

Article

Not peer-reviewed version

Validation of Greulich and Pyle Atlas for Radiological Bone Age Assessment in Pediatric Population from the Canary Islands

[Isidro Miguel Martín Pérez](#)*, [Sebastián Eustaquio Martín Pérez](#), [Jesús María Vega González](#), [Ruth Molina Suárez](#), [Alfonso Miguel García Hernández](#), Fidel Rodríguez Hernández, [Mario Herrera Perez](#)

Posted Date: 2 September 2024

doi: 10.20944/preprints202409.0071.v1

Keywords: diagnostic imaging; radiography; age determination by skeleton; children; Greulich and Pyle Atlas; Canary Islands



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Validation of Greulich and Pyle Atlas for Radiological Bone Age Assessment in Pediatric Population from the Canary Islands

Isidro Miguel Martín Pérez ^{1,2,*}, Sebastián Eustaquio Martín Pérez ^{1,2,3,4}, Jesús María Vega González ⁵, Ruth Molina Suárez ⁶, Alfonso Miguel García Hernández ¹, Fidel Rodríguez Hernández ² and Mario Herrera Perez ^{7,8}

¹ Escuela de Doctorado y Estudios de Posgrado, Universidad de La Laguna, San Cristóbal de La Laguna, 38203 Santa Cruz de Tenerife, Spain

² Departamento de Farmacología y Medicina Física, Área de Radiología y Medicina Física, Sección de Enfermería y Fisioterapia, Facultad de Ciencias de la Salud, Universidad de La Laguna, 38200 Santa Cruz de Tenerife, Spain

³ Musculoskeletal Pain and Motor Control Research Group, Faculty of Health Sciences, Universidad Europea de Canarias, 38300 Santa Cruz de Tenerife, Spain

⁴ Musculoskeletal Pain and Motor Control Research Group, Faculty of Sport Sciences, Universidad Europea de Madrid, 28670 Villaviciosa de Odón, Spain

⁵ Institute of Legal Medicine and Forensic Sciences of Santa Cruz de Tenerife, 38230 San Cristóbal de La Laguna, Spain

⁶ Pediatric Endocrinology Unit, Pediatric Department, Hospital Universitario de Canarias, San Cristóbal de La Laguna, 38320 Santa Cruz de Tenerife, Spain

⁷ School of Medicine (Health Sciences), Universidad de La Laguna, 38200 Santa Cruz de Tenerife, Spain

⁸ Foot and Ankle Unit, Orthopedic Surgery and Traumatology Department, San Cristóbal de La Laguna, 38320 Santa Cruz de Tenerife, Spain

* Correspondence: alu0100713374@ull.edu.es

Abstract: Bone age assessment measures the growth and development of children and adolescents by evaluating their skeletal maturity, which is influenced by various factors like heredity, ethnicity, culture and nutrition. Clinical standards for this assessment should be up-to-date and appropriate for the specific population being studied. This study validates the GP-Canary Atlas for accurately predicting bone age by analyzing posteroanterior left hand and wrist radiographs of healthy children (80 females and 134 males) from the Canary Islands across various developmental stages. The study found strong intra-rater reliability among all three raters, with Raters 1 and 2 indicating very high consistency (Intra-class coefficients = 0.990 to 0.996) and Rater 3 displaying slightly lower but still strong reliability (Intra-class coefficients = 0.921 to 0.976). Inter-rater agreement was excellent between Raters 1 and 2 but significantly lower between Rater 3 and the other two raters, with Intra-class coefficients of 0.408 and 0.463 for Rater 1 and 0.327 and 0.509 for Rater 2. The accuracy analysis revealed a substantial underestimation of bone age compared to chronological age for preschool (Mean Difference = 17.036 months, $p < 0.001$) and school-age males (Mean Difference = 13.298 months, $p < 0.001$). However, this was not observed in females, where the mean difference was minimal (3.949 months, $p < 0.239$). In contrast, the Atlas showed greater accuracy for teenagers, showing only a slight overestimation (Mean Difference = 3.159 months, $p = 0.823$). These results confirm the GP-Canary Atlas's overall precision but suggest caution when using it for preschool children, as it significantly underestimates bone age in this group.

Keywords: diagnostic imaging; radiography; age determination by skeleton; children; Greulich and Pyle Atlas; Canary Islands

1. Introduction

Maturation encompasses the physical and psychological development that occurs from childhood to adulthood [1–3]. Key indicators of biological maturation include sexual maturity [4–6], skeletal maturity [7–9], and morphological maturity [10]. Skeletal maturity is determined by a combination of genetic and environmental factors [11,12]. In optimal conditions, genetic factors account for approximately 80% to 90% of the maturation process; however, in less favorable environments, their influence can decrease to about 60% [13,14]. Various methods are available for assessing skeletal maturity [15–18], with radiographic analyses being among the most widely used [19]. One of the most common techniques for assessing skeletal maturity and predicting growth potential is the radiological evaluation of bone age (BA) using the Greulich and Pyle Atlas (GP Atlas) [11,20]. This method involves comparing the radiographic appearance of bones to standardized maturity levels for specific chronological age (CA) groups [21].

BA is influenced by a range of biological and sociocultural factors [22], including genetics, nutrition, socioeconomic status, and overall health. These factors can vary widely across different populations, leading to significant differences in bone maturation [23]. As a result, the normative data used in clinical practice must be current and specific to the population being assessed to ensure accurate evaluations [24]. The GP Atlas involves comparing children's posteroanterior left hand and wrist radiographs (PA-HW) to reference plates created from a study of white upper-middle-class children conducted between 1932 and 1942 [19]. This Atlas has been validated for use in various ethnic groups [25]. However, a recent systematic review indicates that, while the GP Atlas is generally considered reliable, its accuracy can vary and is not always consistent [26]. This inconsistency is particularly evident in the GP Atlas's tendency to underestimate BA in younger Caucasian and Hispanic populations, which may result in misinterpretations and potentially impact clinical decision-making. [27].

In the Canary Islands (Spain), a region with a diverse genetic composition influenced by indigenous Guanche ancestry, European settlers, and African and American admixture [28], combined with distinct environmental factors could lead to variations in bone development. Furthermore, the archipelago's geographical isolation and specific socio-cultural practices might also contribute to differences in growth patterns and maturation rates compared to other populations. As a result, the GP Atlas may not provide a precise and accurate BA for this pediatric population, suggesting that its applicability may be limited [29]. Therefore, to mitigate the potential biases associated with using the original GP Atlas in the pediatric population, a region-specific adaptation, the *Radiological Reference Atlas for Bone Age in the Canary Islands Population* (GP-Canary Atlas) [2], was launched in 2009, which compiles data from 1978 [24] and has since become the standard reference for pediatricians in the region.

Although the GP-Canary Atlas was developed to improve the BA assessments for the Canary Islands' pediatric population, it has several methodological limitations. These include the reliance on a single set of radiographic images without distinguishing between genders, the insufficient selection of specific radiographs for corresponding age stages, and the absence of formal validation processes. These shortcomings raise concerns about its validity for BA determinations and highlight the need for a thorough evaluation to confirm the Atlas's applicability in the pediatric population of the Canary Islands. Therefore, this study aims to rigorously validate the precision and accuracy of the GP-Canary Atlas for radiological BA assessment in this specific population.

2. Materials and Methods

2.1. Study Design

A cross-sectional study was undertaken between September 1, 2023, and June 20, 2024, within the Departments of Pediatrics and Orthopaedic and Trauma Surgery at Complejo Hospitalario Universitario de Canarias, a tertiary-level referral healthcare center in Tenerife, Spain. The study adhered to the STARD 2015 [30], which provides an updated checklist for reporting diagnostic accuracy studies, to ensure rigorous methodological and reporting standards. Ethical clearance was obtained from the Ethics Committee of Complejo Hospitalario Universitario de Canarias (reference

number CHUC_2023_86, approved on July 13, 2023). The study protocol was strictly in compliance with the ethical principles outlined in the Declaration of Helsinki.

In order to provide detailed contextual data for the analysis, sociodemographic variables—including age and gender—as well as anthropometric measurements such as height, weight and body mass index (BMI), were carefully extracted from the SAP Logon database (IBM®, United States) at the Complejo Hospitalario Universitario de Canarias (Tenerife, Canary Islands). In addition, standardized PA-HW radiographs, which were securely stored in the Centricity PACS system (GE HealthCare®, United States) at the same institution, were systematically analyzed for all participants. This was done according to a rigorously predefined protocol designed to ensure both consistency and precision in data collection and interpretation.

