DOI: 10.4274/jcrpe.galenos.2020.2020.0091

Review

Traditional and New Methods of Bone Age Assessment – An Overview

Prokop-Piotrkowska M et al. Methods of Bone Age Assessment

Monika Prokop-Piotrkowska[1], Kamila Marszałek-Dziuba[1], Elżbieta Moszczyńska[1], Mieczysław Szalecki[1,2], Elżbieta Jurkiewicz[3]
[1]Department of Endocrinology and Diabetology, Children's Memorial Health Institute,
[2]Collegium Medicum, Jan Kochanowski University, Kielce, Poland
[3]Department of Diagnostic Imaging, Children's Memorial Health Institute,

Abstract
Bone age is one of biological indicators of maturity used in clinical practice and it is a very important parameter of a child's assessment, especially in paediatric endocrinology. The most widely used method of bone age assessment is by performing a hand and wrist radiograph and its analysis with Greulich-Pyle or Tanner-Whitehouse atlases, although it has been about 60 years since they were published. Due to the progress in the area of Computer-Aided Diagnosis and application of artificial intelligence in medicine, lately, numerous programs for automatic bone age assessment have been created. Most of them have been verified in clinical studies in comparison to traditional methods showing good precision while eliminating inter- and intra-rater variability and reducing significantly the time of assessment. Also, there are available methods of assessment of bone age without X-ray exposure, like via ultrasound devices or MRI.
Keywords: Maturation, children, radiographs, deep learning, neural networks

Corresponding author: Monika Prokop-Piotrkowska MD,
+48 608 523 869
m.prokop-piotrkowska@ipczd.pl
08.05.2020
15.10.2020
ORCID-ID 0000-0003-3323-6784

Introduction
Maturation indicators
Processes of growth and maturation in children are usually correlated but they cannot be treated as one as they may not be linear and they may proceed at a different pace. Due to numerous disturbances like growth hormone deficiency, thyroid hormones deficiency or delayed puberty, but sometimes also in healthy children, the chronological age doesn't meet the biological age. The reason is that they are regulated by various factors, which are, besides genes and nutrition, many hormones, including growth hormone (GH), IGF-1, sex hormones and adrenal steroids like cortisol, DHEAS, testosterone (1,2). In paediatric endocrinology, it is especially important to assess the child's growth and puberty in relation to biological age, not the metrical one. This is the reasons why clinicians have been looking for a good marker of maturation rate in children for decades (3) .
Age at menarche is a solid biological indicator of maturity, but it is a one-off event and relates to only half of the population (3) . Dentists, mainly orthodontists, use dental age in Demirjian or Willems scale in daily practice, however, it has not been established as a reliable tool for other clinicians (3–5) . On the other hand, sexual characteristics, like an assessment in the Tanner scale, are useful only in the adolescent period and are very subjective.
The only biological indicator of maturity which is available from birth to adulthood is bone age (BA) (3) .
Bone age
In paediatric endocrinology, BA is an important tool used in the clinical assessment of patients, mainly those suffering from growth and puberty disorders. Many parameters correlate better with BA than with chronological age (CA) (e.g. height velocity, menarche, muscle mass and bone mineral mass) (6) . Delayed BA is typical for growth hormone deficiency, constitutional delay of growth, hypothyroidism, malnutrition and chronic illness (6,7) . On the other hand, BA is advanced in, among others, precocious puberty and congenital adrenal hyperplasia, when there is a prolonged elevation of sex steroid levels (6–8) . BA may be also marginally advanced in cases of overweight children, children with tall stature or premature adrenarche (1,6,8) . In genetic overgrowth syndromes (for example Sotos syndrome, Beckwith-Wiedemann syndrome and Marshall-Smith syndrome) BA is usually significantly advanced (6). In all cases it is important to remember that advancement or delay of BA in relation to chronological age is a slow process, thus BA may not be altered in the case of examinations performed shortly after the first manifestations of a disorder (7) .
What is more, BA is used in forensic and legal medicine to estimate chronological age, for example in asylum seekers or unaccompanied minors without documents. In such cases an adequate assessment of age using precise methods is crucial, the consequences of incorrect assessment of a child as an adult may result in more restricted access to education, medical care or other forms of support provided for children (9).

uncorrected proof

This article presents the issue of different methods of BA assessment from the perspective of a paediatrician or paediatric endocrinologist (Tab. 1.).

Traditional Methods

Although there have been attempts to assess BA by examinations of different bones like clavicle or iliac bone (Risser sign) (10–15), in paediatrics and paediatric endocrinology, the established way to obtain BA is by performing a radiograph of the hand and wrist of the non-dominant hand. Assessment of development of the bones can be performed by traditional, manual way or using one of the automated methods. The manual method means a comparison of obtained radiograph with radiographs in atlases. The manual methods can be divided into two groups depending on the type of atlas – holistic or analytic.

The first atlases were published shortly after the discovery of X-rays in 1895. In 1898 John Poland published the first one: "skiagraphic atlas showing the development of bones of the wrist and hand" (16). In his atlas, he depicted skiagraphs (positive reprints) of hand radiographs of 19 British children, aged between 1 and 17 years, with an attached description of each radiograph (16). However, the two most important publications in this field were presented in 1959 by Greulich and Pyle (17) and in 1962 by Tanner, Whitehouse and Healy (18).

Greulich-Pyle atlas

'The Radiographic Atlas of Skeletal Development of the Hand and Wrist' by Greulich and Pyle (GP) has been widely recognised and in use until nowadays. This atlas was created based on radiographs of hands of paediatric patients referred to endocrinologists dr William Walter Greulich and dr Sarah Idell Pyle by paediatricians in years 1931-1942. These patients were Caucasian children from upper middle class living in Cleveland, Ohio, United States (19–21). This atlas consists of separate reference images for boys and girls aged 0-18 (boys) or 0-19 years (girls) in various intervals (3 months -1 year). Images are accompanied by an explanation of the gradual age-related changes in the bones at a given age and separate BA calculated for each bone. Due to the natural variability of BA of different bones in one individual, in some bones, it is often more or less advanced than the standard it is intended to represent. For example, a radiograph representing the age of 3 years 6 month (42 months) includes a 36-month first metacarpal and a 54-month lunate (17). BA is calculated by comparing the non-dominant wrist radiographs of the subject with the nearest matching reference radiographs provided in the atlas, thus this method is called a holistic one.

GP is the most popular method among clinicians and radiologists as the assessment by GP is relatively quick and easy to learn. Although widely recognised, this method has significant drawbacks. Bone age assessment (BAA) using GP is highly inter- and intra-observer variable and GP may not be an appropriate, universal tool nowadays for use in various populations.

BAA by GP is very subjective and the standard error on a single determination in inter-observer studies ranges from 0.45 to 0.83 years (22–26). There is no standardization in how the bones are weighted. Depending on a rater, in clinical practice one may assign different weight to different bones, some raters may ignore the carpals and others may assign even half weight to the carpals during the assessment. Raters using the carpals reduce their importance at higher maturity but again not in a standardized manner (25).

