BoneXpert
Clinical Manual

Versions 3.0 & 3.1

Manufacturer:
Visiana
Agern Allé 24
2970 Hørsholm
Denmark
www.bonexpert.com
support@BoneXpert.com

BoneXpert is a registered trademark
# Table of Contents

1  Introduction ................................................................................................................................... 3  
2  Important journal papers................................................................................................................. 3  
3  Bone age ......................................................................................................................................... 3  
   3.1  The Greulich Pyle (GP) family ................................................................................................. 3  
   3.2  The Tanner-Whitehouse (TW) family ...................................................................................... 4  
   3.3  Relationship between TW and GP bone age ......................................................................... 4  
4  Carpal bone age .............................................................................................................................. 5  
5  Bone health index ............................................................................................................................. 5  
6  The precision and accuracy of BoneXpert ....................................................................................... 6  
   6.1  Precision .................................................................................................................................... 6  
   6.2  Accuracy .................................................................................................................................... 6  
   6.3  Summary .................................................................................................................................... 6  
7  Greulich-Pyle versus Tanner Whitehouse bone age ....................................................................... 7  
8  Adult height prediction ..................................................................................................................... 7  
9  References ........................................................................................................................................ 8  
10 Description of symbols used in this manual .................................................................................. 9  
11 Abbreviations .................................................................................................................................. 9
1 Introduction

This manual explains how the BoneXpert software determines the bone age (BA), the carpal bone age and the Bone Health Index (BHI). It is intended for the clinical users interested in getting insights into how the algorithm works and how it has been validated.

2 Important journal papers

The first paper on BoneXpert, describing it from a medical imaging perspective, is from 2009 (1), and a clinical review appeared the same year (2). A review including numerous validation studies appeared in 2015 (3). BoneXpert’s extension of BA to the end of puberty was presented more recently (4).

There are papers on the validation in Japanese (5) and Chinese (6) children, and in the four ethnicities in USA (7).

3 Bone age

Please refer to the indications for use (ethnicities and disorders) listed in the Instructions for Use.

BoneXpert’s main bone age is defined from the 21 bones, which have epiphyses at some point in time. These bones are the 19 short bones, radius and ulna.

The analysis of each of the 21 bones is validated in two steps:

1. The appearance of the bone in the image is checked against a model of the natural variation of its appearance, and if the match is too poor, the bone is rejected
2. The bone age value of each bone is compared to the average bone age value found among the 21 bones, and if it deviates by more than a certain threshold, the bone is rejected. This threshold is 2.4 years above bone age 7, dropping to 0.6 years at bone age 0 (zero).

If less than 8 bones are present after this double self-validation procedure, the whole image is rejected for bone age determination.

Reasons for a rejection of a bone, or the entire hand, can be

1. It is not a hand image
2. Poor quality of the image due to:
   a. Excessive image postprocessing (too much edge enhancement)
   b. Wrong pose of the hand
   c. Occlusion of a bone by stents or other objects
   d. Defects in the radiographic equipment
3. Abnormal bone structure (skeletal dysplasia, etc.)

All the published validation studies (8) were performed using BoneXpert without any human filter, so the good accuracy and precision found in these studies demonstrate that BoneXpert can be operated safely without any human interference, or by a person who is not an expert in radiology or bone age rating.

The system applies an additional validation on the BHI measurement, so it is possible that the bone age is determined but BHI is not, and there is then a proper BHI error message.

BoneXpert provides several kinds of bone age, and they are organised into two families, GP and TW.

3.1 The Greulich Pyle (GP) family

This family is based on the Greulich-Pyle method, and includes GP and Fels.

BoneXpert’s Fels bone age is an approximation to the true Fels bone age (9) computed from the GP bone age by adding 0.4 years for boys and 0.6 years for girls. This simple relation was derived from the
set of standard images provided by the originators of the Fels method for training. A report documenting this relation can be obtained from Visiana on request (Internal Report 51, version 3, August 2019).