As part of this verification process, each PA-HW radiograph was meticulously reviewed to confirm that the patient's left hand was correctly positioned, with the fingers slightly spread and the wrist properly aligned with the forearm [31,32]. Moreover, the radiographs were carefully examined to ensure that all necessary anatomical landmarks, such as the phalanges, metacarpals, carpal bones, and distal radius and ulna, were clearly visible and appropriately captured. Additionally, the imaging settings—including exposure, focus, and contrast—were thoroughly checked to confirm strict adherence to the established protocol standards.

2.2. Participants

2.2.1. Inclusion and Exclusion Criteria

The inclusion and exclusion criteria for this study were carefully defined to ensure a representative and homogeneous sample of healthy children and adolescents from the Canary Islands, enabling precise assessment of BA using PA-HW radiographs. To be eligible for inclusion: (1) participants had to be healthy children aged 0 to 18 years who were long-term residents of the Canary Islands, defined as having resided there for a minimum of 5 years. Moreover, (2) at least one parent had to be of Canary Island origin, as verified through detailed medical and family history records, to ensure uniformity in the genetic background of the study population. Additionally, it was required (3) that subjects have medical records from 2016 onwards and (4) that their PA-HW radiographs adhered to predefined quality standards, including correct hand positioning, clear visibility of key anatomical landmarks and compliance with standardized imaging protocols.

The exclusion criteria were designed to eliminate any confounding factors that could affect normal bone maturation or impede the accuracy of BA estimation. Participants were excluded if they had (1) medical conditions known to alter bone development, such as endocrine-metabolic disorders (e.g., *growth hormone deficiency*, *hypothyroidism*, *hyperthyroidism*), neurological conditions (e.g., *cerebral palsy*, *muscular dystrophy*), or inherited disorders (e.g., *Down syndrome*, *Turner syndrome*, *Marfan syndrome*). Furthermore, (2) children undergoing medical treatments that could influence skeletal growth (e.g., *growth hormone therapy*, *corticosteroids*, or *chemotherapy*) were excluded from the study. PA-HW radiographs were also excluded if they demonstrated (3) fractures, significant skeletal abnormalities, or (4) were of poor quality, characterized by inadequate resolution, improper exposure, or obscured anatomical landmarks, which could interfere with the BA assessment procedure.

2.2.2. Sample Size Calculation

The sample size for this study was calculated to ensure precise and accurate BA measurements using the GP-Canary Atlas, with a 95% confidence level and a 5% margin of error. Due to the lack of standard deviation data for BA in the Canary Islands' population, estimates from similar studies in comparable populations were used to approximate expected variability [33–35]. Based on these estimates, a minimum of a total sample size of 100 subjects was determined to be sufficient for validating the GP Atlas in this pediatric population, ensuring reliable and generalizable findings.

2.3. Test Methods

During the evaluation process, three blinded raters independently assessed the PA-HW radiographs to ensure objective and unbiased BA determination. The raters included a Radiology expert (Rater 1), a General Practitioner (Rater 2) and a medical student (Rater 3), representing different levels of expertise and training in radiological interpretation. This diversity in raters was deliberately chosen to evaluate the impact of professional experience and training on the accuracy and consistency of BA assessments, thereby providing insights into the generalizability and robustness of the BA estimation method using the GP-Canary Atlas [2]. Each rater evaluated the radiographs by comparing the observed skeletal features with reference images from the GP-Canary Atlas. They used maturity indicators, such as ossification and bone fusion, to estimate the BA. If there was no exact match, the BA was estimated by averaging the ages of two consecutive radiographs from the Atlas. [2].

To assess Intra-rater precision, each rater determined BA at two distinct time points, T1 and T2, separated by less than one and a half months. The PA-HW radiographs were presented in a randomized and blinded sequence during both evaluations to minimize interpretation bias and prevent recall of previous assessments. This method allowed for a reliable examination of Intra-rater reliability by comparing the BA measurements from T1 with those from T2 for each rater, thereby assessing the consistency of each evaluator's assessments over time. With respect to Inter-rater precision, the BA determinations were compared across the three raters to analyze the level of agreement and reliability among different evaluators when interpreting the same set of PA-HW radiographs. This comparison was crucial to determine the reproducibility of the BA assessment method across raters with varying levels of expertise. Additionally, accuracy was determined by comparing the subjects' CA, calculated from the difference between their birth date and the date of the radiological exam, with their estimated BA through GP-Canary Atlas.

2.4. Analysis

Statistical analyses were conducted using IBM® SPSS Statistics software (United States). Descriptive statistics were first calculated for age (in *mos.*), weight (*kg*), height (*m*), and body mass index (BMI) (*kg/m²*). The data were stratified according to developmental stages as defined by Fraga and Fernández (2014) [36]—Preschool (1 to 5 years), School-age (5 to 12 years), and Teenager (12 to 18 years)—and further segmented by gender to account for potential differences in bone maturation between males and females. These descriptive measures included calculations of central tendency (*mean*) and dispersion (*standard deviation, minimum and maximum*) for both CA and BA as estimated by the study's method. To confirm the suitability of the data for further statistical analyses, the Shapiro-Wilk test was applied to assess the normality of the data distribution, while Levene's test was used to evaluate homoscedasticity.

For precision assessment, the Intra-class correlation coefficient (ICC) was calculated to evaluate both Intra-rater and Inter-rater agreement. The ICC provided a robust quantitative measure of consistency within and between raters, indicating the degree of agreement when using the GP-Canary Atlas for BA estimation. Bland-Altman plots were also constructed to visually assess Inter-rater reliability and detect any systematic bias or limits of agreement between the raters' BA measurements. Moreover, the accuracy of the BA estimations was evaluated through a mean difference analysis, comparing the discrepancies between the estimated BA and the actual CA of the children.

3. Results

3.1. Characteristics of Sample

A total of 214 PA-HW radiographs from healthy children were finally included consisting of 80 females and 134 males. In the preschool group, females had an average age of 39.33 *mos.* (SD = 15.18), weight of 14.52 *kg* (SD = 2.05), and height of 0.91 *m* (SD = 0.07), while males had an average age of 46.49 *mos.* (SD = 13.33), weight of 13.09 *kg* (SD = 2.17), and height of 0.94 *m* (SD = 0.05). In the school-age group, females averaged 92.00 *mos.* in age (SD = 26.08), 29.58 *kg* in weight (SD = 7.14), and 1.14

m in height (SD = 0.07), whereas males averaged 100.16 mos. in age (SD = 20.33), 23.67 kg in weight (SD = 4.85), and 1.16 m in height (SD = 0.05). In the teenage group, females and males both averaged 1.33 m in height, with females having an average age of 144.17 mos. (SD = 23.81) and weight of 33.84 kg (SD = 4.62), while males had an average age of 151.53 mos. (SD = 20.17) and weight of 34.21 kg (SD = 3.19). The Shapiro-Wilk test confirmed that all variables were normally distributed across these groups. More details in **Table 1**.

Table 1. Characteristics of sample. **Abbreviation:** BMI = Body Mass Index, mos = mos., Statistical significance: (*) $p < 0.05$, (**) $p < 0.01$, (***) $p < 0.001$. A p -value lower than these thresholds indicates a statistically significant deviation from normality.