It has been reported that currently boys and girls develop secondary sex characteristics earlier than decades ago in United States (27,28), thus GP atlas nowadays may not be as precise as when it was created.

What is more, it has been proven that correlation of BA with CA and consequently the applicability of GP depends on ethnic origin (29,30). According to a recent meta-analysis it has been proven that in African females, in comparison to GP standards, BA is significantly advanced. On the other hand, in Asian males, it is significantly delayed between 6 and 9 years old inclusive and significantly advanced at 17 years (29). This should be taken into consideration while assessing BA in these populations using GP atlas.

There is an online version of GP uploaded by Brazilian Instituto Mineiro de Endocrinologia [29].

Tanner-Whitehouse atlas

The second most popular tool for BA assessment is Tanner-Whitehouse atlas (TW). Its first version has been created in 1962 based on 2600 radiographs collected in the1950s and 1960s of British children coming from average socio-economic class (18). It was later updated in 1983 – Tanner-Whitehouse 2 (TW2) and in 2001 the latest updated version was published – Tanner-Whitehouse 3 (TW3), which takes into consideration the secular trends that influence the relationship between the total bone maturity score and BA (31). In several countries standardized TW methods have been created which change the relationship between the total maturity score and BA to make it suitable for different ethnic groups (32–34).

TW2 is called an analytic or scoring method and it is based on the maturity levels of 20 regions of interest (ROI) in different bones of the hand and wrist. The level of development of each ROI is labelled as a given stage which is calculated to a numerical score. A total maturity score is calculated by adding the scores of the ROIs and it is matched with the age of boys and girls separately.

TW method is considered to be more objective than the holistic method (GP) and to have higher reproducibility than the GP method. It was reported that in the case of GP the intra-observer variation was greater than in TW (95% confidence interval, -2.46 to 2.18 vs -1.48 to 1.43 respectively) (22). On the other hand, assessment using the TW method is more time-consuming. In a study performed by King et al. average time required for TW assessment was calculated as 7.9 min. vs. 1.4 min. in the case of GP assessment (35). In this study the intra-observer variation between GP and TW assessment was concluded as not statistically significant (the average spread of results was 0.74 years for TW and 0.96 years for the GP), however, the number of examinations analysed was much smaller than in the study mentioned before (35).

The comparison of GP and TW methods is presented in Table 1 (Tab. 2)

Other atlases

uncorrected proof

The FELS method was developed in 1988 using 13,823 serial radiographs of the left hand-wrist of boys and girls in the Fels Longitudinal Study performed by WM. Cameron Chumlea, Alex F. Roche and David Thissen from two universities in Kansas and Ohio, US (36). It is based upon maturity indicators that are radiographic features that occur during the maturation of every child (36). The set of maturity indicators is analysed with a computer program that provides the BA and the standard error for that assessment (36). However, it has not gained wide recognition.
In 2005 a digital atlas created by Vicente Gilsanz and Osman Ratib (GR) was presented to the public. It consists of artificially created idealised images of hands and wrists specific for age and sex. These images were produced by an analysis of the size, shape, morphology and density of ossification centres of 522 hand radiographs from healthy Caucasian children from Los Angeles, US (50% girls and 50% boys). Each image includes typical characteristics of development for each of the ossification centres (37). The images are of better quality and precision in comparison to GP. Another advantage is the regular spacing of the images at 6 monthly intervals from ages 2 to 6 and yearly intervals from age 7 to 17 (38). In one study GR atlas was compared to GP and they were concluded to be comparable in terms of precision, however, the study was performed on a small number of examinations (39).

Ultrasound assessment

There has been also some research performed to establish different ways of BAA, like by performing USG (40).
A result of one of such trials is BonAge® (Sunlight Medical Ltd, Tel Aviv, Israel) which consists of a device that performs an ultrasonographic examination and software that calculates the BA on its basis (19,41–44). BonAge® measures the ossifying cartilage structures of the wrist as an ultrasonic wave passes through the subject's distal radius and ulnar epiphysis. According to the producer, BonAge® provides on-the-spot, easy-to-read, immediate results, without exposing children and adolescents to ionizing x-ray radiation, what is more, it is objective and safe (41). The time of the examination is approx. 5 minutes, what can be a problem in case of the smallest children (41).
Several studies have been performed to assess the precision of this instrument. Mentzel et al. and Shimura et al. concluded that the results of BonAge® examinations are highly correlated with BA evaluated conventionally using the GP or TW2 method (42,43). However, in a more recent study performed by Khan et al. on a bigger number of patients it was shown that BonAge® tended to overread delayed BA and underread advanced BA and the authors concluded that ultrasonographic assessment should not yet be considered a valid replacement for radiographic BAA (44).
There has also been a report of ultrasonographic assessment of the thickness of anterior femoral head cartilage, which correlates strongly to child's CA and BA, standing height and body weight, according to the authors of the study (45). Ultrasonic examination of ossification of the iliac crest apophysis, i.e. Risser's sign, was also studied and it presented with high accuracy, specificity and sensitivity in comparison to hand x-ray examination and GP assessment (46).
Although the majority of the authors of the mentioned studies conclude that the ultrasonographic methods they used are of good accuracy in comparison to hand x-ray, they are rarely used in daily practice. The reason for it may be that the examination needs to be performed by a trained radiologist instead of a technician or there is a need for a specific device, in both cases, it takes more time to perform it than an x-ray. Taking into consideration that isolation of the forearm allows for minimal radiation exposure, the radiation during hand x-ray is very low (0.0005 mSv) and the fact that the studies mentioned above were performed on relatively small groups of patients, USG examination appears to be of not much clinical applicability.

MRI assessment

The first research in the field of BAA using magnetic resonance imaging was performed in 2007 to find a tool suitable to establish the age of male football players without unnecessary radiation exposure (30). Since in some Asian and African countries registration at birth is not compulsory, age determination is crucial to prevent participation in the incorrect age group (30).
In 2012 Terada et al. reported a technique of BAA based on MRI examination (47). BA was determined using an open, compact, newly designed MR imager optimized for evaluation of child's hand and wrist and it was scored by two raters using TW Japan system. Evaluation of this method was performed on a group of 93 healthy Japanese children and a strong positive correlation of BA and CA was demonstrated. What is more, the intra- and inter-rater reproducibility rates were significantly high (47).
Another study of these authors was performed in 2014 to improve the performance of this method (48). It was conducted on a group of 88 healthy children with 3 raters assessing BA and it confirmed the reliability and validity of this method (48).
However, a disadvantage of MRI is that it requires a relatively long time to be performed (2 min and 44 sec), therefore it may not be suitable for the youngest children, due to body movement.
Another study was performed by Tomei et al. and its results were published in 2014 (49). They performed hand and wrist MRIs on 179 healthy children aged 11-16 years old and analysed its correlation with chronological age. They concluded that BAA with MRI is feasible and shows good interobserver reproducibility (49).
In 2017 the results of another study were published regarding the use of MRI in BAA. Hojreh et al. performed hand MRI and x-ray examinations in 50 healthy volunteers and 10 patients, all adolescents (aged 15 ± 2 years and 13.5 ± 2.6 years respectively) and assessed both examinations according to GP criteria. They concluded that the correlation between estimated patients' ages on radiographs assessed by GP and MRI was high (the average estimated age difference between the MRIs and radiographs was −0.05/−0.175), however larger, multicenter studies are necessary to confirm the usefulness of this method (50).
There have also been attempts to automate the BAA using MRI instead of radiography (51,52).
The comparison of RTG, USG and MRI methods is presented in Table 3 (Tab. 3).