3.2 The Tanner-Whitehouse (TW) family

The other family is the TW method, which includes the variants: TW2, TW3, and TW-Japan.

These TW bone ages variants all refer to the RUS bones (RUS is Radius, Ulna and the 11 Short bones in ray 1, 3 and 5), and they differ because they correspond to different reference populations (10). Until approx. year 2000, the so-called TW 20-bone method was quite common. It is rarely used nowadays, because the third edition of the TW-method from 2001 (11) recommended using the RUS system instead.

In the manual TW method, the rater attributes a stage to each of 13 bones. These are defined by specific maturity indicators, which are described verbally and with drawings. BoneXpert implements TW rating by predicting the stage from the average of the male and female GP bone age. The stages A through I are represented by numbers 1 through 9 and BoneXpert estimates these stages in terms of real numbers from 1.0 to 9.0.

The stage assignment was calibrated from the First Zürich Longitudinal Study including 3400 images (10). This is used to compute TW2 and TW3.

An alternative calibration of this stage assignment was made in a study of 469 images using a Japanese rater. This is used to compute TW-Japan bone ages.

From the continuous stages, the scores of the TW system are computed using the conventional TW method by interpolating linearly between the stages. The scores from the 13 bones are summed to the conventional Sum Maturity Score (SMS). Finally, the SMS is converted into TW2, TW3 or TW-Japan bone age using the published conversion tables.

There is a fourth variant, TW3x, which extends the range to 19 years for boys and 18 years for girls. When bone age is below 15.5 for boys and below 14 for girls, TW3 and TW3x are identical. Above these bone ages, TW3x becomes increasingly equal to GP BA. The intended audience for this variant is a country (like India), where TW3 is the official bone age method, but where bone age determination up to 19 years is needed.

3.3 Relationship between TW and GP bone age

In the TW bone age method, radius and ulna each contributes with 20% of the information, and this is also the case for BoneXpert’s TW bone age.

In contrast, the BoneXpert GP bone age family assigns equal weight to all the 21 bones (short bones and radius and ulna), i.e. the weight on radius and ulna is 4.8% each.

As explained in section 7, it is recommended to use BoneXpert’s GP bone age family, because it is better validated than the TW bone age, and because it uses a more moderate weight on radius and ulna.

If TW is still preferred, it is recommended to use the TW3 variant, which refers to modern European and American Caucasian children. On average, GP and TW3 give approximately the same results for GP bone ages below 16.0 for boys and below 14.5 for girls (10).


4 Carpal bone age

After BoneXpert has analysed the 21 bones that have epiphyses at some point of time, it analyses the carpals. BoneXpert can find between 1 and 7 carpals, and these are indicated by bone contours and bone age values. An overall “carpal bone age” is reported in the result box. Carpal bone age analysis is usually possible up to bone age 11.5 years for males and 9.5 years for girls. For boys with bone age less than 0.3 years and girls below 0.15 years the carpals have usually not formed yet, and the carpal bone age is therefore typically not reported.

The overall carpal bone age is not always equal to the average of the found carpals. For instance, if the first two carpals, Capitate and Hamate, have appeared and have bone age 4.0, one would expect that carpal 3 has also appeared, since it normally appears around bone age 2.8. So, if this carpal is absent, its absence indicates a maturity of 2.8 years or less. BoneXpert takes this indication into account by adding a “ghost carpal” with bone age 2.8 and including it in the average, so that the overall carpal bone age becomes \((4+4+2.8)/3 = 3.6\) years. Ghost carpals are not shown in the annotation.

