est novice. Pour l'observation des variations dentaires, l'observatrice O1 est expérimentée et les observatrices O2 et O3 sont novices.

Chaque observatrice a réalisé sa prise de données dans des conditions similaires. Une observation macroscopique des squelettes a été réalisée afin de coter les variations suivant les instructions de l'observatrice O1. O1 a établi un catalogue recensant et renseignant chaque variation osseuse et dentaire (définition, mode de cotation, photos).

Lors de la prise de données et de la cotation, deux catégories supplémentaires pouvaient être renseignées : variable non cotable (NA) et variable non observable (NO). Lorsque la variable n'est pas observable, par manque de représentation, du fait de l'absence de l'os ou de la dent sur lequel elle est cotée, la valeur "NA" est inscrite. Pour la suite de l'analyse un seuil de 80 % de cotabilité a été choisi. Quand une variable n'est pas observable, principalement en lien avec l'action des processus taphonomiques (mauvaise conservation de la corticale) ou de l'usure dentaire, la valeur "NO" est inscrite. Un seuil de 50 % d'observabilité a été choisi pour la suite de l'étude.

Analyse statistique

Les analyses statistiques ont été réalisées à l'aide du logiciel R v.4.1.1 (R Foundation for Statistical Computing, Vienna, Austria).

Pour analyser la qualité de l'accord entre les deux séries de cotations (O1.1 et O1.2) de la même observatrice, le choix s'est porté sur l'utilisation du coefficient de Kappa de Cohen (κ) (1)

$$\kappa = \frac{p_o - p_e}{1 - p_e} \quad (1)$$

avec p_o correspondant à la proportion d'accord observée et p_e la proportion globale d'accord attendue.

Zone anatomique	n	Sexe			Classe d'âge						
		M	F	Ind	[5-9]	[10-14]	[15-19]	[20-29]	[30-49]	[≥50]	Adulte
Crâne	35	17	7	11	1	3	0	3	10	11	7
Post-crâne	33	17	14	2	0	0	1	6	14	7	5
Dent	33	8	7	18	1	4	5	6	7	1	9
Total des individus*	77	30	22	25	2	5	5	12	22	14	17

n : nombre d'individus ; Ind : chez les adultes, sexe ostéologique indéterminé, ou indéterminable par absence ou mal conservation de l'os coxal ; chez les immatures, sexe non déterminable à cause de l'immaturité de l'os ; Adulte : absence de signe d'immaturité sur le squelette de l'individu, impossibilité d'estimer l'âge sur l'os coxal de par son absence ou sa mauvaise conservation ; * : pour au moins une des zones anatomiques / n : number of individuals ; Ind: for adult individuals, osteological sex undetermined, or undeterminable because of either absence or poor preservation of the coxal bone; for immature individuals, sex cannot be determined because of bone immaturity; Adult: no sign of immaturity on the individual's skeleton; impossible to estimate age from the coxal bone due to its absence or poor preservation; *: for at least one anatomical area

Tableau 1. Données bio-anthropologiques des individus de la collection ostéo-archéologique de la cathédrale de Notre-Dame-du-Bourg (Digne-les-Bains) étudiés pour l'étude de répétabilité et de reproductibilité / Bio-anthropological data for individuals from the osteo-archaeological collection of the cathedral of Notre-Dame-du-Bourg (Digne-les-Bains) studied for repeatability and reproducibility

Ce calcul a été réalisé à l'aide de la fonction *kappa2* du package *irr* pour évaluer la répétabilité de chacune des variations ; sans pondération pour les variations binaires et nominales, et avec une pondération quadratique dans le cas des variations avec une cotation ordinale (Cohen, 1960 ; Cohen, 1968).

Le calcul du coefficient de Kappa de Fleiss exact (κ') (2) a été utilisé pour évaluer la reproductibilité à trois observatrices (O1.1, O2 et O3) à l'aide de la fonction *kappam.fleiss* du package *irr* (Conger, 1980 ; Fleiss et al., 2003) :

$$\kappa' = \frac{\bar{p}_o - \bar{p}_e}{1 - \bar{p}_e} \quad (2)$$

avec \bar{p}_o : correspondant à la proportion globale d'accord observée et \bar{p}_e : la proportion globale d'accord attendue.

Des calculs du Kappa de Cohen (1) sans pondération ont également été effectués pour évaluer la reproductibilité pour chaque binôme d'observatrices. Les exports RMarkdown sont disponibles dans le matériel supplémentaire.

Les résultats sont considérés comme étant analysables quand ils ont été réalisés à partir d'au moins 20 paires d'observations complètes pour la répétabilité et d'au moins 20 triplets d'observations complètes pour la reproductibilité. Les variables non-polymorphiques, c'est-à-dire à expression constante pour au moins une des observatrices (Santos, 2018), ont été éliminées de l'analyse.

Pour les variables restantes, la qualité de l'accord a été analysée en suivant l'échelle mise en place par Fleiss et al. (2003) (tableau 2).

La répétabilité pourra aussi être trouvée dans cet article sous le terme d'accord intra-observateur, et la reproductibilité sous le terme d'accord inter-observateur.