Automated techniques

Due to the mentioned problems concerning BAA with traditional methods such as inter- and intra-observer variability and the fact that it is time-consuming, a need emerged for new objective tools that would provide immediate results. In the light of

uncorrected proof

numerous attempts to use Computer-Aided Diagnosis (CAD) in the clinical practice, BA was one of the first radiologic examinations to automate. This is not a new issue, first trials date back to 1989 when a semi-automated system called HANDX was introduced by Micheal and Nelson (53). Further on, work on a PROI-Based System was published by Pietka et al. in 1991 (54). In this method, the phalangeal regions of interest (PROI) were detected and the lengths of the distal, middle, and proximal phalanx were measured automatically. BA was estimated using the standard phalangeal length table, presented earlier by Garn et al. (55).

CASAS

However, the first system that was used in studies by different authors was CASAS – a computerized image analysis system for estimating TW2 BA (56). This semi-automated system was introduced in 1994 by Tanner and Gibbons and it used the 13 bones of TW RUS system (radius, ulna and short bones) for BAA. These bones had to be located manually on the screen by a rater (correct positioning was assured by computer templates of each bone stage) and then the automatic scoring could be performed. Based on the conducted research, Tanner concluded that CASAS is more reliable and valid than manual TW RUS rating (56,57). Although also other researchers found it useful and reliable (58,59), it hasn't met a wide recognition. Its major drawback was that it took more time to estimate BA with CASAS than a manual TW assessment. Also, difficulties with BAA in case of abnormally shaped bones restricted its use in some pathological conditions.

Further on, there have been numerous approaches to BAA automation (59–71), the most important ones are described below.

BoneXpert

This automated tool for BAA was created in 2008 by Visiana company based in Holte, Denmark (73,74). This computer program analyses BA automatically, in several steps. The first step is the definition of borders and intensity of the radiologic image of 13 points of interest (of the same 13 bones like in TW RUS system – radius, ulna and 11 short bones). At this point, the system also defines if the picture is complete and of appropriate technical quality. In the next step, BA is assessed for each of the 13 bones separately. The last step is the transformation of the summary BA according to GP and TW criteria (73,74). BAA is available for ages 2.5-19 years for boys and 2-18 years for girls (2.4.7.6. version) (75). The basis for the creation of this program consisted of 1678 hand radiographs of healthy children from Denmark and children from Belgium diagnosed with a range of disorders, e.g. Turner syndrome (74).

Until today several papers have been published that verify the reliability and precision of BAA using BoneXpert in comparison to GP in different populations (Tab. 4). Among European population, a research was conducted on a group of healthy children from the Netherlands (405 patients), children from Germany suffering from short stature (1097 patients), precocious or early puberty (116 patients), congenital adrenal hyperplasia (100 patients) and with various endocrinological disturbances (514 patients) (76–80). Moreover, there was a study conducted on 1100 healthy American children from 4 different ethnic groups (Caucasian, African American, Asian and Hispanic) (23) and another on 515 eutrophic, overweight and obese children from Brasil (81). More research on the validity of BoneXpert has been performed among Asian population, including a study on 397 healthy children from Shanghai, China (82), on 6026 healthy children from 5 different cities in China (83) and among Japanese children based on 185 radiographs coming from 22 healthy children and 284 radiographs coming from 22 patients diagnosed with deficiency of growth hormone (84).

What is more, some studies confirm the validity of BAA via BoneXpert in groups of children suffering from different disorders, like juvenile idiopathic arthritis (85), in severely disabled children (86) and, as mentioned above, on children with short stature (77), precocious puberty (78) and congenital adrenal hyperplasia (79).

All the above-mentioned studies conclude that BoneXpert is a suitable tool to perform BAA, it is faster than traditional methods and eliminates rater variability. However, it is worth mentioning that one of the authors of the vast majority of the mentioned studies is a person connected to the commercial activity of Visiana company, the producer of BoneXpert.

Nevertheless, BoneXpert has several critical limitations. BA is not identified directly, the prediction depends on the relationship between CA and BA (CA is an input to the system) (63). The system is brittle and will reject radiographs when there is excessive noise, in one study it rejected 4.5% of individual bones (82). Finally, BoneXpert does not take the carpal bones into consideration, although in younger children they contain discriminative features (this has been changed in the latest version - BoneXpert 3.0 released in September 2019).

An additional feature that BoneXpert offers is a measurement of a parameter called Bone Health Index (BHI) (87), which is a unique parameter. BHI is a measurement of bone mass counted as a function of cortical thickness of three central metacarpals and their width and length. The program also automatically calculates SD values for BHI based on cohort data of Caucasian children (87). There are several research studies on the comparison of BHI values and traditional methods of bone mass measurement. In one study it was compared to the outcomes of dual-energy X-ray-absorption (DXA) and peripheral quantitative computed tomography (pQCT) in a cohort of paediatric patients from paediatric endocrine or paediatric oncology outpatient clinic and it was concluded that the BHI values showed a strong positive correlation with the DXA readings and total bone mineral density, as assessed via pQCT, also positively correlated with the BHI (88). In another study on a group of patients with juvenile idiopathic arthritis BHI measured by BoneXpert was correlated to measurements of bone mineral density by DXA, however, the correlation of Z-scores of bone mineral density measured by the two methods was weaker (89). Authors of these studies notice that a significant advantage of using BHI in comparison to DXA or pQCT is that radiation exposure is low and in low-risk peripheral areas. Also, BHI is already used in research studies on BA in patients with juvenile idiopathic arthritis (90).

What is more, there is an extension to BoneXpert - Digital X-ray Radiogrammetry (DXR) method that measures the cortical bone thickness in the shafts of the metacarpals and has demonstrated its relevance in the assessment of hand bone loss caused by rheumatoid arthritis (91).

uncorrected proof

Another advantage of BoneXpert is a prediction of the final height of a child (92,93) , which is a vital element of clinical assessment of a child with short stature. Methods used routinely until nowadays take into consideration BAA using traditional methods – GP or TW. The variability of this assessment is the main reason for the variability of predicted final height. In the case of BAA in BoneXpert we can predict final height in an objective, precise way. This program takes into consideration sex, metrical age, height and BA of a child in order to predict their final height. One can also add the height of parents and height at menarche to obtain even more reliable outcome. It is also compulsory to determine one of 9 population groups (5 within Caucasian race, Asian Chinese, Asian American, Hispanic and African American). The result of these calculations is accompanied by an SD value and the true height values will be within the indicated range with 68% probability (94) . This method was compared with height prediction using TW3 (92) .