5 Bone health index

BoneXpert determines the amount of cortical bone in the three middle metacarpal shafts based on radiogrammetry, as shown here (12):

![Image of bone health index](image)

The faint red points delineate the boundaries of the cortex

From the cortical thickness \(T\), the shaft width \(W\), and the bone length \(L\), the transverse area of the cortical bone is computed with the following formula derived for a cylindrical bone shape:

\[
A = \pi T W (1 - T/W)
\]

The Bone Health Index (it was called Pediatric Bone Index (PBI) in (12)) is defined as

\[
BHI = A / (W^{1.33} L^{0.33})
\]

BHI has the unit \(\mu m^{0.33}\), but it is usually quoted as an index with no units. The division by powers of bone width and length is designed to compensate for the highly variable size of children. The BHI has a small relative standard deviation (7.5%) for children of the same bone age and sex, so BHI is sensitive to detection of children with a deviation in the bone content. To interpret a BHI value, it is expressed as a standard deviation score (SDS) relative to children of the same bone age and gender: a BHI SDS around 0 means that the BHI is typical of healthy children. The SDS is computed relative to a Dutch Caucasian population of normal children from 1997 (12). For bone age below 6 years, the Dutch children are supplemented by data from normal Parisian children born 1955 followed from age 1 month.

For bone age above 15 and 17 for boys and girls respectively, the SDS is computed relative to subjects of the same age.
6 The precision and accuracy of BoneXpert

To answer the question: “How reliable is BoneXpert?” requires definition of some concepts.

6.1 Precision

The precision of a measurement device is the ability to obtain the same result in a repeated measurement on the same subject. Firstly, if the same digital radiograph is analysed again with BoneXpert, the result is the same, so in this respect, BoneXpert’s precision is perfect. Secondly, studies have shown (13) that if a second radiograph is recorded on the child after repositioning the hand, the standard deviation between the two determinations is 0.25 years, thus the precision of the measurement is 0.18 years (0.25/√2). Version 3 reduces this to 0.16 years (this result is still unpublished).

It is well documented that two raters can deviate considerably in their rating. The standard deviation between two raters (or two ratings by the same rater) has been found to be 0.82 years (Kaplowitch (14)). It can be smaller or larger depending on experience and the time spent with the rating, but Kaplowitch’s result seems to represent clinical practice. So, the precision of manual raters is 0.58 years (0.82/√2).

6.2 Accuracy

The accuracy of a measurement is defined as the agreement with the true value. The true bone age rating of a hand X-ray is the average of very many manual ratings.

In 2017, RSNA created a de facto benchmark for bone age accuracy by publishing 200 X-rays with associated bone age “reference values” formed as the average of six independent manual ratings by radiologists (15). (So, the reference is not exactly the same as the true rating, but it is a close, and therefore a good benchmark). It was derived that a new rating by a radiologist would have a Mean Absolute Deviation (MAD) of 7.0 months relative to the reference values, reflecting the precision of human raters. BoneXpert has MAD 4.1 months, and a root mean square error (RMSE) of 0.45 years, so BoneXpert is far more accurate than a single rating. Furthermore, it can be estimated that the average of four new manual ratings will also have MAD 4.1 months, so one can state BoneXpert rates as well as four humans working in concert.

In most validation studies, BoneXpert has been compared to a single manual rater with unknown precision, and the typical RMSE between BoneXpert and the manual rating is 0.63 years (16).

6.3 Summary

To conclude, BoneXpert’s consistency with itself is described by the precision error 0.16 years (SD), which is far better than the human rater precision of 0.58 years. BoneXpert’s accuracy is 0.63 years (RMSE) relative to a single human rating and 0.45 years relative to the average of six human ratings.

The precision error is particularly important in longitudinal studies, where one wants to monitor the changes of bone age for disease management.
7 Greulich-Pyle versus Tanner Whitehouse bone age

It is recommended to use BoneXpert’s GP bone age rather than the TW bone age, for following reasons:

1) It is unnatural to force bone maturity into 9 stages, since maturation is a genuinely continuous process. The use of continuous stages in BoneXpert alleviates this issue to some extent, but far from completely.

2) TW uses a remarkably large weight, 20%, on each of radius and ulna. This makes the TW system more vulnerable to errors in the interpretation of the maturity of these two bones, so TW exhibits poorer precision than GP.