Résultats de la répétabilité

Sur les 198 variables, 76 (soit 38,8 %) n'ont pas pu être analysées : 18 variables crâniennes, 25 post-crâniennes, et 33 dentaires (tableaux S2-S3). L'étude de répétabilité a donc été réalisée à partir de 122 variables (figure 1 ; tableau 3).

Le détail des valeurs de Kappa pour chaque variable analysée est disponible dans le tableau S4.

Les variables crâniennes

Parmi les 56 variables crâniennes, toutes comptent au moins 23 paires d'observations complètes et 18 ont été retirées de l'analyse du fait de leur caractère non-polymorphique pour au moins une des deux séries de prises de données. 38 variables ont été analysées (figure 1 ; tableau S4) : 50,0 % (19/38) des variables présentent des $\kappa \geq 0,75$, 42,1 % (16/38) des variables ont un κ compris entre 0,40 et 0,75 et 7,9 % (3/38) des variables présentent un $\kappa < 0,40$.

Valeur de Kappa	Qualité de l'accord
$\kappa < 0,40$	Pauvre
$0,4 \leq \kappa < 0,75$	Moyen à bon
$\kappa \geq 0,75$	Excellent

Tableau 2. Qualité de l'accord en fonction de la valeur du coefficient de Kappa selon l'échelle de Fleiss et al. (2003) / *Quality of agreement between observers based on the Kappa coefficient according to the scale of Fleiss et al. (2003)*

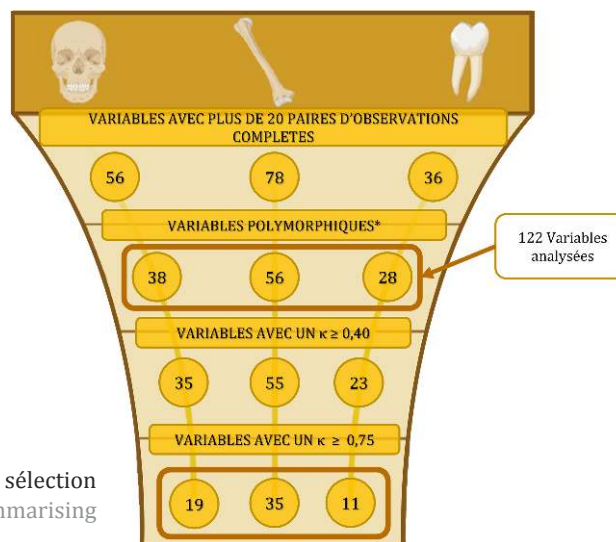


Figure 1. Diagramme en entonnoir récapitulatif du processus de sélection des variables pour l'analyse de répétabilité / *Funnel diagram summarising the variables selection process for repeatability analysis*

Étendue	Crâne				Post-crâne			Dents			
	Base	Face	Voûte	Mandibule	App. supérieur	Axial	App. inférieur	Incisive	Canine	Prémolaire	Molaire
Minimum	0,283	-0,062	0,632	0,298	-0,033	0,592	0,426	0,355	-0,048	0,298	0,39
Maximum	0,891	1	1	0,685	1	1	1	0,838	1	0,826	1
Nb. de variables	14	11	9	4	15	26	15	8	3	3	14

App. : Appendiculaire / *App.: Appendicular*

Tableau 3. Récapitulatif des résultats de l'accord intra-observateur pour l'observatrice O1, en fonction des différentes zones anatomiques / *Summary of findings on intra-observer agreement for the O1 observer, based on the different anatomical areas*

Les variables post-crâniennes

Parmi les 81 variables post-crâniennes, 3 possèdent moins de 20 paires d'observations complètes et ont été retirées de l'analyse, ainsi que 22 variables non-polymorphiques pour au moins une des deux séries de prises de données. 56 variables ont pu être analysées (figure 1 ; tableau S4) : 62,5 % (35/56) des variables ont des $\kappa \geq 0,75$, 35,7 % (20/56) des variables ont un κ compris entre 0,40 et 0,75 et 1,8 % (1/56) des variables présentent des $\kappa < 0,40$.

Les variables dentaires

Parmi les 61 variables dentaires, 25 possèdent moins de 20 paires d'observations complètes et ont été retirées de l'analyse, ainsi que huit variables non-polymorphiques pour au moins une des deux séries de prises de données. 28 variables ont pu être analysées (figure 1 ; tableau S4) : 39,3 % (11/28) de ces variables ont un $\kappa \geq 0,75$, 42,8 % (12/28) des variables ont un κ compris entre 0,40 et 0,75 et 17,9 % (5/28) des variables présentent un $\kappa < 0,40$.

Résultats de la reproductibilité

Dans l'ensemble, les résultats présentent des valeurs de kappa moins élevées que celles de la répétabilité (tableau 4). Tout d'abord, sur les 198 variables sélectionnées sur l'ensemble du squelette et des dents, 102 n'ont pas été retenues pour l'évaluation de la reproductibilité (tableaux S5-S6), soit 51,5 % de nos variables (25 variables crâniennes, 35 variables post-crâniennes, et 42 variables dentaires ; figure 2 ; tableau S7). De plus, parmi les 96 variables retenues, 21 (soit 21,9 %) ont un $\kappa' \geq 0,75$, contre 64 sur les 122 (52,5 %) variables analysées pour la répétabilité.