Artificial Intelligence and Machine Learning

New possibilities of automating BAA emerged with the use of artificial intelligence (AI) and machine learning, especially its specific type known as deep learning. The most popular uses a convolutional neural network (CNN), which has already found application in areas like detection of patterns of interstitial lung disease in CT (95) or segmenting the vascular network of the human eyes on fundus photos (96) . In recent years there has been a tremendous progress in this field and there have numerous publications published on automating BAA using CNN (97–109) .

In 2017 Radiological Society of North America conducted a challenge to assess BA from paediatric hand radiographs (RSNA Pediatric Bone Age Machine Learning Challenge 2017), as part of its efforts to spur the creation of artificial intelligence tools for radiology (110,111) . The goal of the RSNA 2017 Machine Learning Challenge was to develop an algorithm which can most accurately determine BA on a validation set of paediatric hand radiographs. The results were evaluated by determining the mean difference and the mean absolute difference (MAD) between the system's performance and the mean of all reviewers' estimates. A company called 16 Bit entered the competition and achieved MAD of 4.265 months and concordance correlation coefficient (CCC) of 0.991 placing them 1st in the competition (112) . The training data set available for competitors contained 12612 images from two U.S. hospitals with a minimum age of 1 month, maximum age of 19 years and mean age of 10 years and 7 months (SD 3 years 6 months) (112) . Their Paediatric Bone Age Calculator is available on website 16Bit.ai and it is free to use it, although it is provided with a comment that this application is strictly for demonstration purposes and should not be used for clinical decision making (112) .

However, this tool has already been validated by a group of Canadian researchers, who compared its results to BAA using GP atlas on a group of 213 male and 213 female patients and found that the differences between BA assessed by these two methods were not statistically significant (median difference was 0.33 years) and concluded that tool created by 16 Bit company is a suitable one for clinical use (113) .

Another attempt to automate BAA using CNN was described in 2016 by Spampinato et al. (115) They compare performance of several approaches ranging from existing off-the-shelf CNN, trough existing pre-trained on general imagery and fine-tuned ones to custom, trained from scratch only on BA radiographs (114) . All of these CNNs were tested on the same, public data set, provided in 2007 by Gertych et al. called a *Digital Hand Atlas Database System* (115) . This atlas includes 1 391 digitized left hand radiographs from evenly distributed normally developed children of Caucasian, Asian, African-American and Hispanic origin, male and female, ranged from 1 to 18 year old. Spampinato et al. conclude that the best performance was observed in case of BoNet – original, new CNN trained from scratch specifically to assess hand radiographs (114) .

Another study in this area deserving attention, as it is especially thorough and methods used have been precisely described, concerns a system called Fully Automated Deep Learning System for BAA, which has been created in 2017 by a group of researchers from Massachusetts General Hospital, Harvard Medical School. They used a pre-trained, fine-tuned convolutional neural network to create a new tool for BAA on a basis of a large number of hand radiographs (4278 for females and 4047 for males, excluding children aged 0-4 years) (116). It calculates BA and provides a result as a number with representative picture and presents 4 more pictures (of BA +1, +2, -1, -2 years), thus radiologist can verify the result and compare it with closest ones. It achieved accuracy of 57.32 and 61.4% for the female and male cohorts on held-out test images. Female test radiographs were assigned a BAA within 1 year 90.39% and within 2 years 98.11% of the time. Male test radiographs were assigned 94.18% within 1 year and 99.00% within 2 years. It should be taken into consideration, that this system does not reject malformed images (116) . These authors also compared the BAA performance of a cohort of paediatric radiologists with and without the assistance of their tool for automatic BAA (117) . They concluded that AI improves radiologist's performance at BAA by increasing accuracy and decreasing variability and root mean squared error. The best results were achieved in case of radiological assessment assisted by AI, better than in cases of AI alone, a radiologist alone, or a pooled cohort of experts (117) .

Comparison of chosen AI methods and BoneXpert is presented in Table 5 (Tab. 5). Due to small number of radiographs in training and validating data sets all the systems based on CNNs used data augmentation (increasing the number of radiographs by rotating the pictures, adding noise, etc.). In some studies authors tested more than one type of CNN, in such cases the CNN with best performance observed is described in the table.

Conclusion

For clinicians, mainly paediatric endocrinologists, it is a very important issue to assess BA as precisely as possible to be able to make the right diagnosis and monitor closely the development of a child, the progress of a disease or effects of treatment. The traditional methods used until nowadays have very significant drawbacks – they are very time consuming, there is a high inter- and intra-rater variability (what makes it difficult to compare the examinations of one patient performed at different times) and there is a need to possess a physical copy of the atlas. The new automated techniques of BAA provide us with instant results, eliminate inter- and intra-rater variability and all only need is access to the software. A lot of research is being performed at this time in this field and the results are very promising. Most of the mentioned programs have been validated in clinical studies in

uncorrected proof

comparison to traditional BAA and they show very good precision while eliminating inter- and intra-rater variability and providing us with instant results. There are already some options available for wide clinical use like BoneXpert and Paediatric Bone Age Calculator on 16Bit.ai. It is to be expected that these automated tools are going to gain more and more recognition worldwide and soon the traditional atlases will be put aside.

Ethics

Ethics Committee Approval: -

Informed Consent: -

Authorship Contributions

Surgical and Medical Practices: -

Concept: Elżbieta Jurkiewicz, Mieczysław Szalecki, Elżbieta Moszczyńska, Monika Prokop-Piotrkowska

Design: Elżbieta Jurkiewicz, Mieczysław Szalecki, Elżbieta Moszczyńska, Monika Prokop-Piotrkowska

