

# COMPARISON OF MANUAL AND AUTOMATED BONE AGE ASSESSMENT IN 1285 CHILDREN AND ADOLESCENTS AGED 5 TO 16 YEARS

K. Maratova, D. Zemkova, J. Lebl, O. Soucek, S. Pruhova, S. Kolouskova, M. Snajderova, H. Krasnicanova, Z. Sumnik  
Department of Pediatrics Second Faculty of Medicine and Motol University Hospital, Prague, Czech Republic

## BACKGROUND

Skeletal maturation is the most reliable indicator of biological age in children and adolescents. The evaluation of hand and wrist X-Ray according to Tanner-Whitehouse (TW3) or Greulich-Pyle (GP) are the most commonly used methods for biological age assessment. Automated bone age assessment has recently become increasingly popular, however a large independent study comparing automated and manual evaluation of bone age is still missing.

## AIMS

1. The assessment of the differences between automated and manual evaluation of bone age using TW3 and GP method.
2. The assessment of applicability of automated bone age assessment in clinical practice.

## PATIENTS AND METHODS

In this cross-sectional study we evaluated bone age scans using TW3 and GP methods in 1285 children and adolescents (659 boys, range 5.0 - 15.9 years, median 10.3, IQR 4.9 years) with various endocrine conditions in parallel manually and using automated bone age assessment software BoneXpert version 2.4.5.1. (Visiana, Holte, Denmark).

Bone age scan of left hand and wrist was taken and the image in DICOM format was sent for evaluation. From the radiograph BoneXpert automatically reconstructs borders of 15 bones and calculates bone age for 13 bones. These values are then transformed into bone age according to GP or TW3 method. For the TW3 method BoneXpert also transforms bone age into bone age stages<sup>1</sup>.