Les variables crâniennes

Parmi les 56 variables crâniennes, une variable possède moins de 20 triplets d'observations complètes et a été retirée de l'analyse, ainsi que 24 variables non-polymorphiques pour au moins une des trois séries de prises de données.

31 variables ont pu être analysées (figure 2 ; tableau S7) : 12,9 % (4/31) des variables de VA ont un $\kappa' \geq 0,75$, 45,2 % (14/31) des variables ont un κ' compris entre 0,40 et 0,75 et 41,9 % (13/31) des variables présentent des $\kappa' < 0,40$.

Les variables post-crâniennes

Parmi les 81 variables post-crâniennes, 3 possèdent moins de 20 triplets d'observations complètes et sont retirées de l'analyse, ainsi que 12 variables non-polymorphiques pour au moins une des trois séries de prises de données. 46 variables ont été analysées (figure 2 ; tableau S7) : 37 % (17/46) de ces variables ont un $\kappa' \geq 0,75$, 54,3 % (25/46) des variables ont un κ' compris entre 0,40 et 0,75 et 8,7 % (4/46) des variables présentent des $\kappa' < 0,40$.

Les variables dentaires

Parmi les 61 variables dentaires, 36 possèdent moins de 20 triplets d'observations complètes et ont été retirées de l'analyse, ainsi que six variables non-polymorphiques pour au moins une des trois séries de prises de données. 19 variables ont été analysées (figure 2 ; tableau S7) : Aucune variable ne présente un $\kappa' \geq 0,75$, 15,8 % (3/19) des variables ont un κ' compris entre 0,40 et 0,75 et 84,2 % (16/19) des variables présentent des $\kappa' < 0,40$.

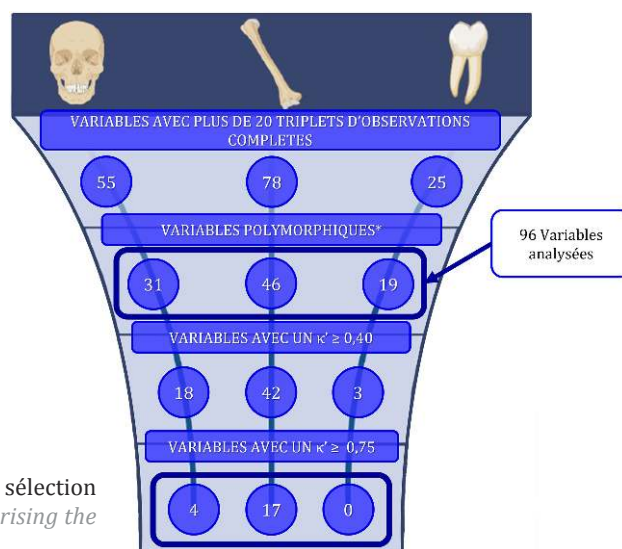


Figure 2. Diagramme en entonnoir récapitulatif du processus de sélection des variables pour la reproductibilité / *Funnel diagram summarising the variables selection process for reproducibility analysis*

Étendue	Crâne				Post-crâne			Dents		
	Base	Face	Voûte	Mandibule	App. supérieur	Axial	App. inférieur	Incisive	Prémolaire	Molaire
Minimum	0,064	-0,061	0,291	0,089	0,474	0,592	0,318	-0,076	0,071	-0,033
Maximum	0,646	1	1	0,432	1	0,871	1	0,505	0,338	0,575
Nb. de variables	9	11	7	4	10	22	14	8	2	9

App. : Appendiculaire / *App.: Appendicular*

Tableau 4. Récapitulatif des résultats de l'accord inter-observateurs pour les trois observatrices, en fonction des différentes zones anatomiques / *Summary of findings on inter-observer agreement for the three observers, according to the different anatomical areas*

Les résultats de reproductibilité obtenus entre les trois binômes d'observatrices

Les trois binômes présentent un nombre plutôt équivalent de variables pour lequel l'accord est "Excellent". En revanche, la répartition des résultats pour les variables pour lesquelles l'accord est "Moyen" et "Pauvre" entre les binômes O1.1 vs O3 et O2 vs O3 est inversée. Le nombre de variables pour lesquelles l'accord est "Moyen" est équivalent entre les binômes O1.1 vs O2 et O1.1 vs O3 (tableau 5).

	O1.1 vs O2 (%)	O1.1 vs O3 (%)	O2 vs O3 (%)
Accord Excellent	22 (22,9%)	25 (26,0%)	27 (28,1%)
Accord Moyen	39 (40,6%)	40 (41,7%)	29 (30,2%)
Accord Pauvre	35 (36,5%)	31 (32,3%)	40 (41,7%)

Tableau 5. Répartition des 96 variables retenues pour l'étude de la reproductibilité selon la qualité de l'accord pour chaque binôme d'observatrices en fonction de l'échelle de Fleiss et al. (2003) / *Distribution of the 96 variables selected for the reproducibility study according to the quality of agreement for each pair of observers according to the scale of Fleiss et al. (2003)*

Variations répétables et reproductibles

Les résultats de répétabilité et de reproductibilité pris en compte simultanément permettent de mettre en évidence 19 variables avec un kappa supérieur ou égal à 0,75. Ceci correspond à 11 variations asymptomatiques à la fois répétables et reproductibles (tableau 6).