Data Collection or Processing: -

Analysis or Interpretation: -

Literature Search: Monika Prokop-Piotrkowska, Kamila Marszałek-Dziuba

Writing: Monika Prokop-Piotrkowska, Kamila Marszałek-Dziuba

Conflict of Interest: No conflict of interest

Financial Disclosure: No financial disclosure

References

1. Reinehr T, de Sousa G, Wabitsch M. Relationships of IGF-I and androgens to skeletal maturation in obese children and adolescents. J Pediatr Endocrinol Metab. 2006;19(9):1133–40.
2. Phillip M, Moran O, Lazar L. Growth without growth hormone. J Pediatr Endocrinol Metab. 2002;15 Suppl 5:1267–72.
3. Cox LA. The biology of bone maturation and ageing. Acta Paediatr Suppl [Internet]. 1997;423(3):107–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9401555
4. Demirjian A, Goldstein H, Tanner JM. Demirjian, A., A NEW SYSTEM OF DENTAL AGE ASSESSMENT , Human Biology, 45:2 (1973:May) p.211. Hum Biol. 1973;45(2):211–27.
5. Sehrawat JS, Singh M. Willems method of dental age estimation in children: A systematic review and meta-analysis. J Forensic Leg Med [Internet]. 2017;52:122–9. Available from: http://dx.doi.org/10.1016/j.jflm.2017.08.017
6. Martin DD, Wit JM, Hochberg Z, Sävendahl L, van Rijn RR, Fricke O, et al. The Use of Bone Age in Clinical Practice – Part 1. Horm Res Paediatr [Internet]. 2011;76(1):1–9. Available from: https://www.karger.com/Article/FullText/329372
7. Spadoni GL, Cianfarani S. Bone age assessment in the workup of children with endocrine disorders. Horm Res Paediatr. 2010;73(1):2–5.
8. Kim SE, Jang JW, Ahn MB, Kim S-H, Cho WK, Cho KS, et al. The association between skeletal maturation and adrenal androgen levels in obese children and adolescents. Ann Pediatr Endocrinol Metab. 2017;22(2):108.
9. Menjívar C, Perreira KM. Undocumented and unaccompanied : children of migration in the European Union and the United States. 2019;9451.
10. Kaur G, Khandelwal N, Jasuja OP. Computed tomographic studies on ossification status of medial epiphysis of clavicle : Effect of slice thickness and dose distribution . Original research paper Computed Tomographic Studies on Ossification Status of Medial Epiphysis of Clavicle : Effect of. 2010;(October).
11. Schmidt S, Mühler M, Schmeling A. Magnetic resonance imaging of the clavicular ossification. 2007;321–4.
12. Hillewig E, Tobel J De, Cuche O, Vandemaele P. Magnetic resonance imaging of the medial extremity of the clavicle in forensic bone age determination : a new four-minute approach. 2011;757–67.
13. Bitan FD, Veliskakis KP, Campbell BC. Differences in the Risser Grading Systems in the United States and France. 2005;(436):190–5.
14. Wittschieber D, Vieth V, Domnick C, Pfeiffer H, Schmeling A. The iliac crest in forensic age diagnostics : evaluation of the apophyseal ossification in conventional radiography. 2013;473–9.
15. Kreitner K, Schweden FJ, Riepert T, Nafe B, Thelen M. Musculoskeletal radiology Review article Bone age determination based on the study of the medial extremity of the clavicle. 1998;1122:1116–22.
16. Poland J. Skiagraphic atlas showing the development of bones of the wrist and hand. London Smith, Elder Co. 1898;
17. Greulich, William Walter; Pyle SI. Radiographic atlas of skeletal development of the hand and wrist. Stanford. Stanford University Press; 1959.
18. Tanner, JM; Whitehouse, RH; Healy M. A new system for estimating skeletal maturity from the hand and wrist with standards derived from a study of 2600 healthy british children. 1962. Paris: Centre International de L'enfance.
19. Satoh M. Bone age: assessment methods and clinical applications. Clin Pediatr Endocrinol case reports Clin Investig Off J Japanese Soc Pediatr Endocrinol [Internet]. 2015;24(4):143–52. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4628949&tool=pmcentrez&rendertype=abstract
20. Mughal AM, Hassan N, Ahmed A. Bone age assessment methods: A critical review. Pakistan J Med Sci. 2014;30(1):211–5.
21. Greulich WW, Pyle SI. Radiographic atlas of the hand and wrist. Stanford: Stanford University Press; 1959.
22. Bull RK, Edwards PD, Kemp PM, Fry S, Hughes IA. Bone age assessment : a large scale comparison of the Greulich and Pyle , and Tanner and Whitehouse (TW2) methods. 1999;172–3.

uncorrected proof

23. Thodberg HH, Sävendahl L. Validation and reference values of automated bone age determination for four ethnicities. Acad Radiol. 2010;17(11):1425–32.
24. Johnson F, Dorst J, Kuhn J, Roche A, Davila G. Reliability of skeletal age assessments. Am J Roentgenol Radium Ther Nucl Med. 1973;118:320–7.
25. Roche AF, Rohmann G, French NY, Davila H. Effect of training on replicability of assessments of skeletal maturity (Greulich-Pyle). Am J Roentgenol Radium Ther Nucl Med. 108(3):511–5.
26. Kim SY, Oh YJ, Shin JY, Rhie YJ, Lee KH. Comparison of the Greulich-Pyle and Tanner Whitehouse (TW3) Methods in Bone Age Assessment. J Korean Soc Pediatr Endocrinol. 2008;13:50–5.
27. Euling SY, Herman-giddens ME, Lee PA, Selevan SG, Juul A. Examination of US Puberty-Timing Data from 1940 to 1994 for Secular Trends : Panel Findings. 2015;
28. Herman-giddens ME, Steffes J, Harris D, Slora E, Dowshen SA, Wasserman R, et al. Secondary Sexual Characteristics in Boys : Data From the Pediatric Research in Of fi ce Settings Network. 2012;
29. Alshamrani K. Is the Greulich and Pyle atlas applicable to all ethnicities ? A systematic review and meta-analysis. 2019;2910–23.
30. Dvorak J, George J, Junge A, Hodler J. Age determination by magnetic resonance imaging of the wrist in adolescent male football players. 2007;45–52.
31. Tanner JM, Healy M, Goldstein H, Cameron N. Assessment of skeletal maturity and prediction of adult height (TW3 method). 3rd ed. London, WB Saunders, Harcourt Publishers Ltd; 2001.
32. Beunen G, Lefevre J, Ostyn M, Renson R. Skeletal maturity in Belgian youths assessed by the Tanner-Whitehouse method (TW2). 1990;17(5).
33. Murata M. Japanese specific bone age standard on the TW2. Clin Pediatr Endocrinol. 1993;(Suppl 3):35–41.
34. Tanner J, Oshman D, Bahhage F, Healy M. Tanner-Whitehouse bone age reference values for North American children. J Pediatr. 1997;131:34–40.
35. King DG, Steventon DM, O'Sullivan MP, Cook AM, Hornsby VP, Jefferson IG, et al. Reproducibility of Bone Ages when Performed by Radiology Registrars: an Audit of Tanner and Whitehouse II versus Greulich and Pyle Methods. Br J Radiol. 1994;67(September):848–51.
36. Chumlea WMC, Roche AF, Thissen ANDD. The FELS Method of Assessing the Skeletal Maturity of the Hand-Wrist. 1989;183:175–83.
37. Gilsanz V, Ratib O. Hand Bone Age: A digital Atlas of Skeletal Maturity. Springer; 2005.
38. Adler BH, Us I. Vicente Gilsanz , Osman Ratib : Bone age atlas. 2005;2005.
39. Kaplowitz P, Srinivasan S, He J, Mccarter R, Hayeri MR, Sze R. Comparison of bone age readings by pediatric endocrinologists and pediatric radiologists using two bone age atlases. 2011;690–3.
40. Bilgili Y, Hizel S, Kara SA, Sanli C, Altinok D. Accuracy of Skeletal Age Assessment in Children From Birth to 6 Years of Age With the Ultrasonographic Version of the Greulich-Pyle Atlas. 2003;683–90.
41. https://www.beammed.com/wp-content/uploads/2017/05/boneage.pdf.
42. Mentzel H, Vilser C, Vogt S, Bo J, Kaiser WA. Assessment of skeletal age at the wrist in children with a new ultrasound device. 2005;429–33.
43. Shimura N, Koyama S, Arisaka O, Imataka M. Assessment of Measurement of Children ' s Bone Age Ultrasonically with Sunlight BonAge. 2005;14(Suppl 24):17–20.
44. Khan K, Miller B, Hoggard E, Somani A, Sarafoglou K. Application of Ultrasound for Bone Age Estimation in Clinical Practice. J Pediatr. 2009;243–7.
45. Castriota-Scanderbeg A, De Micheli V. Ultrasound of femoral head cartilage: a new method of assessing bone age. Skeletal Radiol. 1995;24:197–200.
46. Wagner UA, Diedrich V, Schmitt O. Skeletal Radiology Determination of skeletal maturity by ultrasound : a preliminary report. 1995;417–20.
47. Terada Y, Kono S, Tamada D, Uchiumi T, Kose K, Miyagi R, et al. Skeletal Age Assessment in Children Using an Open Compact MRI System. 2012;000:1–6.
48. Terada Y, Kono S, Uchiumi T, Kose K. Improved Reliability in Skeletal Age Assessment using a Pediatric Hand MR Scanner with a 0 . 3T Permanent Magnet. Magn Reson Med Sci. 2014;13(3):215–9.
49. Tomei E, Sartori A, Nissman D, Al N, Battisti S, Rubini A, et al. Value of MRI of the Hand and the Wrist in Evaluation of Bone Age : Preliminary Results. 2014;1205:1198–205.
50. Hojreh A, Gamper J, Schmook MT, Weber M, Prayer D, Herold CJ, et al. Hand MRI and the Greulich-Pyle atlas in skeletal age estimation in adolescents. 2018;
51. Ebner T, Darko S, Rene D, Horst B, Martin U. Towards Automatic Bone Age Estimation from MRI: Localization of 3D Anatomical Landmarks. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI. 2014. p. 421–8.
52. Štern D, Payer C, Urschler M. Automated age estimation from MRI volumes of the hand. 2019;58.
53. Michael DJ, Nelson AC. HANDX: a model-based system for automatic segmentation of bones from digital hand radiographs. IEEE Trans Med Imaging. 1989;8(1):64–9.
54. Pietka E, McNitt-Gray MF, Kuo ML, Huang HK. Computer-assisted phalangeal analysis in skeletal age assessment. IEEE Trans Med Imaging. 1991;10(4):616–20.
55. Garn SM, Hertzog KP, Poznanski AK, Nagy JM. Metacarpophalangeal Length in the Evaluation. Radiology. 1972;105:375–81.