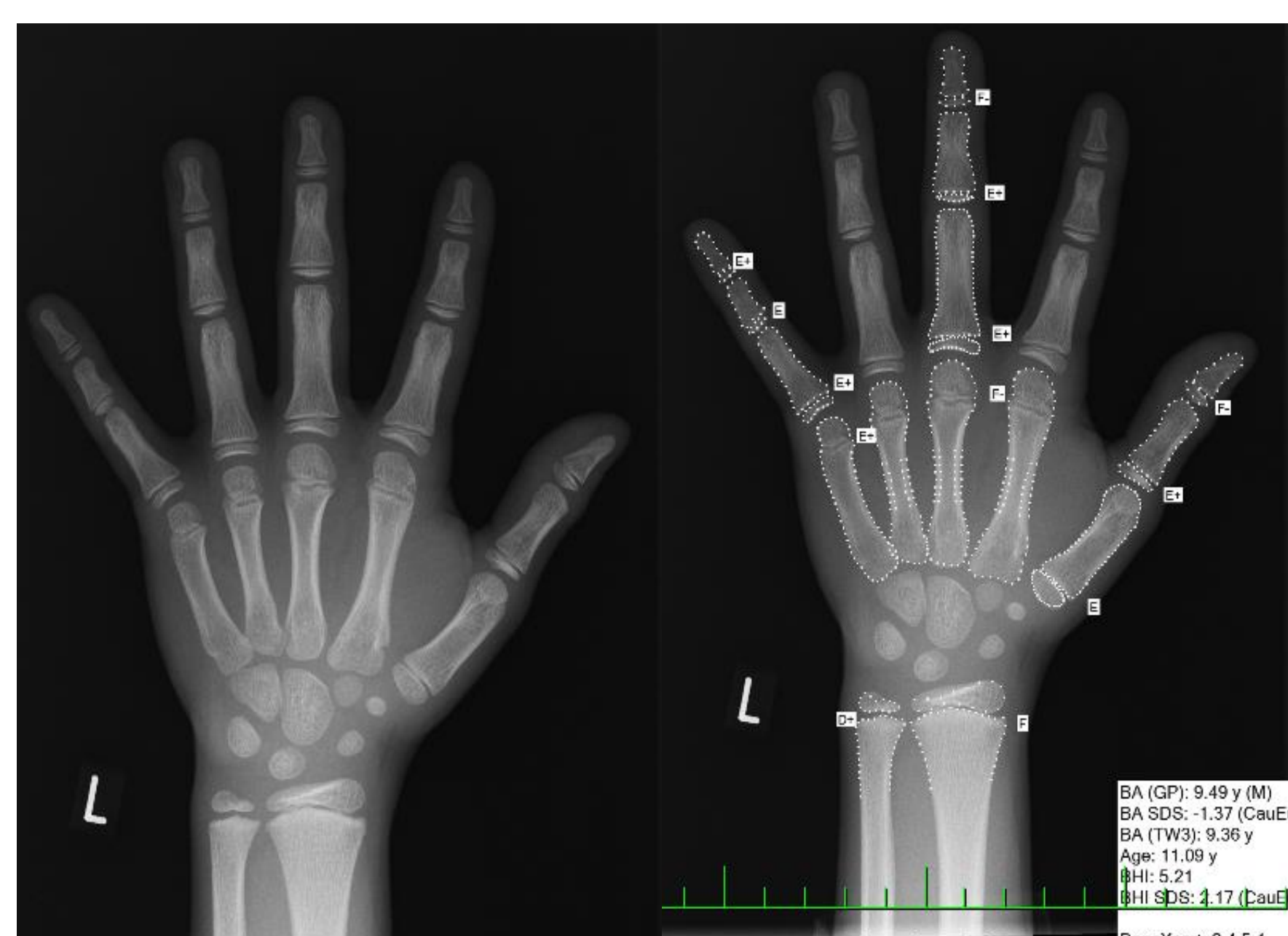


Figure 1 Bone age scan of 11.1 year old boy and the evaluation provided by BoneXpert (GP=9.5 y, TW3 = 9.4 y).

## STATISTICAL ANALYSIS

1. Root mean square errors (RMSE) were calculated for the whole group and for sex-specific one-year age categories (girls between 5 and 15 years, boys between 5 and 16 years, over 50 children in each category).
2. Sex-specific Bland-Altman analysis was performed.

## RESULTS

In total RMSE was 0.61 years and 0.58 years in boys and 0.79 years and 0.60 years in girls, respectively for TW3 and GP. Sex- and age-specific analysis showed the greatest differences between manual and automated TW3 evaluation in girls between 6-7, 12-13 and 13-14 years with RMSE 0.90, 0.90 and 1.05 years, respectively. Manual and automated evaluation differed by more than 1 year in 9.7% and 7.0% boys and 18.2% and 8.6% girls, respectively for TW3 and GP.

Table 1 Differences between manual and automated bone age assessment calculated using RMSE analysis (data are in years).

| Age categories | Boys (N = 659) |      | Girls (N = 626) |      |
|----------------|----------------|------|-----------------|------|
|                | TW3            | GP   | TW3             | GP   |
| 5-6            | 0.52           | 0.55 | 0.66            | 0.55 |
| 6-7            | 0.61           | 0.43 | 0.90            | 0.74 |
| 7-8            | 0.60           | 0.58 | 0.63            | 0.61 |
| 8-9            | 0.69           | 0.63 | 0.65            | 0.58 |
| 9-10           | 0.65           | 0.80 | 0.57            | 0.58 |
| 10-11          | 0.56           | 0.65 | 0.86            | 0.49 |
| 11-12          | 0.72           | 0.71 | 0.75            | 0.50 |
| 12-13          | 0.52           | 0.62 | 0.90            | 0.65 |
| 13-14          | 0.60           | 0.48 | 1.05            | 0.63 |
| 14-15          | 0.58           | 0.36 | 0.88            | 0.70 |
| 15-16          | 0.66           | 0.46 | -               | -    |

Sex-specific Bland-Altman analysis showed that the automated bone age assessment tends to be underestimated when compared to manual rating in boys and girls alike. Differences between automated and manual rating are more pronounced in girls during puberty when using TW3 method. Further analysis is required to determine the source of these differences, for the definition of bone stages according to the TW3 method is sex-independent.

