Traditional and New Methods of Bone Age Assessment – An Overview

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

Monika Prokop-Piotrkowska1, Kamila Marszałek-Dziuba1, Elżbieta Moszczyńska1, Mieczysław Szalecki1,2, Elżbieta Jurkiewicz3
1Department of Endocrinology and Diabetology, Children’s Memorial Health Institute,
2Collegium Medicum, Jan Kochanowski University, Kielce, Poland
3Department 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

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).

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): height velocity, menarche, muscle mass, bone mineral mass (6). Delayed BA is typical for growth hormone deficiency, constitutional delay of growth and puberty, 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).
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 recognized 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. Those 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 4 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 the 1950s 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, -1.46 to 2.38 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**

This article presents the issue of different methods of BA assessment from the perspective of a paediatrician or paediatric endocrinologist (Tab. 1).
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 ultrasongraphic 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 apophysy, 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 99 healthy children aged 11-16 years old and analysed its correlation with chronological age. They concluded that BAA with MRI was feasible showing 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
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 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 introduced by Micheal and Nelson (53) 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.

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).
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 in 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 are 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). They compare performance of several approaches ranging from existing off-the-shelf CNN, through using 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 1391 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 years 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 previously 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 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 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).

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). They compare performance of several approaches ranging from existing off-the-shelf CNN, through using 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 1391 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 years 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 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, through using 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 1391 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 years 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).

### 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
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**

65. Tiu CHÆTJÆC. Bone age estimation based on phalanx information with fuzzy constrain of carpals. 2007;283–95.


94. Adult Height Predictor [Internet]. Available from: https://www.bonexpert.com/documentation/adult-height-predictor


Table 1. BAA Methods

<table>
<thead>
<tr>
<th>Method</th>
<th>Manual</th>
<th>Automatic</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRI</td>
<td>- Pediatric Hand MR Scanner (46, 47), - Method of Tomei et al. (48), - Method of Hojreh et al. (49).</td>
<td>- Method of Stern et al. (51).</td>
</tr>
<tr>
<td>USG</td>
<td>- Femoral head cartilage thickness (45), - Risser’s stage (46).</td>
<td>- BonAge (41).</td>
</tr>
</tbody>
</table>

Table 2. Comparison of GP and TW methods

<table>
<thead>
<tr>
<th>Atlas</th>
<th>Greulich-Pyle</th>
<th>Tanner-Whitehouse</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advantages</td>
<td>widely recognised, BAA relatively quick, easy to learn</td>
<td>latest version from 2001, higher reproducibility than GP</td>
</tr>
<tr>
<td>Disadvantages</td>
<td>high intra- and interrater variability, not applicable to some populations since 1959</td>
<td>BAA time consuming, not applicable to one version</td>
</tr>
</tbody>
</table>

Table 3. Comparison of RTG, USG and MRI methods

<table>
<thead>
<tr>
<th>Method</th>
<th>RTG</th>
<th>USG</th>
<th>MRI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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

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

<table>
<thead>
<tr>
<th>Study</th>
<th>Population</th>
<th>Validity claimed</th>
<th>Author</th>
<th>Year</th>
<th>Size</th>
<th>Origin</th>
<th>Health status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Van Rijn et al. (76)</td>
<td>2009</td>
<td>405</td>
<td>Netherlands</td>
<td>healthy</td>
<td>yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Martin et al. (77)</td>
<td>2008</td>
<td>1097</td>
<td>Germany</td>
<td>short stature</td>
<td>yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Martin et al. (78)</td>
<td>2011</td>
<td>176</td>
<td>Germany</td>
<td>precocious or early puberty</td>
<td>yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Martin et al. (79)</td>
<td>2013</td>
<td>104</td>
<td>Germany</td>
<td>congenital adrenal hyperplasia</td>
<td>yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Booz et al. (80)</td>
<td>2020</td>
<td>514</td>
<td>Germany</td>
<td>various endocrinological disturbances</td>
<td>yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thodberg et al. (23)</td>
<td>2010</td>
<td>1100</td>
<td>American (4 ethnic groups)</td>
<td>healthy</td>
<td>yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Artioli et al. (81)</td>
<td>2019</td>
<td>515</td>
<td>Brasil</td>
<td>healthy, overweight and obese</td>
<td>yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zhang et al. (82)</td>
<td>2016</td>
<td>397</td>
<td>Shanghai</td>
<td>healthy</td>
<td>yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zhang et al. (83)</td>
<td>2013</td>
<td>6026</td>
<td>China</td>
<td>healthy</td>
<td>yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Martin et al. (84)</td>
<td>2010</td>
<td>44</td>
<td>Japan</td>
<td>healthy, deficiency of growth hormone</td>
<td>yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anink et al. (85)</td>
<td>2014</td>
<td>69</td>
<td>Netherlands</td>
<td>juvenile idiopathic arthritis</td>
<td>yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mergler et al. (86)</td>
<td>2016</td>
<td>95</td>
<td>Netherlands</td>
<td>severely disabled</td>
<td>yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Name of tool / author</td>
<td>BoneXpert (73)</td>
<td>Spampinato et al. (114)</td>
<td>Bilbily A. and Cicero M. (112)</td>
<td>Lee. H et al. (116)</td>
<td>Van Steenkiste T. et al. (108)</td>
<td>Liu et al. (103)</td>
<td></td>
</tr>
<tr>
<td>-----------------------</td>
<td>----------------</td>
<td>-------------------------</td>
<td>-------------------------------</td>
<td>-------------------</td>
<td>-----------------------------</td>
<td>-----------------</td>
<td></td>
</tr>
<tr>
<td><strong>Year of creation / last update</strong></td>
<td>2008/2019</td>
<td>2016</td>
<td>2017</td>
<td>2017</td>
<td>2018</td>
<td>2019</td>
<td></td>
</tr>
<tr>
<td><strong>Method</strong></td>
<td>Conventional (nondeep) Machine Learning</td>
<td>CNN BoNet</td>
<td>CNN (pre-trained Inception V3)</td>
<td>CNN (pre-trained GoogLeNet)</td>
<td>CNN (pre-trained VGGNet)</td>
<td>CNN (pre-trained VGGNet)</td>
<td></td>
</tr>
<tr>
<td><strong>Input</strong></td>
<td>radiograph, race, CA and gender</td>
<td>radiograph, race and gender</td>
<td>radiograph and gender</td>
<td>radiograph and gender</td>
<td>radiograph and gender</td>
<td>radiograph and gender</td>
<td></td>
</tr>
<tr>
<td><strong>Data set (no. of radiographs)</strong></td>
<td>1 678* (Digital Hand Atlas)</td>
<td>1 391 (RSNA Challenge)</td>
<td>8 325</td>
<td>12 611 (RSNA Challenge)</td>
<td>1 391 (Digital Hand Atlas)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Age range (years)</strong></td>
<td>2.5-19 for boys 2-18 for girls</td>
<td>0-18</td>
<td>1-19</td>
<td>5-18</td>
<td>1-19</td>
<td>0-18</td>
<td></td>
</tr>
<tr>
<td><strong>Reported accuracy (MAD in months)</strong></td>
<td>4.5 (4th place in RSNA challenge)</td>
<td>9.6</td>
<td>4.265 (1st place in RSNA challenge)</td>
<td>11.16 (females) / 9.84 (males) **</td>
<td>6.8</td>
<td>8.28</td>
<td></td>
</tr>
</tbody>
</table>

* - 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