uncorrected proof

56. Tanner JM, Gibbons RD. A computerized image analysis system for estimating Tanner-Whitehouse 2 bone age. Horm Res. 1994;42:282–7.
57. Tanner JM, Oshman D LG. Reliability and validity of computer-assisted estimates of Tanner-Whitehouse skeletal maturity (CASAS): comparison with the manual method. Horm Res. 1994;288–94.
58. Frisch H, Riedl S, Waldhor T. Computer-aided estimation of skeletal age and comparison with bone age evaluations by the method of Greulich-Pyle and Tanner-Whitehouse. Pediatr Radiol. 1996;26:226–31.
59. van Teunenbroek A, de Waal W, Chinafo P, Fokker M, Mulder P, de Muinck Keiser-Schrama S, et al. Computer-Aided Skeletal Age Scores in Healthy Children, Girls with Turner Syndrome, and in Children with Constitutionally Tall Stature. Pediatr Res. 1996;39:360–367.
60. Hill K, Pynsent PB. A fully automated bone-ageing system. Acta Paediatr Suppl. 1994;406:81–3.
61. Rucci M, Coppini G, Nicoletti I, Cheli D, Valli G. Automatic analysis of hand radiographs for the assessment of skeletal age: a subsymbolic approach. Comput Biomed Res. 1995;28:239–56.
62. Somkantha K, Theera-umpon N. Bone Age Assessment in Young Children Using Automatic Carpal Bone Feature Extraction and Support Vector Regression. 2011;1044–58.
63. Seok J, Hyun B, Kasa-Vubu J, Girard A. Automated Classification System for Bone Age X-ray Images. In: IEEE International Conference on Systems, Man, and Cybernetics. 2012. p. 208–13.
64. Cao F, Huang HK, Pietka E, Gilsanz V. Digital hand atlas and web-based bone age assessment : system design and implementation. 2000;24:297–307.
65. Gross G, Boone J, Bishop D. Pediatric Determination Skeletal Age : with Neural Networks '. Radiology. 1995;195:689–95.
66. Sato K, Ashizawa K, Anzo M, Otsuki F, Kaneko S, Tanaka T, et al. Setting Up an Automated System for Evaluation of Bone Age. Endocr J. 1999;46:97–100.
67. Mahmoodi S, Sharif BS, Chester EG, Owen JP. Skeletal Growth Estimation Using Radiographic Image Processing and Analysis. 2000;4(4):292–7.
68. Zhang A, Gertych A, J LB. Automatic bone age assessment for young children from newborn to 7-year-old using carpal bones. Comput Med Imaging Graph. 2007;31:299–310.
69. Tiu CHÆTJÆC. Bone age estimation based on phalanx information with fuzzy constrain of carpals. 2007;283–95.
70. Liu J, Qi J, Liu Z, Ning Q, Luo X. Automatic bone age assessment based on intelligent algorithms and comparison with TW3 method. 2008;32:678–84.
71. Tristan-Vega A, Arribas JI. A Radius and Ulna TW3 Bone Age Assessment System. IEEE Trans Biomed Eng. 2008;55(5):1463–76.
72. Jong CHTLT. A fuzzy-based growth model with principle component analysis selection for carpal bone-age assessment. 2010;(201):579–88.
73. Thodberg HH, Kreiborg S, Juul A, Pedersen KD. The BoneXpert method for automated determination of skeletal maturity. IEEE Trans Med Imaging. 2009;28(1):52–66.
74. Thodberg HH. An Automated Method for Determination of Bone Age. J Clin Endocrinol Metab [Internet]. 2009;94(7):2239–44. Available from: https://academic.oup.com/jcem/article-lookup/doi/10.1210/jc.2008-2474
75. Thodberg HH, van Rijn RR, Jenni OG, Martin DD. Automated determination of bone age from hand X-rays at the end of puberty and its applicability for age estimation. Int J Legal Med. 2017;131(3):1–10.
76. Van Rijn RR, Lequin MH, Thodberg HH. Automatic determination of Greulich and Pyle bone age in healthy Dutch children. Pediatr Radiol. 2009;39(6):591–7.
77. Martin DD, Deusch D, Schweizer R, Binder G, Thodberg HH, Ranke MB. Clinical application of automated Greulich-Pyle bone age determination in children with short stature. Pediatr Radiol. 2009;39(6):598–607.
78. Martin DD, Meister K, Schweizer R, Ranke MB, Thodberg HH, Binder G. Validation of automatic bone age rating in children with precocious and early puberty. J Pediatr Endocrinol Metab. 2011;24(11–12):1009–14.
79. Martin DD, Heil K, Heckmann C, Zierl A, Schaefer J, Ranke MB, et al. Validation of automatic bone age determination in children with congenital adrenal hyperplasia. Pediatr Radiol. 2013;43(12):1615–21.
80. Booz C, Yel I, Wichmann JL, Boettger S, Al Kamali A, Albrecht MH, et al. Artificial intelligence in bone age assessment: accuracy and efficiency of a novel fully automated algorithm compared to the Greulich-Pyle method. Eur Radiol Exp. 2020;4(1).
81. Artioli TO, Alvares MA, Macedo VSC, Silva TS, Avritchir R, Kochi C, et al. Bone age determination in eutrophic , overweight and obese Brazilian children and adolescents : a comparison between computerized BoneXpert and Greulich-Pyle methods. 2019;1185–91.
82. Zhang J, Lin F, Ding X. Maturation disparity between hand-wrist bones in a chinese sample of normal children: An analysis based on automatic bonexpert and manual greulich and pyle atlas assessment. Korean J Radiol. 2016;17(3):435–42.
83. Zhang S, Liu G, Ma C, Han Y, Shen X, Xu R, et al. Automated Determination of Bone Age in a Modern Chinese Population. 2013;2013.
84. Martin DD, Sato K, Sato M, Thodberg HH, Tanaka T. Validation of a new method for automated determination of bone age in japanese children. Horm Res Paediatr. 2010;73(5):398–404.
85. Anink J, Nusman CM, van Suijlekom-Smit LW, van Rijn RR, Maas M, van Rossum MA. Automated determination of bone age and bone mineral density in patients with juvenile idiopathic arthritis: a feasibility study. Arthritis Res Ther [Internet]. 2014;16(4):424. Available from: http://arthritis-research.biomedcentral.com/articles/10.1186/s13075-014-0424-1