Discussion

Les résultats de répétabilité (figure 2) montrent une bonne cohérence dans cette étude, avec plus de 50 % des variables étudiées ayant une valeur de kappa supérieure ou égale à 0,75 et dont la grande majorité est représentée par des variations osseuses. Ainsi, le protocole de prise de données peut être considéré comme bon.

Les variables osseuses présentent une meilleure répétabilité et reproductibilité que les variables dentaires. L'explication est en lien avec le type de cotation. La cotation binaire (exclusive pour les VA osseuses) s'avère moins subjective que la cotation catégorielle (nominale ou ordinale) qui est majoritaire pour les VA dentaires. Une observation dichotomique est plus facile à reproduire qu'une observation graduée. En effet, ce mode d'observation augmente le risque de désaccord. Ce constat ne fait que renforcer les données de la littérature (Pilloud et al., 2019 ; Corron et al., 2022). Pour diminuer le risque de désaccord, il semble important d'avoir en amont un temps d'échange, pendant lequel les observateurs puissent s'entendre sur la prise de données (B-Hardy, 2015). Il faut donc mettre en place un temps de formation spécifique sur la prise de données des variations asymptomatiques, en particulier les variations dentaires, à partir de référentiels connus.

En effet, lors de la deuxième prise de données de la première observatrice (O1.2), un gain d'expérience est constaté pour les VA osseuses et dentaires. O1 s'est formée pendant ses études en anthropologie biologique (master et doctorat en cours) sur les variations asymptomatiques osseuses et dentaires. Par ailleurs, les résultats de reproductibilité du binôme O1.1-O2 confrontés à ceux du binôme O1.2-O2 montrent que la reproductibilité pour les variables osseuses est meilleure contrairement à la reproductibilité pour les variables dentaires qui est moins bonne. O2 est une chercheuse dont la spécialité est l'étude des VA osseuses et non dentaires. Les résultats sont en cohérence avec un gain d'expérience avec le temps. La formation à l'étude des traits macroscopiques et qualitatifs en général est souvent considérée comme importante pour la réalisation d'études bio-anthropologiques fiables et reproductibles (Kamnkar et al., 2018).

Un autre point semble important à relever et à discuter, la conséquence de la présence de données manquantes sur les échantillons ostéo-archéologiques. 76 variables n'ont pas pu être évaluées pour la répétabilité, soit 38,4 % des variables ; et 102 variables pour la reproductibilité, soit 51,5 %.

Zone anatomique	Variations asymptomatiques
3 variations crâniennes	Suture métopique complète Épine trochléaire Foramen pariétal
8 variations post-crâniennes	Perforation du nerf supra-claviculaire Foramen supra-scapulaire Perforation humérale Perforation sternale Foramen transverse double des vertèbres cervicales Foramen transverse incomplet des vertèbres cervicales Facette articulaire supéro-antérieure double du calcaneum Facette articulaire distale du cunéiforme médial

Tableau 6. Répartition anatomique des variations asymptomatiques répétables et reproductibles / *Anatomical distribution of repeatable and reproducible asymptomatic variations*

La cause principale des données manquantes pour les VA dentaires est l'usure et la perte *ante mortem* avant même l'action des processus taphonomiques, qui, elle, impactera majoritairement les VA osseuses. Afin de pallier ces effets limitants, deux propositions sont formulées ici, appuyées par d'autres études de la littérature. Concernant l'évaluation de la R&R, pour les VA dentaires il s'avère indispensable d'inclure un critère d'âge dans la sélection des individus. La classe d'âge à retenir se situe entre 15 et 30 ans afin d'avoir l'ensemble des dents permanentes en bouche (excepté la 3^e molaire) (AlQahtani et al., 2010 ; Liversidge, 2016) et un niveau d'usure dentaire relativement faible (score ≤ 2) (Burnett et al., 2013 ; Burnett, 2016 ; Schmidt, 2016). Dans l'échantillon sélectionné ici, pour l'analyse des VA dentaires, les individus dont l'âge varie entre 15 et 30 ans représentent seulement 42 %.

Concernant les VA osseuses pour limiter l'impact des données manquantes, inhérentes à l'étude de collections ostéo-archéologiques (Bello et al., 2003), la solution proposée est d'augmenter la taille de l'échantillon sur lequel sont réalisées les études de R&R (Corron et al., 2022). Il faut s'intéresser en début d'étude à la taille de l'échantillon à sélectionner afin d'obtenir une puissance statistique élevée. Cette solution pourrait également diminuer le nombre de variations considérées comme non-polymorphiques dans l'échantillon au vu de leur prévalence rare, voire très rare, et de pouvoir ainsi utiliser le coefficient de Kappa efficacement (Desquilbet, 2012 ; Bujang et Baharum, 2017 ; Santos, 2018 ; Obertová et Stewart, 2020). Cependant, les études de R&R réalisées sur un effectif supérieur à 100 individus restent rares (Klales et Kenyhercz, 2015 ; Lewis et Garvin, 2016 ; Lesciotto et Doershuk, 2018).