3) The prediction of adult height has been found to be more accurate when using a smaller weight on radius and ulna, so GP bone age gives more accurate prediction than TW rating. This holds for both manual and BoneXpert rating.

4) The TW3 bone age scale terminates before the bones are fully mature, namely at 16.5 and 15 years for boys and girls respectively, whereas the GP system continues to 19 and 18 years respectively.

5) The manual TW3 bone age scale is not defined below 2 years, and therefore BoneXpert’s implementation of TW3 also cannot give values lower than 2 years. In contrast, the GP scale starts at new-borns, also in the BoneXpert implementation.

6) The number of training examples used for BoneXpert’s GP bone age method is now (Version 3 of BoneXpert, October 2019) 24000, while there are only 3400 for the TW bone age. So BoneXpert is better trained to do GP than to do TW.

7) It has been found that different TW raters can show large systematic differences – of the order of one year. BoneXpert’s TW was trained on data from a group of four raters working closely together in Zurich, so it is likely that TW raters in other cities exhibit a systematic difference with the Zurich raters, and therefore BoneXpert’s TW rating. In contrast, no such effect has been seen with GP rating, and furthermore, BoneXpert’s GP rating was estimated on data from more than 8 raters from five cities, including both Europe and USA. So, one expects, in general, a far better agreement between BoneXpert and manual rating with GP than with TW.

With all these arguments in favour of the GP method, it is remarkable that many users of BoneXpert still prefer TW3. It is presumable based on the good reputation that the TW method gained in the era of manual rating, where the TW method was found to convey standardization of the rating, by forcing a certain relative weight of the bones, and by forcing the rater to rate 13 bone individually, thereby also using more time. However, these advantages do not carry over to the automated world, since BoneXpert analyses all bones individually for the GP method as well as for the TW method, and the standardization is imminent for both methods. TW is left with only disadvantages.

8 Adult height prediction

This section is only relevant for BoneXpert Server version.

As an extension of the workflow, one can use the BoneXpert online adult height predictor tool. The simplest workflow is this:

- The clinician requests bone age
- Radiology provides bone age with BoneXpert Server
- The requesting clinician receives the result from radiology, either via the PACS viewer or some electronic reporting or journal system
- The requesting clinician accesses the tool at www.BoneXpert.com/ahp and enters the bone age, chronological age and height of the patient
- The result can be stored as a PDF to be used for reporting.

The tool is described in more detail in the manual available here:

www.BoneXpert.com/docs/BoneXpert_AdultHeightPredictor_v1.1.pdf
9 References


10 Description of symbols used in this manual

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Name</th>
<th>Description of symbols used in this manual</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Manufacturer</td>
<td>The device is produced by the manufacturer whose name and address are stated next to the symbol. It indicates the medical device manufacturer, as defined in EU Directives 93/42/EEC. Date of manufacture is listed below.</td>
</tr>
<tr>
<td>2019-10</td>
<td>Electronic instruction for use</td>
<td>Indicates that the user needs to consult the instructions for use. The instructions for use can be found under the “Help (PDF)” menu/button in the BoneXpert Server Control Panel.</td>
</tr>
<tr>
<td></td>
<td>CE mark</td>
<td>The device complies with Medical Device Directive 93/42/EEC.</td>
</tr>
</tbody>
</table>

11 Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BA</td>
<td>Bone age</td>
</tr>
<tr>
<td>BHI</td>
<td>Bone Health Index</td>
</tr>
<tr>
<td>DICOM</td>
<td>Digital Imaging and Communications in Medicine</td>
</tr>
<tr>
<td>GP</td>
<td>Greulich-Pyle</td>
</tr>
<tr>
<td>MAD</td>
<td>Mean Absolute Deviation</td>
</tr>
<tr>
<td>RMSE</td>
<td>Root Mean Square Error</td>
</tr>
<tr>
<td>SD</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td>SDS</td>
<td>Standard Deviation Score</td>
</tr>
<tr>
<td>TW</td>
<td>Tanner-Whitehouse</td>
</tr>
</tbody>
</table>