	Stage	Gender	N	Mean	SD	Min	Max	p -value
Age (mos.)	Preschool	Female	24	39.33	15.18	20.00	67.00	0.235
		Male	45	46.49	13.33	18.00	69.00	0.105
	Scholar	Female	40	92.00	26.08	85.00	118.00	0.310
		Male	62	100.16	20.33	75.00	109.00	0.089
	Teenager	Female	16	144.17	23.81	102.00	168.00	0.150
		Male	27	151.53	20.17	107.00	192.00	0.080
Weight (kg)	Preschool	Female	24	14.52	2.05	9.80	18.60	0.215
		Male	45	13.09	2.17	7.40	18.00	0.175
	Scholar	Female	40	29.58	7.14	17.60	40.00	0.200
		Male	62	23.67	4.85	14.20	44.00	0.115
	Teenager	Female	16	33.84	4.62	22.00	39.50	0.250
		Male	27	34.21	3.19	23.80	45.70	0.140
Height (m)	Preschool	Female	24	0.91	0.07	0.77	1.05	0.289
		Male	45	0.94	0.05	0.80	1.10	0.175
	Scholar	Female	40	1.14	0.07	0.99	1.30	0.200
		Male	62	1.16	0.05	0.94	1.40	0.115
	Teenager	Female	16	1.33	0.04	1.21	1.37	0.250
		Male	27	1.33	0.03	1.16	1.45	0.140
BMI (kg/m ²)	Preschool	Female	24	17.53	2.47	8.32	19.49	0.180
		Male	45	14.81	2.45	18.81	18.87	0.120
	Scholar	Female	40	22.76	5.49	13.45	20.29	0.175
		Male	62	17.59	3.60	12.57	20.92	0.150
	Teenager	Female	16	19.13	2.66	15.02	21.73	0.240
		Male	27	19.33	1.80	14.61	20.99	0.130

3.2. Main Results

3.2.1. Precision

- Intra-rater agreement

The ICC indicated strong precision and consistency in Intra-rater reliability across all three raters when assessing BA using the GP-Canary Atlas, with minor variations between genders. Specifically, Rater 1 showed high consistency, with ICCs of 0.995 (95% CI: 0.990–0.998) for females and 0.996 (95% CI: 0.992–0.998) for males. Similarly, Rater 2 demonstrated strong reliability, with ICCs of 0.990 (95% CI: 0.979–0.995) for females and 0.992 (95% CI: 0.982–0.996) for males. In contrast, Rater 3 reported slightly lower but still strong ICCs, with 0.921 (95% CI: 0.832–0.964) for females and 0.976 (95% CI: 0.947–0.989) for males. More details are in **Table 2**.

Table 2. Intra-rater agreement by time of measurement and gender. This table shows the mean BA values, Intra-class Correlation Coefficient (ICC), and 95% Confidence Intervals (CI) for lower and

upper limits for each Rater (Rater 1, Rater 2, and Rater 3) at two different times of measurement (T1 and T2) for both female and male participants.

Group	Time of measurement	Gender	Mean	ICC	95% CI Lower	95% CI Upper
Rater 1	T1	Female	77.65			
		Male	78.33			
	T2	Female	75.25	0.995	0.990	0.998
		Male	76.21	0.996	0.992	0.998
Rater 2	T1	Female	74.10			
		Male	82.47			
	T2	Female	70.57	0.990	0.979	0.995
		Male	80.94	0.992	0.982	0.996
Rater 3	T1	Female	78.79			
		Male	78.62			
	T2	Female	80.67	0.921	0.832	0.964
		Male	81.83	0.976	0.947	0.989

- Inter-rater agreement
The Inter-rater agreement in determining BA using the GP-Canary Atlas showed notable differences between female and male participants. For females, there was excellent agreement between Rater 1 and Rater 2, with an ICC of 0.982 (95% CI: 0.968, 0.990). However, the agreement was significantly lower between Rater 1 and Rater 3, and between Rater 2 and Rater 3, with ICCs of 0.463 (95% CI: 0.216, 0.654) and 0.509 (95% CI: 0.273, 0.688), respectively. For males, Rater 1 and Rater 2 demonstrated strong consistency with an ICC of 0.944 (95% CI: 0.902, 0.968). In contrast, the agreement between Rater 1 and Rater 3, and between Rater 2 and Rater 3 was lower, with ICCs of 0.408 (95% CI: 0.145, 0.618) and 0.327 (95% CI: 0.052, 0.557), respectively. These findings suggest that while there is high agreement between trained and General Practitioner radiologists, the lower agreement with the student emphasizes the need for standardized training for evaluators using the GP-Canary Atlas. Further details can be found in **Table 3**.

Table 3. Inter-rater agreement of BA assessment using the GP-Canary Atlas by gender. This table presents the mean BA values, Intra-class Correlation Coefficient (ICC), and 95% Confidence Intervals (CI) for lower and upper limits of agreement between different pairs of Raters (Rater 1 vs. Rater 2, Rater 1 vs. Rater 3, and Rater 2 vs. Rater 3) for both female and male participants.

Groups	Gender	Mean	ICC	95% CI Lower	95% CI Upper
Rater 1 - Rater 2	Female	75.73			
		72.34	0.982	0.968	0.990
	Male	78.33			
		81.70	0.944	0.902	0.968
Rater 1 - Rater 3	Female	75.73			
		79.73	0.463	0.216	0.654
	Male	78.33			
		80.22	0.408	0.145	0.618
Rater 2 - Rater 3	Female	72.34			
		79.73	0.509	0.273	0.688
	Male	81.70			
		80.22	0.327	0.052	0.557

The Bland-Altman plots in **Figure 1** illustrate the agreement among the three raters (Rater 1, Rater 2, and Rater 3) for BA assessment using the GP-Canary Atlas. For female participants, Rater 1 and Rater 2 showed high agreement with a narrow range of differences, indicating strong

consistency. In contrast, the agreement between Rater 1 and Rater 3, and between Rater 2 and Rater 3, was moderate, with wider limits of agreement, suggesting more variability due to differences in training. A similar pattern was observed for male participants: Rater 1 and Rater 2 demonstrated strong agreement, while Rater 1 and Rater 3, and especially Rater 2 and Rater 3, exhibited lower agreement with broader ranges of differences, highlighting the challenges of achieving consistent assessments among less experienced raters.

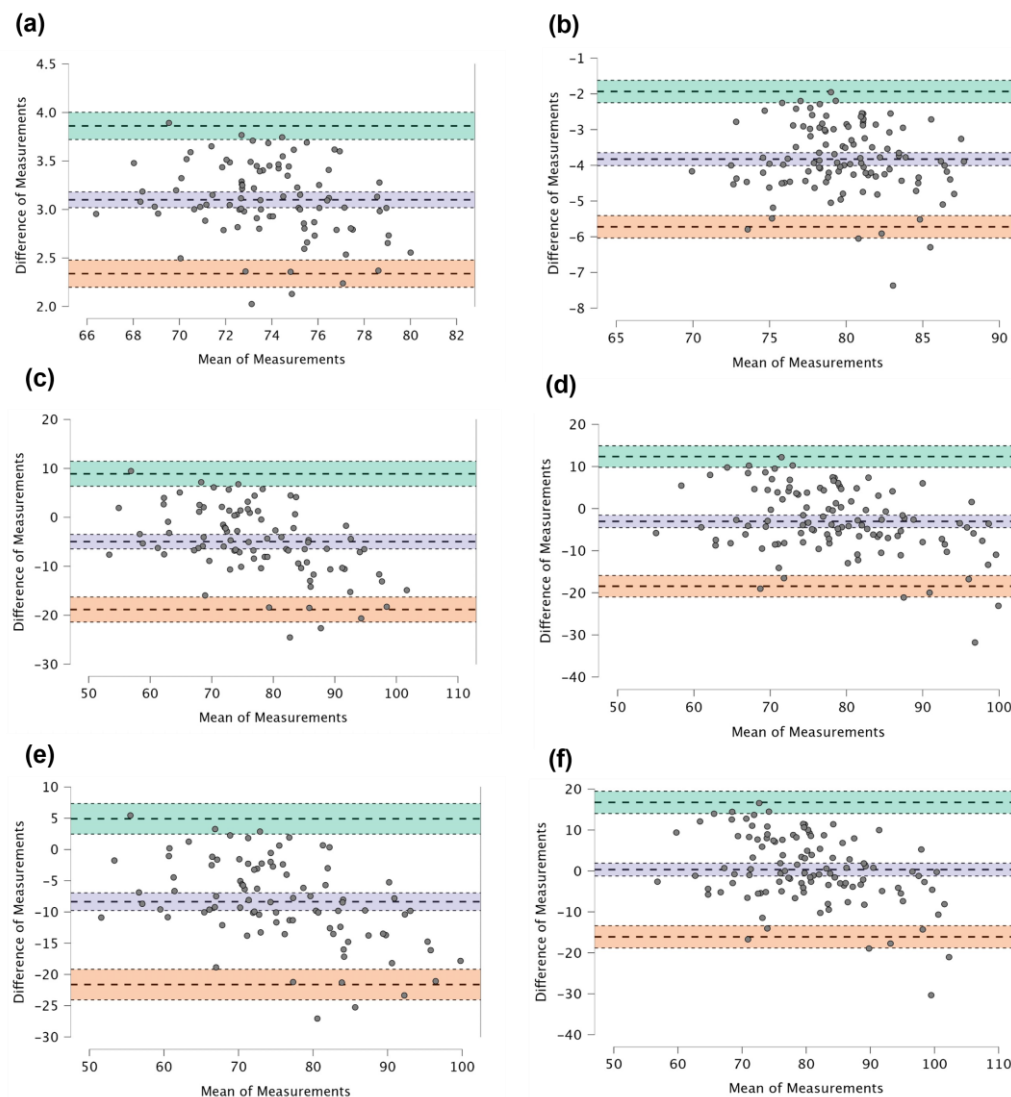


Figure 1. Bland-Altman plots illustrate BA assessments using the GP-Canary Atlas. The plots compare the assessments of Rater 1 with Rater 2 for both females (a) and males (b), Rater 1 with Rater 3 for females (c) and males (d), and Rater 2 with Rater 3 for females (e) and males (f). The dashed lines represent the mean differences, while the shaded areas show the limits of agreement (± 1.96 standard deviations).