D'un point de vue statistique, le choix s'est porté sur le coefficient de Kappa, coefficient le plus utilisé dans la littérature sur les données qualitatives en anthropologie biologique (Kamnkar et al., 2018 ; Corron et al., 2022), ce qui permet par la suite de faciliter les comparaisons entre les études. De plus, il est utilisable peu importe la nature des variables qualitatives. Le coefficient alpha de Krippendorff (α) est lui davantage utilisé sur des variables binaires et nominales, et le coefficient de concordance de Kendall (W), de même que le coefficient de corrélation intra-classe (ICC), sur des variables ordinales ayant au moins trois niveaux, ou lorsque le nombre d'observateurs est supérieur à deux (Pilloud et al., 2019 ; Zeman et Beňuš, 2020 ; de Raadt et al., 2021 ; Corron et al., 2022). Le coefficient Kappa de Cohen présente toutefois un inconvénient lorsque la prévalence de la variation est rare dans l'échantillon. En effet, le résultat du coefficient de Kappa sera réduit de façon importante, pour une petite diminution du pourcentage d'accord observé (P_o). Une réduction notable de la valeur du Kappa est obtenue lorsqu'un déséquilibre important des prévalences pour les cotations binaires est observé par rapport au pourcentage d'accord attendu (P_e). Ceci n'a pas lieu lorsque les fréquences des deux cotations sont similaires (Zeman et Beňuš, 2020).

Cette influence de la prévalence sur les résultats du Kappa se remarque notamment pour les traits bilatéraux. En se concentrant sur nos résultats de répétabilité, deux cas de figure existent. Tout d'abord des variations rares qui n'apparaissent pas sur un des côtés, ce qui a été observé pour six variations. Ensuite des variations rares pour lesquelles très peu de désaccord entre O1.1 et O1.2 est observé ; mais un seul désaccord supplémentaire entre la prise de données du côté gauche ou du côté droit, entraîne une diminution notable de la valeur du kappa. Ce cas de figure a été observé pour dix variations. Le foramen transverse incomplet de l'atlas (FTI1) représente bien ce cas. Aucun désaccord à droite n'est présent, le kappa est de 1 et l'accord est donc "Excellent". Un désaccord est observé à gauche ce qui donne un kappa à 0,65 donc un accord "Bon".

Concernant l'interprétation du Kappa de Cohen, dans les études de répétabilité et de reproductibilité, l'échelle de la qualité d'accord établie par Landis et Koch (1977) est souvent favorisée, avec $\kappa > 0,60$ traduisant une bonne qualité d'accord (L'Abbé et al., 2011 ; Klales et Kenyhercz, 2015 ; Kamnkar et al., 2018 ; Pilloud et al., 2019). Pour cette étude, l'échelle de Fleiss et al. (2003) a été privilégiée car elle est plus stricte (Cicchetti, 1994 ; Fleiss et al., 2003). Avec cette échelle, l'accord entre les observatrices est considéré comme "Bon" voire "Excellent" quand le kappa est $\geq 0,75$. De plus, d'après McHuch (2012), la valeur du Kappa obtenu nous informe sur le pourcentage de fiabilité des données acquises, respectivement 35 % avec un Kappa de 0,60, et 56 % avec un Kappa de 0,75. Ce critère semble très pertinent notamment dans les applications en anthropologie médico-légale (McHugh, 2012 ; Shirley et Ramirez Montes, 2015 ; Adalian, 2020). En effet, en anthropologie médico-légale, la validité des résultats obtenus est évaluée par la mesure de la fiabilité et de la précision des méthodes utilisées pour obtenir ces résultats (Adalian, 2020). Par conséquent, l'utilisation du coefficient de Kappa est à favoriser, car il est utilisable peu importe la nature des variables qualitatives, et peu importe le nombre de séries d'observations incluses dans l'analyse.

À la suite des différents points évoqués et discutés, une démarche commune est proposée afin de sélectionner ces VA (caractères qualitatifs) sur les mêmes critères en testant leur répétabilité et reproductibilité. Cette démarche (figure 3) permettra de rendre plus facilement comparables les études sur différents échantillons (d'époques différentes) tout en prenant en compte l'état de conservation des individus.

Conclusion

Cette étude avait pour objectif d'évaluer la répétabilité et la reproductibilité de 94 variations asymptomatiques osseuses et dentaires afin de proposer une démarche normalisée pour les futures études. De ces 94 variations asymptomatiques, 198 variables ont été observées sur une trentaine d'individus. Sur les 122 variables analysées pour la répétabilité, 52,5 % ont un kappa supérieur ou égal à 0,75.

Protocole de sélection de caractères qualitatifs, de type VA, par l'évaluation de la Répétabilité et de la Reproductibilité de leur prise de données

Choix des caractères qualitatifs d'intérêt dans une étude d'anthropologie biologique :

- Variations asymptomatiques,
- Traits macromorphoscopiques,
- Marqueurs de posture,...