uncorrected proof

86. Mergler S, de Man SA, Boot AM, Heus KGCBB de, Huijbers WAR, van Rijn RR, et al. Automated radiogrammetry is a feasible method for measuring bone quality and bone maturation in severely disabled children. Pediatr Radiol [Internet]. 2016;46(7):1017–22. Available from: http://dx.doi.org/10.1007/s00247-016-3548-4
87. Thodberg HH, Van Rijn RR, Tanaka T, Martin DD, Kreiborg S. A paediatric bone index derived by automated radiogrammetry. Osteoporos Int. 2010;21(8):1391–400.
88. Schündeln MM, Marschke L, Bauer JJ, Hauffa PK, Schweiger B, Führer-Sakel D, et al. A Piece of the Puzzle: The Bone Health Index of the BoneXpert Software Reflects Cortical Bone Mineral Density in Pediatric and Adolescent Patients. PLoS One. 2016;11(3):e0151936.
89. Nusman CM, Anink J, Van Rossum MAJ, Van Rijn RR, Maas M, Van Suijlekom-Smit LWA. Bone health assessment of patients with juvenile idiopathic arthritis: A comparison between DXA and BoneXpert. Pediatr Radiol [Internet]. 2014;44(Suppl 1):S321–2. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&CSC=Y&NEWS=N&PAGE=fulltext&D=emed12&AN=71550407
90. Twilt M, Pradsgaard D, Spannow AH, Horlyck A, Heuck C, Herlin T. Joint cartilage thickness and automated determination of bone age and bone health in juvenile idiopathic arthritis. Pediatr Rheumatol [Internet]. 2017;15(1):63. Available from: http://ped-rheum.biomedcentral.com/articles/10.1186/s12969-017-0194-9
91. Pfeil A, Thodberg HH, Renz DM, Reinhardt L, Oelzner P, Wolf G, et al. Metacarpal bone loss in patients with rheumatoid arthritis estimated by a new Digital X-ray Radiogrammetry method - Initial results. BMC Musculoskelet Disord [Internet]. 2017;18(1):1–10. Available from: http://dx.doi.org/10.1186/s12891-016-1348-5
92. Thodberg HH, Jenni OG, Caflisch J, Ranke MB, Martin DD. Prediction of adult height based on automated determination of bone age. J Clin Endocrinol Metab. 2009;94(12):4868–74.
93. Satoh M. Bone age : assessment methods and clinical applications. 2015;24(4):143–52.
94. Adult Height Predictor [Internet]. Available from: https://www.bonexpert.com/documentation/adult-height-predictor
95. Anthimopoulos M, Christodoulidis S, Ebner L, Christe A, Mougiakakou S. Lung Pattern Classification for Interstitial Lung Diseases Using a Deep Convolutional Neural Network. IEEE Trans Med Imaging. 2016;35(5):1207–16.
96. Liskowski P, Krawiec K. Segmenting Retinal Blood Vessels with Deep Neural Networks. IEEE Trans Med Imaging. 2016;0062(c):1–12.
97. Lee JH, Kim KG. Applying Deep Learning in Medical Images : The Case of Bone Age Estimation. 2018;24(1):86–92.
98. Hao P, Chokuwa S, Xie X, Wu F, Wu J, Bai C. Skeletal bone age assessments for young children based on regression convolutional neural networks. 2019;16(June):6454–66.
99. Duc T, Lee J, Shin J. Incorporated region detection and classification using deep convolutional networks for bone age assessment. Artif Intell Med [Internet]. 2019;97(April):1–8. Available from: https://doi.org/10.1016/j.artmed.2019.04.005
100. Wang F, Gu X, Chen S, Liu Y, Shen Q, Pan H, et al. Artificial intelligence system can achieve comparable results to experts for bone age assessment of Chinese children with abnormal growth and development. PeerJ. 2020;8:e8854.
101. Zhou XL, Wang EG, Lin Q, Dong GP, Wu W, Huang K, et al. Diagnostic performance of convolutional neural network-based Tanner-Whitehouse 3 bone age assessment system. Quant Imaging Med Surg. 2020;10(3):657–67.
102. Mutasa S. MABAL : a Novel Deep-Learning Architecture for Machine-Assisted Bone Age Labeling. 2018;513–9.
103. Liu Y, Zhang C, Cheng J, Chen X, Wang ZJ. A multi-scale data fusion framework for bone age assessment with convolutional neural networks. Comput Biol Med [Internet]. 2019;108(November 2018):161–73. Available from: https://doi.org/10.1016/j.compbiomed.2019.03.015
104. Ren X, Member S, Li T, Yang X, Wang S, Ahmad S, et al. Regression Convolutional Neural Network for Automated Pediatric Bone Age Assessment from Hand Radiograph. IEEE J Biomed Heal Informatics. 2018;PP(c):1.
105. Spampinato C, Palazzo S, Giordano D, Aldinucci M, Leonardi R. Deep learning for automated skeletal bone age assessment in X-ray images R. 2017;36:41–51.
106. Tong C, Liang B, Li J, Zheng Z. A Deep Automated Skeletal Bone Age Assessment Model with Heterogeneous Features Learning. 2018;6–13.
107. Kim J, Shim W, Yoon H, Hong S, JS L, Cho Y, et al. Computerized Bone Age Estimation Using Deep Learning – Based Program : Evaluation of the Accuracy and Efficiency. 2017;(December):1374–80.
108. Steenkiste T Van, Ruyssinck J, Janssens O, Vandersmissen B, Vandecasteele F, Devolder P, et al. Automated Assessment of Bone Age Using Deep Learning and Gaussian Process Regression. 2018;674–7.
109. Larson DB, Chen MC, Lungren MP, Stence N V, Langlotz CP. Performance of a Deep-Learning Neural Network Model in Assessing Skeletal Maturity on. Radiology. 2018;287(1):313–22.
110. Halabi SS, Prevedello LM, Kalpathy-cramer J, Mamonov AB. The RSNA Pediatric Bone Age Machine Learning Challenge. Radiology. 2018;290(2):498–503.
111. Siegel EL. What Can We Learn from the RSNA Pediatric Bone Age Machine Learning Challenge ? Radiology. 2018;290(2):504–5.
112. https://www.16bit.ai/bone-age.
113. Gerges M, Eng H, Chhina H, Cooper A. Modernization of bone age assessment: comparing the accuracy and reliability of an artificial intelligence algorithm and shorthand bone age to Greulich and Pyle. Skeletal Radiol. 2020;49(9):1449–57.
114. Spampinato C, Palazzo S, Giordano D, Aldinucci M, Leonardi R. Deep learning for automated skeletal bone age assessment in X-ray images. Med Image Anal [Internet]. 2017;36:41–51. Available from: http://dx.doi.org/10.1016/j.media.2016.10.010