3.2.2. Accuracy

The GP-Canary Atlas assessment method demonstrated a lack of accuracy in estimating BA compared to CA in both preschool and school-age groups. Specifically, in the preschool group (ages >1 to 5 years), the method significantly underestimated BA with a mean difference (MD) of 17.036 mos. ($p < 0.001$). This underestimation was even more pronounced in females (MD = 15.081 mos., $p < 0.001$) than in males (MD = 14.898 mos., $p < 0.001$). Similarly, in the school-age group (ages >5 to 12 years), the Atlas continued to underestimate BA, although to a lesser extent, with an MD of 8.165 mos. ($p < 0.001$). Notably, the underestimation was more significant in males (MD = 13.298 mos., $p < 0.001$).

0.001) compared to females (MD = 3.949 mos., $p = 0.239$). In contrast, the GP-Canary Atlas showed the highest accuracy in the teenage group (ages >12 to 18 years), with only a slight overestimation of BA (MD = 3.159 mos., $p = 0.823$). Interestingly, this overestimation was more pronounced in females (MD = 4.497 mos., $p = 0.980$) than in males (MD = 4.85 mos., $p = 0.094$). More details are provided in **Table 4** and visually summarized in **Figure 2**.

Table 4. Accuracy of BA assessments of GP-Canary Atlas. **Abbreviation:** BA = Bone Age, CA = Chronological Age, MD = Mean difference CA - BA, SD = Standard deviation; W = Paired Samples Test Wilcoxon signed-rank statistic. **Statistical significance:** (*) $p < 0.05$, (**) $p < 0.01$, (***) $p < 0.001$.

Stage		Mean	SD	MD	W	Z	p
Preschool (n = 69)	CA	43.485	14.476				
	BA	26.449	15.409	17.036	2297.5	6.517	< 0.001***
Female	CA	39.331	15.182				
	BA	24.250	16.896	15.081	390.0	3.730	< 0.001***
Male	CA	46.496	13.333				
	BA	31.598	24.881	14.898	776.0	4.920	< 0.001***
Scholar (n = 102)	CA	95.684	23.906				
	BA	87.519	35.572	8.165	3306.5	3.346	< 0.001***
Female	CA	92.001	26.086				
	BA	88.052	37.203	3.949	849.0	1.182	0.239
Male	CA	100.168	20.338				
	BA	86.870	33.876	13.298	829.0	3.898	< 0.001***
Teenager (n = 43)	CA	148.883	23.665				
	BA	152.042	29.943	-3.159	339.00	-0.954	0.823
Female	CA	144.170	23.810				
	BA	148.667	24.231	- 4.497	69.0	0.052	0.980
Male	CA	151.53	20.176				
	BA	156.38	18.179	- 4.85	91.50	-1.686	0.094

4. Discussion

4.1. Precision of GP-Canary Atlas

- Intra-rater agreement

Our results show that the GP-Canary Atlas exhibits high Intra-rater precision in BA assessments. The ICC for evaluations by the radiology specialist (Rater 1) was nearly perfect, with ICCs of 0.995 for females and 0.996 for males. The General Practitioner (Rater 2) also demonstrated high precision, with ICCs of 0.990 for females and 0.992 for males. However, the medical student (Rater 3) showed slightly lower precision, with ICCs of 0.921 for females and 0.976 for males.

These findings align with previous research in pediatric populations from Anglo-Saxon countries. Hackman and Black (2012) [37] reported an ICC of 0.969 for Scottish children, and Maggio et al. (2016) [38] found ICCs of 0.970 for males and 0.972 for females in Australia. Similarly, high correlations were reported in Germany and the Netherlands, with Schmidt et al. (2007) [39] finding an ICC of 0.96 for both genders and Van Rijn et al. (2001) [40] reporting ICCs of 0.979 for males and 0.974 for females. Our results also slightly exceed those reported in Southern European countries. Santos et al. (2011) [41] observed excellent Intra-rater agreement in Portugal, with ICCs of 0.99 for both boys and girls, while Pinchi et al. (2014) [42] reported ICCs of 0.907 for males and 0.928 for females in Italy. However, Santoro et al. (2012) [43] found moderate Intra-rater concordance in Southern Italy, with ICCs of 0.88 for males and 0.81 for females. Studies in the United States and Sweden have shown lower reliability, with Calfee et al. (2010) [44] reporting an ICC of 0.890 in a Latin American sample, and Kullman (1995) [45] finding ICCs ranging from 0.64 to 0.74 in Swedish teenagers.

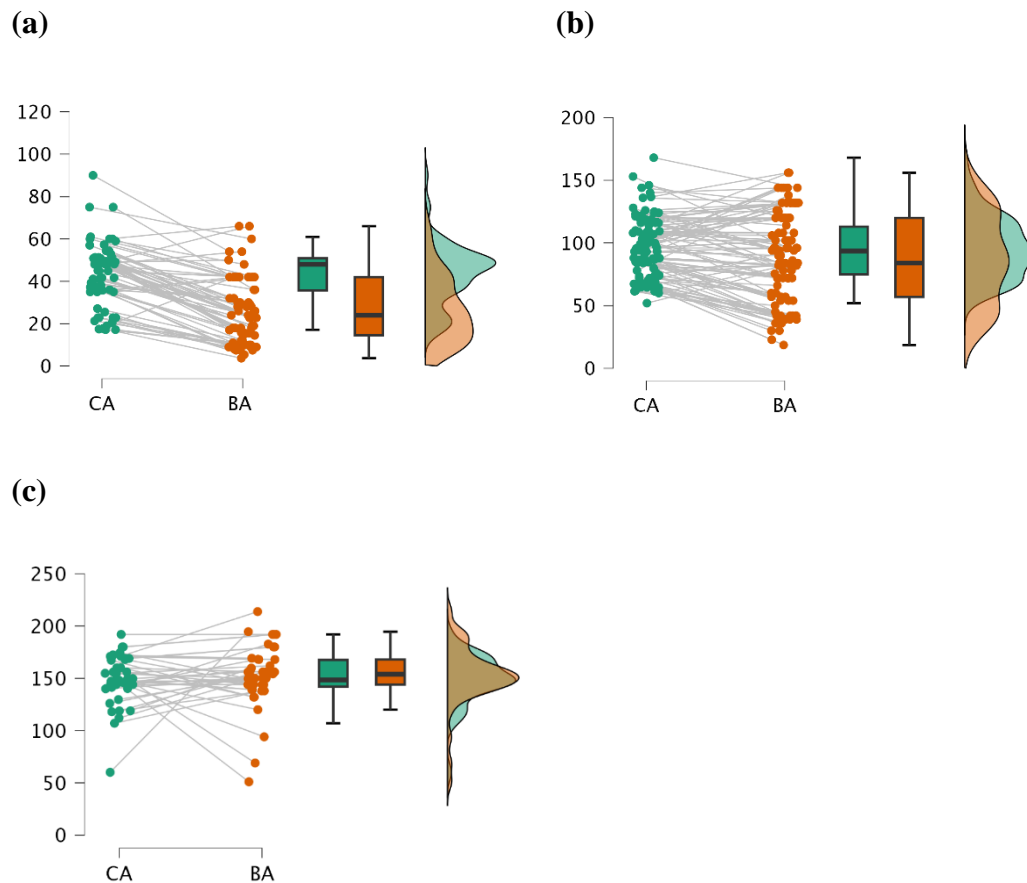


Figure 2. Accuracy of BA determination using the GP-Canary Atlas across different developmental stages. Raincloud plots display BA accuracy in (a) preschool (1 to 5 years), (b) school-age (5 to 12 years), and (c) teenage (12 to 18 years) groups. The method shows significant BA underestimation and variability in the preschool and school-age groups, while accuracy improves in the teenage group with no significant differences between BA and CA.