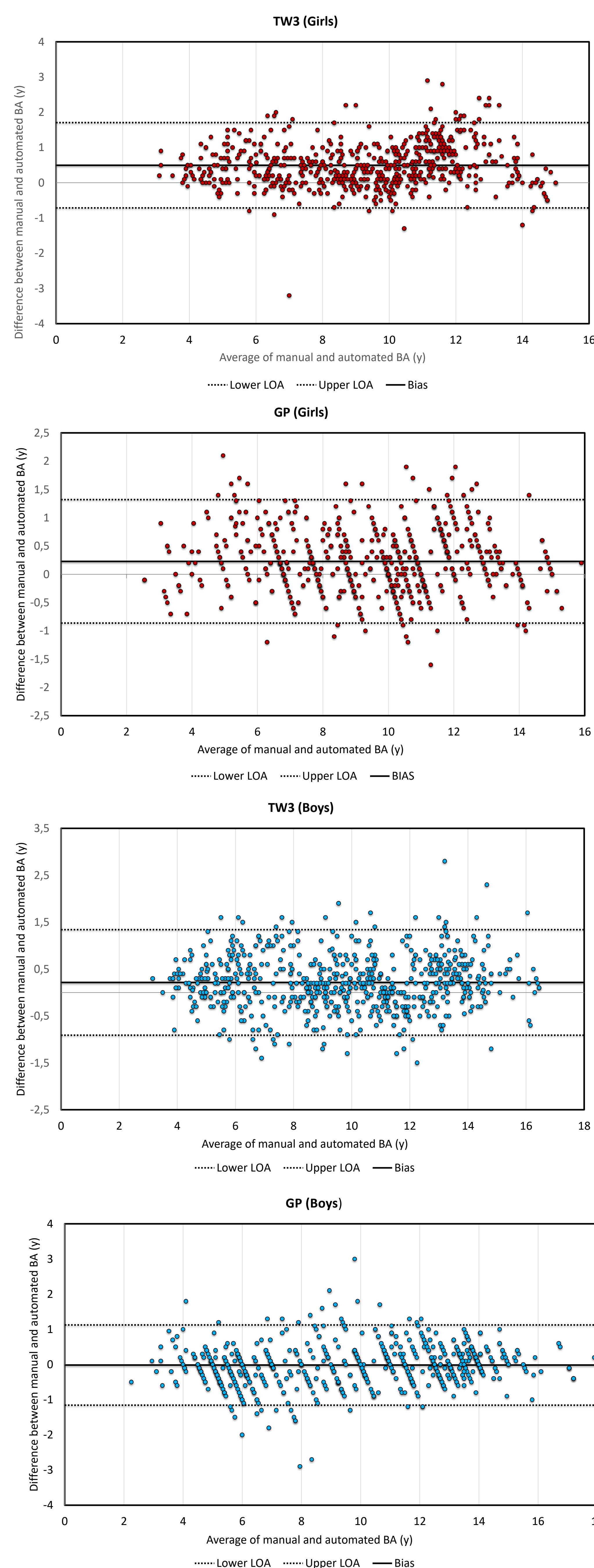


Figure 2 A-D Sex-specific Bland-Altman plot of the relation between automated and manual bone age assessment for TW3 and GP method. The difference between automated and manual assessment is shown against the average of the two methods. The dotted lines indicate the upper and lower limits of agreement (LOA).

## CONCLUSIONS

Automated bone age assessment provides sufficient agreement with manual evaluation in most scans of children with common endocrine disorders. Bone age assessment provided by BoneXpert tends to be underestimated, especially in girls during puberty using TW3 method. Further analysis is required to identify the source of these differences.

<sup>1</sup>H.H. Thodberg, S. Kreiborg, A. Juul, K.D. Pedersen, The BoneXpert Method For Automated Determination of Skeletal Maturity, IEEE Transactions on medical imaging, 28(1), 2009, 52-66.