Validation of an automated method (BoneXpert) for the determination of bone age in paediatric endocrinology

A single centre experience

Calder A1, Dastamani A2, Spoudeas HA2, Dattani M 2,3

1. Radiology Department, Great Ormond Street Hospital for Children, London  
2. Endocrinology Department, Great Ormond Street Hospital for Children, London  
3. GGM Programme, UCL GOS, Institute of Child Health, University College London

Background:

✓ Manual BA rating is associated with a considerable rater variability and it is difficult to ensure consistency between raters.  
✓ The BoneXpert method is an automated determination of bone age which employs deformable models of each bone to locate the bones and extracts the component of the bone appearance related to maturity in a holistic, statistical manner.  
✓ The BoneXpert has been validated on normal children and children with diagnoses typical of pediatric endocrinology.  
✓ Multiple clinical studies suggested that BoneXpert has adequate accuracy, precision, and efficiency to be clinically useful (1,2).

Objective

✓ To investigate accuracy, precision and bias of conventional methods of bone age evaluation compared with BoneXpert.

Design

✓ Comparison of bone age assessed using the method of Greulich and Pyle, Tanner-White-house (TW-RUS) or BoneXpert.  
✓ 79 radiographs of the left hand were obtained from children with various diagnoses between the ages of 4.8 to 16.6 years.  
✓ The films were rated using:  
  ▪ GP atlas method by two different raters (Human GP (HGP))  
  ▪ TW-RUS method by two different raters TW-RUS scores (Rater 1,2)  
  ▪ BoneXpert (BGP)  
✓ Results are evaluated using Rank and Intraclass correlation, and Bland-Altman analysis.

Results

✓ Human GP (HGP) versus BoneXpert GP (BGP) score shows bias towards chronological age (CA) for HGP ratings  
✓ HGP scores being on average 0.3 years higher than BGP scores  
✓ Similar results when comparing HGP with Human TW-RUS scores (Rater 1,2)  
✓ No evidence of bias in TW-RUS scores towards chronological age  
✓ Human TW-RUS scores were on average 0.51 years higher than BGP scores  
✓ Rank and intraclass correlation stats showed high correlation between all methods  
✓ Human TW RUS showed higher correlation with BGP than HGP with BGP

Conclusion

✓ The results of our study suggest that BoneXpert has adequate accuracy, precision and efficiency.  
✓ The automated method provides a reliable and efficient standard for bone age determination.  
✓ Introduction of BoneXpert into clinical practice provides precise standardized bone age determination, vital for long term outcome and comparator studies of growth and growth interventions.

References