Similar high Intra-rater agreements have been observed in African studies. Govender and Goodier (2018) [33] in South Africa reported an ICC of 0.99, and Olaotse et al. (2023) [46] in Botswana found ICCs of 0.97 for males and 0.98 for females. Dembetembe et al. (2012) [47] observed moderate precision ($r = 0.76$) using the GP Atlas in Cape Town. Comparable agreements have been reported elsewhere, such as in Saudi Arabia, with Albaker et al. (2021) [48] finding ICCs of 0.995 for males and 0.996 for females, and in Malaysia, with Nang et al. (2023) [49] reporting ICCs of 0.947 for males and 0.933 for females.

The excellent Intra-rater agreement observed with both the GP-Canary Atlas and the GP Atlas can be attributed to the quick and direct visual comparisons they allow, facilitating efficient BA assessments across various pediatric populations. However, the slight variability observed when applying the GP-Canary Atlas may be due to individual cognitive biases and potential misinterpretations of the Atlas [50,51]. Biases such as anchoring, confirmation bias, experience-based bias, overconfidence, the availability heuristic, and the observer expectancy effect can impact the rater's judgment, resulting in inconsistencies and longer times for BA assessments [52,53]. Additionally, errors may occur due to the limited number of maturity indicators available for evaluation, especially when assessing young children. As a child grows, the number of ossification points increases, but when fewer points are present, the potential for assessment errors also becomes higher.

- Inter-rater agreement

Our findings demonstrate a high level of agreement among different raters when using the GP-Canary Atlas for BA determinations. The concordance between the radiology specialist (Rater 1) and the General Practitioner (Rater 2) was remarkably high for both women (ICC = 0.982) and men (ICC = 0.944), indicating that both raters consistently produced similar BA assessments.

On the one hand, these results align with previous studies on the degree of agreement between two expert evaluators when determining the BA of children from Anglo-Saxon countries. For instance, Alshamrani et al. (2019) [54] observed high agreement between two raters in a sample of British children aged 8.80 to 9.59 years. In northern Europe, Zabet et al. (2015)[55] identified an excellent level of Inter-rater concordance among assessors in France (ICC = 0.94, 95% CI: 0.91-0.96, $p < 0.05$). Similarly, Calfee et al. (2010) [44] found very high Inter-rater reliability (ICC = 0.982) in a study involving children from Washington, United States. Additionally, significant agreement among examiners was reported in Oceania, with a Cohen's kappa of 0.887 ($p < 0.001$) when the GP Atlas was used to assess BA in Western Australian children [38]. On the other hand, in Africa, there was a remarkable similarity between the GP-Canary Atlas and GP Atlas Inter-rater agreement. Olaotse et al. (2023) [46] reported that the degree of agreement between two expert raters in assessing BA in the Palapye region of Botswana reached an ICC of 0.94 for girls and 0.93 for boys.

However, the Inter-rater reliability significantly declines when comparing the scores assigned by Rater 1 and Rater 3, as well as Rater 2 and Rater 3. This results in a noticeable reduction in agreement for both girls (ICC = 0.463 and 0.509, respectively) and boys (ICC = 0.408 and 0.327, respectively). The significant decrease in concordance among evaluators with varying levels of experience suggests that less experienced raters might interpret the characteristics of the images differently or make errors when applying the scoring criteria of the GP-Canary Atlas. As reported in other radiological diagnostic tests, this lack of precision may be due to limited familiarity with the specific methodology [56–59]. This highlights the need for more comprehensive training and rigorous standardization of evaluation procedures to ensure that all raters, regardless of experience, apply the criteria consistently and accurately. Such measures are crucial for preventing inconsistent diagnostic decisions in clinical practice.

4.2. Accuracy of GP-Canary Atlas

In preschool-aged children, the GP-Canary Atlas underestimates BA for all genders, showing statistically significant differences (MD = 17.036 mos., $p < 0.001$). This level of underestimation is considerably greater than that reported in several European studies. For example, Martrille et al. (2023) [60] found a significant underestimation using the GP Atlas in Caucasian children from southern France, with a mean difference (MD) of 1.27 mos. (SD = 1.56 mos., $p < 0.05$). Similarly, Santoro et al. (2012) [43] reported underestimations in a southern Italian cohort, with MDs of 1.2 mos. for boys (SD = 15.6 mos., $p = 0.18$) and 4.8 mos. for girls (SD = 12.0 mos., $p < 0.001$). Also, Kullman (1995) [45] also noted a smaller mean underestimation (MD = 4.8 mos.) in Swedish children.

The GP-Canary Atlas may lack accuracy in assessing BA in preschool-aged children due to several reasons. Firstly, during this period, known as "*turgor primus*", children experience rapid and significant growth influenced by thyroid hormones, leading to high variability in ossification points that the Atlas evaluates. This variability makes it challenging to capture changes accurately. Secondly, the Atlas may not be well-designed to reflect the developmental changes influenced by genetic and familial factors [61], rather than following a uniform pattern [62]. Additionally, the presence of children with constitutional delay of growth and puberty (CDGP) could introduce further growth variations that the Atlas may not capture due to insufficient calibration [63]. Finally, errors in reference PA-HW radiographs may also contribute to the lack of accuracy, suggesting that the Atlas might not be reliable for evaluating BA in preschool children.

For school-aged children, the GP-Canary Atlas also underestimates BA but less than that observed in preschoolers, with a mean difference (MD) of 8.165 mos. ($p < 0.001$). The Atlas is more accurate for school-aged girls than boys, largely due to a significant reduction in the underestimation of BA for girls (MD = 3.949 mos., $p = 0.239$). This closer alignment with the Atlas's developmental

stages leads to measurements that are within the normal accuracy range established for this age group, which is up to 12 mos. [2].

During this age, alternating changes in bone maturation, such as increases in length "*proceritas prima*" and "*proceritas secunda*" and weight "*turgor secundus*", occur. In girls, these changes may be accelerated by early puberty [64,65], which is often associated with lifestyle factors, exposure to endocrine disruptors, or genetic determinants [66,67]. Obesity is a significant factor contributing to early puberty, particularly among Hispanic girls [68–70]. The girls in our sample have a high body mass index (BMI = 22.76, SD = 5.49), indicating obesity, which likely leads to an earlier onset of puberty and adolescence. In this phase, bone maturation becomes more regular and standardized compared to preschool children, resulting in earlier developmental stages for girls than boys. This reduces individual differences in bone growth patterns, making them more consistent and predictable, thereby allowing the GP-Canary Atlas to provide more accurate assessments of BA in girls compared to boys.

With respect to teenagers, it has been demonstrated that GP-Canary Atlas increases its accuracy as children mature. However the Atlas slightly overestimates BA with a mean difference (MD) of 3.159 mos. overall ($p = 0.823$), 4.497 mos. for girls ($p = 0.980$), and 4.85 mos. for boys ($p = 0.095$), though these overestimations are not statistically significant. This trend is consistent with studies in geographically similar regions, such as Portugal and Spain, where the GP Atlas also showed progressive overestimation of BA. For instance, Santos et al. [41] reported an increasing MD from 2 to 7 mos. in Portuguese adolescents, while comparisons with the Spanish-adapted Ebri method showed overestimations from 5 to 6.5 mos. (both $p < 0.05$) [71].

Other studies across Europe, including in Lower Saxony, Germany (Schmidt et al., 2007) [39], and the Loire Valley, France (Zabet et al., 2015) [55], demonstrated similar overestimations of BA in teenagers when used GP Atlas, with MDs ranging from 2.29 to 5.8 mos. (all $p < 0.05$). In Anglo-Saxon countries, Hackman and Black [37] found BA overestimations from 1.62 to 11.05 mos. in adolescents aged 13 to 14 years in the northern UK ($p < 0.05$), while Paxton et al. [34] observed an underestimation of 0.81 mos. in early childhood ($p = 0.719$) but a significant overestimation of 3.8 mos. in adolescence ($p = 0.001$) in Caucasian Australian children. Similar trends have been reported in Middle Eastern countries. For example, Soudack et al. (2012) [72] found significant underestimations in Israeli Caucasian children across various age groups: 6-10 years (MD = 2.3 mos., $p < 0.0001$), 10-15 years (MD = 5.4 mos., $p < 0.0001$), and 15-18 years (MD = 3.7 mos., $p < 0.0001$). However, a slight overestimation was noted in those over 18 years (MD = 2.9 mos., $p = 0.0043$). Similarly, Cantekin et al. (2012) [73] reported comparable results in Turkish Caucasian children, with underestimations of 1.32 to 5.76 mos. ($p < 0.05$) in the 7-10 years age group and overestimations up to 9 mos. ($p < 0.05$) in the 10-17 years age group.