uncorrected proof

115. Gertych A, Zhang A, Sayre J, Pospiech-Kurkowska S, Huang HK. Bone Age Assessment of Children using a Digital Hand Atlas. Comput Med Imaging Graph [Internet]. 2007;31(4–5):322–31. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3624763/pdf/nihms412728.pdf
116. Lee H, Tajmir S, Lee J, Zissen M, Yeshiwas BA, Alkasab TK, et al. Fully Automated Deep Learning System for Bone Age Assessment. J Digit Imaging. 2017;30(4):427–41.
117. Tajmir SH, Lee H, Shailam R, Gale HI, Nguyen JC, Westra SJ, et al. Artificial intelligence-assisted interpretation of bone age radiographs improves accuracy and decreases variability. 2019;275–83.

Table 1. BAA Methods

	Manual	**Automatic**
RTG	• Greulich-Pyle Atlas (17), • Tanner-Whitehouse Atlas (31), • FELS Method (37), • Gilsanz and Ratib Atlas (38).	• CASAS (56), • BoneXpert (72,73) • AI methods (98-110).
MRI	• Pediatric Hand MR Scanner (46, 47), • Method of Tomei et al. (48), • Method of Hojreh et al. (49).	• Method of Stern et al. (51).
USG	• Femoral head cartilage thickness (45), • Risser's stage (46).	• BonAge (41)

Table 2. Comparison of GP and TW methods

Atlas	Greulich-Pyle	Tanner-Whitehouse
Advantages	• widely recognised • BAA relatively quick • easy to learn	• latest version from 2001 • higher reproducibility than GP
Disadvantages	• high intra- and interrater variability • not applicable to some populations • one version since 1959	• BAA time consuming

Table 3. Comparison of RTG, USG and MRI methods

Method	RTG	USG	MRI

uncorrected proof

Advantages	• the most frequently used • many recognised atlases • easy to perform • quick • accessible • doesn't require a radiologist to perform, only to assess • automated methods available	• no X-ray exposure	• no X-ray exposure • accuracy validated in studies • there are attempts to automate BAA using MRI
Disadvantages	• X-ray exposure	• presence of radiologist required to perform • time consuming • only few studies on its accuracy	• not easily accessible • relatively time consuming (quicker than USG)

Table 4. Studies on the validity of BoneXpert vs. GP

Study	**Population**	**Validity claimed**			
Author	**Year**	**Size**	**Origin**	**Health status**	
Van Rijn et al. (76)	2009	405	Netherlands	healthy	yes
Martin et al. (77)	2008	1097	Germany	short stature	yes
Martin et al. (78)	2011	116	Germany	precocious or early puberty	yes
Martin et al. (79)	2013	100	Germany	congenital adrenal hyperplasia	yes
Booz et al. (80)	2020	514	Germany	various endocrinological disturbances	yes
Thodberg et al. (23)	2010	1100	American (4 ethnic groups)	healthy	yes
Artioli et al. (81)	2019	515	Brasil	healthy, overweight and obese	yes
Zhang et al. (82)	2016	397	Shanghai	healthy	yes
Zhang et al. (83)	2013	6026	China	healthy	yes
Martin et al. (84)	2010	44	Japan	healthy, deficiency of growth hormone	yes
Anink et al. (85)	2014	69	Netherlands	juvenile idiopathic arthritis	yes
Mergler et al. (86)	2016	95	Netherlands	severely disabled	yes

uncorrected proof

Table 5. Comparison of chosen AI methods and BoneXpert

Name of tool / author	**BoneXpert (73)**	**Spampinato et al. (114)**	**Bilbily A. and Cicero M. (112)**	**Lee. H et al. (116)**	**Van Steenkiste T. at al. (108)**	**Liu et al. (103)**
Year of creation / last update	2008/2019	2016	2017	2017	2018	2019
Method	Conventional (nondeep) Machine Learning	CNN BoNet	CNN (pre-trained Inception V3)	CNN (pre-trained GoogLeNet)	CNN (pre-trained VGGNet)	CNN (pre-trained VGGNet)
Input	radiograph, race, CA and gender	radiograph, race and gender	radiograph and gender	radiograph and gender	radiograph and gender	radiograph and gender
Data set (no. of radiographs)	1 678*	1 391 (Digital Hand Atlas)	12 611 (RSNA Challenge)	8 325	12 611 (RSNA Challenge)	1 391 (Digital Hand Atlas)
Age range (years)	2.5-19 for boys 2-18 for girls	0-18	1-19	5-18	1-19	0-18
Reported accuracy (MAD in months)	4.5 (4th place in RSNA) challenge)	9.6	4.265 (1st place in RSNA) challenge)	11.16 (females) / 9.84 (males) **	6.8	8.28

* - validation on numerous groups of patients healthy and with various conditions and of various ethnic origin (Tab. 2)
** - result reported in RMSE (root mean square error) instead of MAD

uncorrected proof

MALE STANDARD 16

SKELETAL AGE: 7 YEARS

Figure 1.

uncorrected proof

Figure 2.

uncorrected proof