Standardiser la prise de données

- Définition, illustrations et prévalence (locale, régionale et/ou mondiale)
- Critères de cotation :
 - Binaire, nominale ou ordinale
 - Définir un stade seuil de présence

Constituer un échantillon pour l'étude R&R

- Déterminer le nombre d'individus idéal en fonction de la prévalence des caractères
 - Guide pratique de Desquilbet [v.12.9] pour les variables binaires (Desquilbet, 2012)
- Privilégier les individus jeunes (14-29 ans) pour maximiser les observations
- Bon état de représentation et de conservation de l'os ou de la dent

Critères d'inclusion des variables dans l'analyse statistique

- Nombre de n-uplets (paires, triplets, ...) d'observations complètes suffisant pour la R&R
- Variables polymorphiques

Analyse statistique

- Utilisation du Kappa de Cohen pour 2 séries de cotations :
 - Sans pondération pour les variables binaires ou nominales
 - Avec pondération quadratique pour les variables ordinales
- Utilisation du Kappa de Fleiss pour plus de 2 séries de cotations
- Qualifier la qualité de l'accord grâce à l'échelle de Fleiss et al. (2003)

Normaliser les études sur les VA afin de faciliter leur comparaison

Figure 3. Protocole d'étude amélioré de caractères qualitatifs en fonction de la qualité de leur répétabilité et reproductibilité / Improved protocol for the selection of qualitative variables according to the quality of their repeatability and reproducibility

Concernant la reproductibilité, sur les 96 variables analysées, 21,9 % ont un $kappa \geq 0,75$. Au total, 19 variables sont répétables et reproductibles, ce qui correspond à 11 variations asymptomatiques. Le nombre s'avère faible mais plusieurs éléments ont réduit le nombre de variables avant même l'analyse (données manquantes, variables non-poly-morphiques...). Finalement ce résultat n'est qu'un exemple. Il ne s'agit pas de conclure que seules ces 11 VA sont exploitables par la suite. Le nombre variera en fonction des caractéristiques et spécificités des différentes études. À partir de la démarche commune proposée, les données obtenues seront comparables peu importe l'état de conservation de l'échantillon, l'époque, ou l'origine géographique.

Matériel supplémentaire

Le matériel supplémentaire, les données brutes, les exports RMarkdown sont disponibles respectivement au format .docx, .csv et .html sur le site des BMSAP (<https://journals.openedition.org/bmsap/11864?file=1>).

Références

- Adalian P (2020) General considerations about data and selection of statistical approaches. In: Obertová Z, Stewart A, Cattaneo C (eds) *Statistics and Probability in Forensic Anthropology*. Academic Press, London, pp 59-72
- AlQahtani SJ, Hector MP, Liversidge HM (2010) Brief communication: The London atlas of human tooth development and eruption. *American Journal of Physical Anthropology* 142(3): 481-490 [<https://doi.org/10.1002/ajpa.21258>]
- Alt KW, Pichler S, Vach W (Unpublished article) Dental traits for population studies: the Freiburg University dental trait system (FU-DTS)
- B-Hardy M-H (2015) Étude de la diversité des populations historiques de Montréal et de Québec par l'analyse de la morphologie dentaire : le cimetière catholique de la première église Notre-Dame (1691-1796) et le cimetière protestant Saint-Matthew de Québec (1771-1860). Mémoire, Université de Montréal, 152 p
- Bello S, Thomann A, Massa ER et al (2003) Quantification de l'état de conservation des collections ostéoarchéologiques et ses champs d'application en anthropologie. *Antropo* 5:21-37 [www.didac.ehu.es/antropo]
- Berry AC, Berry RJ (1972) Origins and relationships of the ancient Egyptians. Based on a study of non-metrical variations in the skull. *Journal of Human Evolution* 1(2):199-208 [[https://doi.org/10.1016/0047-2484\(72\)90020-6](https://doi.org/10.1016/0047-2484(72)90020-6)]
- Brůžek J, Santos F, Dutailly B et al (2017) Validation and reliability of the sex estimation of the human os coxae using freely available DSP2 software for archaeology and forensic anthropology. *American Journal of Physical Anthropology* 164(2):440-449 [<https://doi.org/10.1002/ajpa.23282>]
- Bujang MA, Baharum N (2017) Guidelines of the minimum sample size requirements for Kappa agreement test. *Epidemiology, Biostatistics and Public Health* 14(2):e12267-1-e12267-10 [<https://doi.org/10.2427/12267>]
- Burnett SE, Irish JD, Fong MR (2013) Wear's the problem? Examining the effect of dental wear on studies of crown morphology. In: Scott GR, Irish JD (eds) *Anthropological Perspectives on Tooth Morphology*. Cambridge University Press, Cambridge, pp 535-554
- Burnett SE (2016) Crown wear: Identification and categorization. In: Irish JD, Scott GR (eds) *A companion to dental anthropology*. Wiley Blackwell companions to anthropology, John Wiley & Sons, pp 415-432
- Cicchetti DV (1994) Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment* 6(4):284-290 [<https://doi.org/10.1037/1040-3590.6.4.284>]
- Cohen J (1960) A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 20(1):37-46 [<https://doi.org/10.1177/001316446002000104>]
- Cohen J (1968) Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. *Psychological Bulletin* 70(4):213-220 [<https://doi.org/10.1037/h0026256>]
- Conger AJ (1980) Integration and generalization of kappas for multiple raters. *Psychological Bulletin* 88(2):322-328 [<https://doi.org/10.1037/0033-2909.88.2.322>]
- Coqueugniot H, Weaver TD, Houët F (2010) Brief communication: A probabilistic approach to age estimation from infracranial sequences of maturation. *American Journal of Physical Anthropology* 142(4):655-664 [<https://doi.org/10.1002/ajpa.21312>]
- Corron LK, Stock MK, Cole SJ et al (2021) Standardizing ordinal subadult age indicators: Testing for observer agreement and consistency across modalities. *Forensic Science International* 320:110687 [<https://doi.org/10.1016/j.forsciint.2021.110687>]
- Corron LK, Broehl KA, Chu EY et al (2022) Agreement and error rates associated with standardized data collection protocols for skeletal and dental data on 3D virtual subadult crania. *Forensic Science International* 334:111-272 [<https://doi.org/10.1016/j.forsciint.2022.111272>]
- De Raadt A, Warrens MJ, Bosker RJ et al (2021) A comparison of reliability coefficients for ordinal rating scales. *Journal of Classification* 38(3):519-543 [<https://doi.org/10.1007/s00357-021-09386-5>]
- Démians d'Archimbaud G, Pelletier J-P, Flavigny F et al (2010) Notre-Dame du Bourg à Digne : Fouilles, restauration et aménagement liturgiques, une chronique monumentale de 20 siècles, Plan Patrimoine Antique, 133 p
- Desquilbet L (2012) Guide pratique de validation statistique de méthodes de mesure : répétabilité, reproductibilité, et concordance, 44 p
- Eyal S, Kult S, Rubin S et al (2019) Bone morphology is regulated modularly by global and regional genetic programs. *Development* 146:dev.167882 [<https://doi.org/10.1242/dev.167882>]
- Fleiss JL, Levin B, Paik MC (2003) Chapter 18 The measurement of interrater agreement. In: Shewart WA, Wilks SS, Fleiss JL et al (eds) *Statistical Methods for Rates and Proportions*, Third Edition. Wiley Series in Probability and Statistics, New Jersey, pp 589-626
- Grüneberg H (1963) *The pathology of development; a study of inherited skeletal disorders in animals*. Blackwell Scientific Publications, Oxford, 309 p