Furthermore, GP-Canary Atlas appears more accurate from puberty onwards compared to children from the nearby African continent. However, due to limited data, direct comparisons with North African regions influenced by the Berber ethnic group, the ancestors of the Canary Islands' Guanches, are not possible. In other African countries, Tsehay et al. (2017) [74] reported that the GP Atlas overestimated BA in children aged 10 to 22 years, with a mean difference (MD) of 8.7 mos. for males and 11.8 mos. for females (both $p < 0.05$). Similarly, Olaotse et al. (2023) [46] found overestimations in Botswana ranging from 3 mos. in early bone development to 11.2 mos. for adolescents aged 15 to 18 years ($p < 0.05$). Kowo-Nyakoko et al. (2023) [75] also reported that the GP Atlas overestimates BA by approximately 9.12 mos. in peripubertal children in Zimbabwe.

At this period of development, the GP-Canary Atlas shows increased accuracy in predicting BA for both boys and girls during the final phase of development, known as "*turgor tertius*," which is characterized by rapid growth and hormonal changes driven by sex steroids. This is followed by the "*post-pubertal period*" or "*internubil-puberal of Godin*" marked by the closure of the epiphyseal growth plates, indicating the end of bone growth and the attainment of full skeletal maturity. In this phase of childhood, the differences in development between boys and girls decrease, leading to more synchronized and predictable maturation patterns. The GP-Canary Atlas captures this synchronization, as evidenced by the mean differences in BA for females (MD = -4.497 mos.) and

males (MD = -4.85 mos.). These consistent changes allow the Atlas to predict BA accurately within the normal range of up to 24 mos. [2], making it a reliable tool for assessing BA in adolescents across genders.

5. Conclusion

The study confirms the GP-Canary Atlas as a valid diagnostic tool for assessing BA in children and adolescents, demonstrating high Intra-rater and good Inter-rater reliability. However, its accuracy varies by age group, with significant underestimations in preschool and school-aged children and slight overestimations in teenagers, suggesting a need for age-specific adjustments.

Author Contributions: Conceptualization, I.M.M.P., and S.E.M.P.; methodology, I.M.M.P., and S.E.M.P.; resources, I.M.M.P., S.E.M.P.; J.M.V.G.; Formal analysis, I.M.M.P. and S.E.M.P.; writing—original draft, I.M.M.P. and S.E.M.P.; writing—review & editing, A.G.H., M.H.P. and R.M.S.; supervision, A.G.H., M.H.P., F.R.H.; Project administration, I.M.M.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board (or Ethics Committee) of Complejo Hospitalario Universitario de Canarias (CHUC_2023_86—13 July 2023).

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Tanner, J.M. *Growth at Adolescence*; Blackwell Scientific Publications: Oxford, UK, **1962**.
2. Toledo Trujillo, F.M.; Others. *Atlas Radiológico de Referencia de la Edad Ósea en la Población Canaria*; Fundación Canaria de Salud y Sanidad de Tenerife: Santa Cruz de Tenerife, España, **2009**.
3. Nandiraju, D.; Ahmed, I. Human skeletal physiology and factors affecting its modeling and remodeling. *Fertil. Steril.* **2019**, 112(5), 775–781. doi:10.1016/j.fertnstert.2019.10.005.
4. Macias, H.; Hinck, L. Mammary Gland Development. *Wiley Interdiscip. Rev. Dev. Biol.* **2012**, 1, 533–557. doi:10.1002/wdev.35.
5. Susman, E.J.; Houts, R.M.; Steinberg, L.; Belsky, J.; Cauffman, E.; Dehart, G.; Friedman, S.L.; Roisman, G.I.; Halpern-Felsher, B.L.; Eunice Kennedy Shriver NICHD Early Child Care Research Network. Longitudinal Development of Secondary Sexual Characteristics in Girls and Boys Between Ages 9½ and 15½ Years. *Arch. Pediatr. Adolesc. Med.* **2010**, 164, 166–173. doi:10.1001/archpediatrics.2009.261.
6. Bangalore Krishna, K.; Witchel, S.F. Normal Puberty. *Endocrinol. Metab. Clin. North Am.* **2024**, 53, 183–194. doi:10.1016/j.ecl.2024.01.001.
7. Niwczyk, O.; Grymowicz, M.; Szczęsnowicz, A.; Hajbos, M.; Kostrzak, A.; Budzik, M.; Maciejewska-Jeske, M.; Bala, G.; Smolarczyk, R.; Męczekalski, B. Bones and Hormones: Interaction Between Hormones of the Hypothalamus, Pituitary, Adipose Tissue and Bone. *Int. J. Mol. Sci.* **2023**, 24, 6840. doi:10.3390/ijms24076840.
8. Ulijaszek, S.J. The International Growth Standard for Children and Adolescents Project: Environmental Influences on Preadolescent and Adolescent Growth in Weight and Height. *Food Nutr. Bull.* **2006**, 27, S279–S294. doi:10.1177/15648265060274S510.
9. Johnson, A.B. Genetic Determinants of Maturation Under Favorable Environmental Conditions. *Genet. Dev.* **2017**, 5, 112–125.
10. Cavallo, F.; Mohn, A.; Chiarelli, F.; Giannini, C. Evaluation of Bone Age in Children: A Mini-Review. *Front. Pediatr.* **2021**, 9, 580314. doi:10.3389/fped.2021.580314.
11. Navarro, M.M.; Tejedor, B.M.; Siguero, J.P.L. El uso de la edad ósea en la práctica clínica. *Anales de Pediatría Continuada* **2014**, 12(6), 275–283.
12. Roberts, C.D. Influence of Environment on Genetic Control of Maturation: A Longitudinal Study. *Environ. Genet.* **2019**, 28, 78–91.
13. Díaz Gómez, M.N. *Crecimiento y desarrollo físico del niño*; Tenerife, **1992**; (18).
14. Liu, X.; Zhang, J.; Zheng, Z. Clinical Methods for Bone Age Assessment in Pediatrics. *J. Pediatr. Endocrinol. Metab.* **2018**, 31, 487–495.
15. Smith, R.; Johnson, M.; Williams, L. Hormonal Profiling in Pediatric Endocrinology. *Endocr. Rev.* **2020**, 42, 301–318.