- Hanihara T, Ishida H (2001a) Frequency variations of discrete cranial traits in major human populations. I. Supernumerary ossicle variations. *Journal of Anatomy* 198(6):689-706 [https://doi.org/10.1046/j.1469-7580.2001.19860689.x]
- Hanihara T, Ishida H (2001b) Frequency variations of discrete cranial traits in major human populations. II. Hypostotic variations. *Journal of Anatomy* 198(6):707-725 [https://doi.org/10.1046/j.1469-7580.2001.19860707.x]
- Hanihara T, Ishida H (2001c) Frequency variations of discrete cranial traits in major human populations. III. Hyperostotic variations. *Journal of Anatomy* 199(3):251-272 [https://doi.org/10.1046/j.1469-7580.2001.19930251.x]
- Hanihara T, Ishida H (2001d) Frequency variations of discrete cranial traits in major human populations. IV. Vessel and nerve related variations. *Journal of Anatomy* 199(3):273-287 [https://doi.org/10.1046/j.1469-7580.2001.19930273.x]
- Hefner JT (2009) Cranial nonmetric variation and estimating ancestry. *Journal of Forensic Sciences* 54(5):985-995 [https://doi.org/doi:10.1111/j.1556-4029.2009.01118.x]
- Hefner JT, Pilloud MA, Buikstra JE et al (2016) A brief history of biological distance analysis. In: Pilloud MA, Hefner JT (eds) *Biological Distance Analysis Forensic and bioarchaeological perspectives*. Academic Press, London, pp 3-22
- Hefner JT, Linde KC (2018) *Atlas of Human Cranial Macromorphoscopic Traits*. Academic Press, London, 356 p
- Jernvall J, Jung H (2000) Genotype, phenotype, and developmental biology of molar tooth characters. *American Journal of Physical Anthropology* 43:171-190 [https://doi.org/10.1002/1096-8644(2000)43:31+%3C171::aid-ajpa6%3E3.0.co;2-3]
- Jernvall J, Thesleff I (2012) Tooth shape formation and tooth renewal: Evolving with the same signals. *Development* 139(19):3487-3497 [https://doi.org/10.1242/dev.085084]
- Kamnikar KR, Plemons AM, Hefner JT (2018) Intraobserver error in macromorphoscopic trait data. *Journal of Forensic Science* 63(2):361-370 [https://doi.org/10.1111/1556-4029.13564]
- Kimura R, Watanabe C, Kawaguchi A et al (2015) Common polymorphisms in WNT10A affect tooth morphology as well as hair shape. *Human Molecular Genetics* 24:2673-2680 [https://doi.org/10.1093/hmg/ddv014]
- Klales AR, Kenyhercz MW (2015) Morphological assessment of ancestry using cranial macromorphoscopies. *Journal of Forensic Science* 60(1):13-20 [https://doi.org/10.1111/1556-4029.12563]
- L'Abbé EN, Van Rooyen C, Nawrocki SP et al (2011) An evaluation of non-metric cranial traits used to estimate ancestry in a South African sample. *Forensic Science International* 209(1-3):195.e1-195.e7 [https://doi.org/10.1016/j.forsciint.2011.04.002]
- Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. *Biometrics* 33(1):159-174 [https://doi.org/10.2307/2529310]
- Lantéri L (2016) *Recrutement, paléodémographie et cémento-chronologie. Application à un contexte d'inhumation paroissial d'Ancien Régime : Notre-Dame-du-Bourg à Digne (04)*, Doctorat, Aix-Marseille Université, 351 p
- Lesciotto KM, Doershuk LJ (2018) Accuracy and reliability of the Klales et al. (2012) Morphoscopic Pelvic Sexing Method. *Journal of Forensic Sciences* 63(1):214-220 [https://doi.org/10.1111/1556-4029.13501]
- Lewis CJ, Garvin HM (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]
- Liversidge HM (2016) Tooth Eruption and Timing. In: Irish JD, Scott GR (eds) *A Companion to Dental Anthropology*. Wiley Blackwell, Malden, pp 159-171
- McHugh ML (2012) Interrater reliability: the kappa statistic. *Biochemia Medica* 22(3):276-282 [https://doi.org/10.11613/BM.2012.031]
- Murshed M (2018) Mechanism of bone mineralization. *Cold Spring Harbor Perspectives in Medicine* 8:a031229 [https://doi.org/10.1101/cshperspect.a031229]
- Obertová Z, Stewart A (2020) Probability distributions, hypothesis testing, and analysis. In: Obertová Z, Stewart A, Cattaneo C (eds) *Statistics and Probability in Forensic Anthropology*. Academic Press, London, pp 73-86
- Pilloud MA, Edgar HJJ, George R et al (2016) Dental morphology in biodistance analysis. In: Pilloud MA, Hefner JT (eds) *Biological Distance Analysis Forensic and bioarchaeological perspectives*. Academic Press, London, pp 109-133
- Pilloud MA, Adams DM, Hefner JT (2019) Observer error and its impact on ancestry estimation using dental morphology. *International Journal of Legal Medicine* 133(3):949-962 [https://doi.org/10.1007/s00414-018-1985-3]
- Pink CM, Maier C, Pilloud MA et al (2016) Cranial nonmetric and morphoscopic data sets. In: Pilloud MA, Hefner JT (eds) *Biological Distance Analysis Forensic and bioarchaeological perspectives*. Academic Press, London, pp 91-107
- Pontius RG, Millones M (2011) Death to Kappa: Birth of quantity disagreement and allocation disagreement for accuracy assessment. *International Journal of Remote Sensing* 32(15):4407-4429 [https://doi.org/10.1080/01431161.2011.552923]
- Santos F (2018) *AnthropMMD: An R package with a graphical user interface for the mean measure of divergence*. *American Journal of Physical Anthropology* 165(1):200-205 [https://doi.org/10.1002/ajpa.23336]
- Sarfo TA (2014) *A statistical investigation of nonmetric vertebral traits with a skeletal population sample from the Dakhleh Oasis, Egypt*. Electronic Thesis and Dissertation Repository, 142 p
- Saunders SR, Rainey DL (2008) Nonmetric trait variation in the skeleton: Abnormalities, anomalies, and atavisms. In: Katzenberg AM, Saunders SR (eds) *Biological Anthropology of the Human Skeleton*. John Wiley & Sons, pp 533-559
- Schmidt CW (2016) Estimating age, sex, and individual ID from teeth. In: Irish JD, Scott GR (eds) *A companion to dental anthropology*. Wiley Blackwell companions to anthropology. John Wiley & Sons, pp 362-376
- Schmitt A (2005) Une nouvelle méthode pour estimer l'âge au décès des adultes à partir de la surface sacro-pelvienne iliaque. *Bulletins et Mémoire de la Société d'Anthropologie de Paris* 17(1-2):89-101 [https://doi.org/10.4000/bmsap.943]
- Sjøvold T (1973) The occurrence of minor non-metrical variants in the skeleton and their quantitative treatment for population comparison. *HOMO - Journal of Comparative Human Biology* 24:204-233