16. Gilsanz, V.; Ratib, O. *Hand Bone Age: A Digital Atlas of Skeletal Maturity*; Springer: **2011**.
17. Dwyer, A.A.; Hayes, F.J. Evaluation of Endocrine Disorders of the Hypothalamic-Pituitary-Gonadal (HPG) Axis. In: Llahana, S., Follin, C., Yedinak, C., Grossman, A. (Eds.) *Advanced Practice in Endocrinology Nursing*; Springer: Cham, **2019**. doi:10.1007/978-3-319-99817-6_45.
18. Jones, A.; Brown, B. *Pediatric Bone Age Assessment: A Practical Guide*; Springer: New York, NY, USA, **2019**.
19. Greulich, W.W.; Pyle, S.I. *Radiographic Atlas of Skeletal Development of the Hand and Wrist*, 2nd ed.; Stanford University Press: Stanford, CA, USA, **1959**.
20. Prokop-Piotrkowska, M.; Marszałek-Dziuba, K.; Moszczyńska, E.; Szalecki, M.; Jurkiewicz, E. Traditional and New Methods of Bone Age Assessment—An Overview. *J. Clin. Res. Pediatr. Endocrinol.* **2021**, *13*, 251–262. doi:10.4274/jcrpe.galenos.2020.2020.0091.
21. Satoh, M.; Hasegawa, Y. Factors affecting prepubertal and pubertal bone age progression. *Front. Endocrinol. (Lausanne)* **2022**, *13*, 967711. doi:10.3389/fendo.2022.967711.
22. Grgic O, Shevroja E, Dharmo B, Uitterlinden AG, Wolvius EB, Rivadeneira F, Medina-Gomez C. Skeletal maturation in relation to ethnic background in children of school age: The Generation R Study. *Bone*. **2020** Mar;132:115180. doi: 10.1016/j.bone.2019.115180.
23. Khadilkar, V.; Oza, C.; Khadilkar, A. Relationship between Height Age, Bone Age and Chronological Age in Normal Children in the Context of Nutritional and Pubertal Status. *J. Pediatr. Endocrinol. Metab.* **2022**, *35*, 767–775. doi:10.1515/jpem-2021-0698.
24. Toledo Trujillo, F.M. *Maduración ósea en una muestra de población urbana de las islas Canarias*; Doctoral Thesis, Universidad La Laguna: San Cristóbal de La Laguna, España, **1978**.
25. Zhang, A.; Sayre, J.W.; Vachon, L.; Liu, B.J.; Huang, H.K. Racial Differences in Growth Patterns of Children Assessed on the Basis of Bone Age. *Radiology* **2009**, *250*, 228–235. doi:10.1148/radiol.2493080468.
26. Martín Pérez, S.E.; Martín Pérez, I.M.; Vega González, J.M.; Molina Suárez, R.; León Hernández, C.; Rodríguez Hernández, F.; Herrera Perez, M. Precision and Accuracy of Radiological Bone Age Assessment in Children Among Different Ethnic Groups: A Systematic Review. *Diagnostics (Basel)* **2023**, *13*, 3124. doi:10.3390/diagnostics13193124.
27. Ontell, F.K.; Ivanovic, M.; Ablin, D.S.; Barlow, T.W. Bone Age in Children of Diverse Ethnicity. *AJR Am. J. Roentgenol.* **1996**, *167*, 1395–1398. doi:10.2214/ajr.167.6.8956565.
28. Fregel, R.; Ordóñez, A.C.; Serrano, J.G. The Demography of the Canary Islands from a Genetic Perspective. *Hum. Mol. Genet.* **2021**, *30*, R1. doi:10.1093/hmg/ddaa262.
29. Alshamrani, K.; Messina, F.; Offiah, A.C. Is the Greulich and Pyle Atlas Applicable to All Ethnicities? A Systematic Review and Meta-Analysis. *Eur. Radiol.* **2019**, *29*, 2910–2923. doi:10.1007/s00330-018-5792-5.
30. Bossuyt, P.M.; Reitsma, J.B.; Bruns, D.E.; Gatsonis, C.A.; Glasziou, P.P.; Irwig, L.; Lijmer, J.G.; Moher, D.; Rennie, D.; de Vet, H.C.; Kressel, H.Y.; Rifai, N.; Golub, R.M.; Altman, D.G.; Hooft, L.; Korevaar, D.A.; Cohen, J.F.; STARD Group. STARD 2015: An Updated List of Essential Items for Reporting Diagnostic Accuracy Studies. *BMJ* **2015**, *351*, h5527. doi: 10.1136/bmj.h5527.
31. Bhat, A.K.; Kumar, B.; Acharya, A. Radiographic Imaging of the Wrist. *Indian J. Plast. Surg.* **2011**, *44*, 186–196. doi:10.4103/0970-0358.85339.
32. Hardy, D.C.; Totty, W.G.; Reinus, W.R.; Gilula, L.A. Posteroanterior Wrist Radiography: Importance of Arm Positioning. *J. Hand Surg. Am.* **1987**, *12*, 504–508. doi:10.1016/s0363-5023(87)80196-x.
33. Govender, D.; Goodier, M. Bone of Contention: The Applicability of the Greulich–Pyle Method for Skeletal Age Assessment in South Africa. *South Afr. J. Radiol.* **2018**, *22*, 6. doi:10.4102/sajr.v22i1.1348.
34. Tiwari, P.K.; Gupta, M.; Verma, A.; Pandey, S.; Nayak, A. Applicability of the Greulich-Pyle Method in Assessing the Skeletal Maturity of Children in the Eastern Uttar Pradesh (UP) Region: A Pilot Study. *Cureus* **2020**, *12*, e10880. doi:10.7759/cureus.10880.
35. KKim, J.R.; Lee, Y.S.; Yu, J. Assessment of Bone Age in Prepubertal Healthy Korean Children: Comparison Among the Korean Standard Bone Age Chart, Greulich-Pyle Method, and Tanner-Whitehouse Method. *Korean J. Radiol.* **2015**, *16*, 201–205. doi:10.3348/kjr.2015.16.1.201.
36. Fraga Bermúdez, J.M.; Fernández Lorenzo, J.R. La Pediatría, el Niño y el Pediatra: Una Aproximación General. In: *Tratado de Pediatría*, 1st ed.; Moro Serrano, M., Málaga Guerrero, S., Madero López, L., Eds.; Editorial Médica Panamericana: Madrid, Spain, **2014**; Volume 1, pp. 1–18.
37. Hackman, L.; Black, S. The Reliability of the Greulich and Pyle Atlas When Applied to a Modern Scottish Population. *J. Forensic Sci.* **2012**, *58*, 114–119. doi:10.1111/j.1556-4029.2012.02294.x.
38. Maggio, A.; Flavel, A.; Hart, R.; Franklin, D. Assessment of the Accuracy of the Greulich and Pyle Hand-Wrist Atlas for Age Estimation in a Contemporary Australian Population. *Aust. J. Forensic Sci.* **2016**, *50*, 385–395. doi:10.1080/00450618.2016.1251970.
39. Schmidt, S.; Koch, B.; Schulz, R.; Reisinger, W.; Schmeling, A. Comparative Analysis of the Applicability of the Skeletal Age Determination Methods of Greulich–Pyle and Thiemann–Nitz for Forensic Age Estimation in Living Subjects. *Int. J. Leg. Med.* **2007**, *121*, 293–296. doi:10.1007/s00414-007-0165-7.
40. Van Rijn, R.R.; Lequin, M.H.; Robben, S.G.F.; Hop, W.C.J.; van Kuijk, C. Is the Greulich and Pyle Atlas Still Valid for Dutch Caucasian Children Today? *Pediatr. Radiol.* **2001**, *31*, 748–752. doi:10.1007/s002470100531.