- Shirley NR, Ramirez Montes PA (2015) Age estimation in forensic anthropology: Quantification of observer error in phase versus component-based methods. *Journal of Forensic Sciences* 60(1): 107-111 [<https://doi.org/10.1111/1556-4029.12617>]
- Turner CG, Nichol CR, Scott GR (1991) Scoring procedures for key morphological traits of permanent dentition: The Arizona State University dental anthropology system. In: Kelley MA, Larsen CS (eds) *Advances in Dental Anthropology*, New York, pp 13-31
- Verna E (2014) Les variations osseuses asymptomatiques du squelette postcranien : leur contribution à l'identification en Anthropologie médico-légale. Doctorat, Aix-Marseille Université, 409 p
- Verna E, Piercecchi-Marti M-D, Chaumoitre K et al (2014) Mise au point sur les caractères discrets du membre inférieur : définition, épidémiologie, étiologies. *Bulletins et Mémoires de la Société d'Anthropologie de Paris* 26:52-66 [<https://doi.org/10.1007/s13219-013-0090-x>]
- Zeman T, Beňuš R (2020) Initial assessment: Measurement errors and interrater reliability. In: Obertová Z, Stewart A, Cattaneo C (eds) *Statistics and Probability in Forensic Anthropology*. Academic Press, London, pp 47-56



BMSAP

Bulletins et Mémoires de la Société d'Anthropologie de Paris