41. Santos, C.; Ferreira, M.; Alves, F.C.; Cunha, E. Comparative Study of Greulich and Pyle Atlas and Matusos 4.0 Program for Age Estimation in a Portuguese Sample. *Forensic Sci. Int.* **2011**, *212*, 276.e1–276.e7. doi:10.1016/j.forsciint.2011.05.032.
42. Pinchi, V.; De Luca, F.; Ricciardi, F.; Focardi, M.; Piredda, V.; Mazzeo, E.; Norelli, G.-A. Skeletal age estimation for forensic purposes: A comparison of GP, TW2 and TW3 methods on an Italian sample. *Forensic Sci Int.* **2014**, *238*, 83–90. doi: 10.1016/j.forsciint.2014.02.030.
43. Santoro, V.; Roca, R.; De Donno, A.; Fiandaca, C.; Pinto, G.; Tafuri, S.; Introna, F. Applicability of Greulich and Pyle and Demirjian aging methods to a sample of Italian population. *Forensic Sci Int.* **2012**, *221*, 153.e1–153.e5. doi: 10.1016/j.forsciint.2012.04.014.
44. Calfee, R.P.; Sutter, M.; Steffen, J.A.; Goldfarb, C.A. Skeletal and chronological ages in American adolescents: Current findings in skeletal maturation. *J Child Orthop.* **2010**, *4*, 467–470. doi: 10.1007/s11832-010-0289-z.
45. Kullman, L. Accuracy of two dental and one skeletal age estimation method in Swedish adolescents. *Forensic Sci Int.* **1995**, *75*, 225–236. doi: 10.1016/0379-0738(95)01792-5.
46. Olaotse B., Norma P.G., Kaone P.-M., Morongwa M., Janes M., Kabo K., Shathani M., Thato P. Evaluation of the suitability of the Greulich and Pyle atlas in estimating age for the Botswana population using hand and wrist radiographs of young Botswana population. *Forensic Sci. Int. Rep.* **2023**, *7*. doi: 10.1016/j.fsr.2023.100312.
47. Dembetembe K.A., Morris A.G. Is Greulich–Pyle age estimation applicable for determining maturation in male Africans? *South Afr. J. Sci.* **2012**, *108*. doi: 10.4102/sajs.v108i9/10.1036.
48. Albaker A.B., Aldhilan A.S., Alrabai H.M., AlHumaid S., AlMogbil I.H., Alzaidy N.F.A., Alsaadoon S.A.H., Alobaid O.A., Alshammary F.H. Determination of Bone Age and its Correlation to the Chronological Age Based on the Greulich and Pyle Method in Saudi Arabia. *J. Pharm. Res. Int.* **2021**, *1186–1195*. doi: 10.9734/jpri/2021/v33i60B34731.
49. Nang K.M., Ismail A.J., Tangaperumal A., Wynn A.A., Thein T.T., Hayati F., Teh Y.G. Forensic age estimation in living children: How accurate is the Greulich-Pyle method in Sabah, East Malaysia? *Front. Pediatr.* **2023**, *11*, 1137960. doi: 10.3389/fped.2023.1137960.
50. Yoon, S.Y.; Lee, K.S.; Bezuidenhout, A.F.; Kruskal, J.B. Spectrum of Cognitive Biases in Diagnostic Radiology. *Radiographics* **2024**, *44*, e230059. doi:10.1148/rq.230059.
51. Chen, J.; Gandomkar, Z.; Reed, W.M. Investigating the Impact of Cognitive Biases in Radiologists' Image Interpretation: A Scoping Review. *Eur. J. Radiol.* **2023**, *166*, 111013. doi:10.1016/j.ejrad.2023.111013.
52. Busby, L.P.; Courtier, J.L.; Glastonbury, C.M. Bias in Radiology: The How and Why of Misses and Misinterpretations. *Radiographics* **2018**, *38*, 236–247. doi:10.1148/rq.2018170107.
53. Berst, M.J.; Dolan, L.; Bogdanowicz, M.M.; Stevens, M.A.; Chow, S.; Brandser, E.A. Effect of Knowledge of Chronologic Age on the Variability of Pediatric Bone Age Determined Using the Greulich and Pyle Standards. *AJR Am. J. Roentgenol.* **2001**, *176*, 507–510. doi:10.2214/ajr.176.2.1760507.
54. Alshamrani, K.; Offiah, A.C. Applicability of Two Commonly Used Bone Age Assessment Methods to Twenty-First Century UK Children. *Eur. Radiol.* **2019**, *30*, 504–513. doi:10.1007/s00330-019-06300-x.
55. Zabet, D.; Rérolle, C.; Puchaux, J.; Telmon, N.; Saint-Martin, P. Can the Greulich and Pyle Method Be Used on French Contemporary Individuals? *Int. J. Leg. Med.* **2014**, *129*, 171–177. doi:10.1007/s00414-014-1028-7.
56. Dawes, T.J.; Vowler, S.L.; Allen, C.M.; Dixon, A.K. Training Improves Medical Student Performance in Image Interpretation. *Br. J. Radiol.* **2004**, *77*, 775–776. doi:10.1259/bjr/66388556.
57. Vincent, C.A.; Driscoll, P.A.; Audley, R.J.; Grant, D.S. Accuracy of Detection of Radiographic Abnormalities by Junior Doctors. *Arch. Emerg. Med.* **1988**, *5*, 101–109. doi:10.1136/emj.5.2.101.
58. Christiansen JM, Gerke O, Karstoft J, Andersen PE. Poor interpretation of chest X-rays by junior doctors. *Dan Med J.* **2014** Jul, *61*(7), A4875.
59. Cheung, T.; Harianto, H.; Spanger, M.; Young, A.; Wadhwa, V. Low Accuracy and Confidence in Chest Radiograph Interpretation Amongst Junior Doctors and Medical Students. *Intern. Med. J.* **2018**, *48*, 864–868. doi:10.1111/imj.13946.
60. Martrille, L.; Papadodima, S.; Venegoni, C.; Molinari, N.; Gibelli, D.; Baccino, E.; Cattaneo, C. Age Estimation in 0–8-Year-Old Children in France: Comparison of One Skeletal and Five Dental Methods. *Diagnostics* **2023**, *13*, 1042.
61. Al-Khater, K.M.; Hegazi, T.M.; Al-Thani, H.F.; Al-Muhanna, H.T.; Al-Hamad, B.W.; Alhuraysi, S.M.; Alsfyani, W.A.; Alessa, F.W.; Al-Qwairi, A.O.; Al-Qwairi, A.O.; Bayer, S.B.; Siddiqui, F.B. Time of appearance of ossification centers in carpal bones: A radiological retrospective study on Saudi children. *Saudi Med. J.* **2020**, *41*, 938–946. <https://doi.org/10.15537/smj.2020.9.25348>.
62. Reynolds E. Degree of kinship and pattern of ossification. A longitudinal X-ray study of the appearance pattern of ossification centers in children of different kinship groups. *AJBA.* **1943**, *1*(4), 405–416. <https://doi.org/10.1002/ajpa.1330010411>

63. Gaudino, R.; De Filippo, G.; Bozzola, E.; Gasparri, M.; Bozzola, M.; Villani, A.; Radetti, G. Current clinical management of constitutional delay of growth and puberty. *Ital. J. Pediatr.* **2022**, *48*(1), 45. doi: 10.1186/s13052-022-01242-5.
64. Rosenfield, R.L.; Lipton, R.B.; Drum, M.L. Thelarche, Pubarche, and Menarche Attainment in Children with Normal and Elevated Body Mass Index. *Pediatrics* **2009**, *123*, 84–88. doi:10.1542/peds.2008-0146. Erratum in: *Pediatrics* **2009**, *123*, 1255.
65. De Bont, J.; Díaz, Y.; Casas, M.; García-Gil, M.; Vrijheid, M.; Duarte-Salles, T. Time Trends and Sociodemographic Factors Associated with Overweight and Obesity in Children and Adolescents in Spain. *JAMA Netw. Open* **2020**, *3*, e201171. doi:10.1001/jamanetworkopen.2020.1171.
66. Li, W.; Liu, Q.; Deng, X.; Chen, Y.; Liu, S.; Story, M. Association between Obesity and Puberty Timing: A Systematic Review and Meta-Analysis. *Int. J. Environ. Res. Public Health* **2017**, *14*, 1266. doi:10.3390/ijerph14101266.
67. Huang, A.; Reinehr, T.; Roth, C.L. Connections Between Obesity and Puberty: Invited by Manuel Tena-Sempere, Cordoba. *Curr. Opin. Endocr. Metab. Res.* **2020**, *14*, 160–168. doi:10.1016/j.coemr.2020.08.004.
68. Gavela-Pérez, T.; Garcés, C.; Navarro-Sánchez, P.; López Villanueva, L.; Soriano-Guillén, L. Earlier Menarcheal Age in Spanish Girls Is Related with an Increase in Body Mass Index Between Pre-Pubertal School Age and Adolescence. *Pediatr. Obes.* **2015**, *10*, 410–415. doi:10.1111/ijpo.277.
69. Pérez-Rodrigo, C.; Aranceta Bartrina, J.; Serra Majem, L.; Moreno, B.; Delgado Rubio, A. Epidemiology of Obesity in Spain. Dietary Guidelines and Strategies for Prevention. *Int. J. Vitam. Nutr. Res.* **2006**, *76*, 163–171. doi:10.1024/0300-9831.76.4.163.
70. Shi, L.; Jiang, Z.; Zhang, L. Childhood Obesity and Central Precocious Puberty. *Front. Endocrinol. (Lausanne)* **2022**, *13*, 1056871. doi:10.3389/fendo.2022.1056871.
71. Ebri Torné, B. Comparative Study Between Bone Ages: Carpal, Metacarpophalangeal, Carpometacarpophalangeal Ebri, Greulich and Pyle, and Tanner Whitehouse2. *Med. Res. Arch.* **2021**, *9*, e2625. doi:10.18103/mra.v9i12.2625.
72. Soudack, M.; Ben-Shlush, A.; Jacobson, J.; Raviv-Zilka, L.; Eshed, I.; Hamiel, O. Bone Age in the 21st Century: Is Greulich and Pyle's Atlas Accurate for Israeli Children? *Pediatr. Radiol.* **2012**, *42*, 343–348.
73. Cantekin, K.; Celikoglu, M.; Miloglu, O.; Dane, A.; Erdem, A. Bone Age Assessment: The Applicability of the Greulich-Pyle Method in Eastern Turkish Children. *J. Forensic Sci.* **2011**, *57*, 679–682. doi:10.1111/j.1556-4029.2011.02035.x.
74. Tsehay, B.; Afework, M.; Mesifin, M. Assessment of Reliability of Greulich and Pyle (GP) Method for Determination of Age of Children at Debre Markos Referral Hospital, East Gojjam Zone. *Ethiop. J. Health Sci.* **2017**, *27*, 631–640.
75. Kowo-Nyakoko, F.; Gregson, C.L.; Madanhire, T.; Stranix-Chibanda, L.; Rukuni, R.; Offiah, A.C.; Micklesfield, L.K.; Cooper, C.; Ferrand, R.A.; Rehman, A.M.; et al. Evaluation of Two Methods of Bone Age Assessment in Peripubertal Children in Zimbabwe. *Bone* **2023**, *170*, 116725. doi: 10.4314/ejhs.v27i6.8.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.