The validation of an automated bone age assessment in girls with Turner syndrome – a pilot study.

Soucek Ondrej, Lebl Jan, Maratova Klara, Zemkova Dana, Sumnik Zdenek

Department of Pediatrics, Second Faculty of Medicine,
Charles University and Motol University Hospital, Prague, Czech Republic

FN MOTOL

Objective and hypotheses:
The aim of this study was to compare the manual and automated bone age analysis in a pilot group of girls with Turner syndrome of different age. We expected good concordance between the two methods.

Disclosure	None of the authors have any conflict of interest.

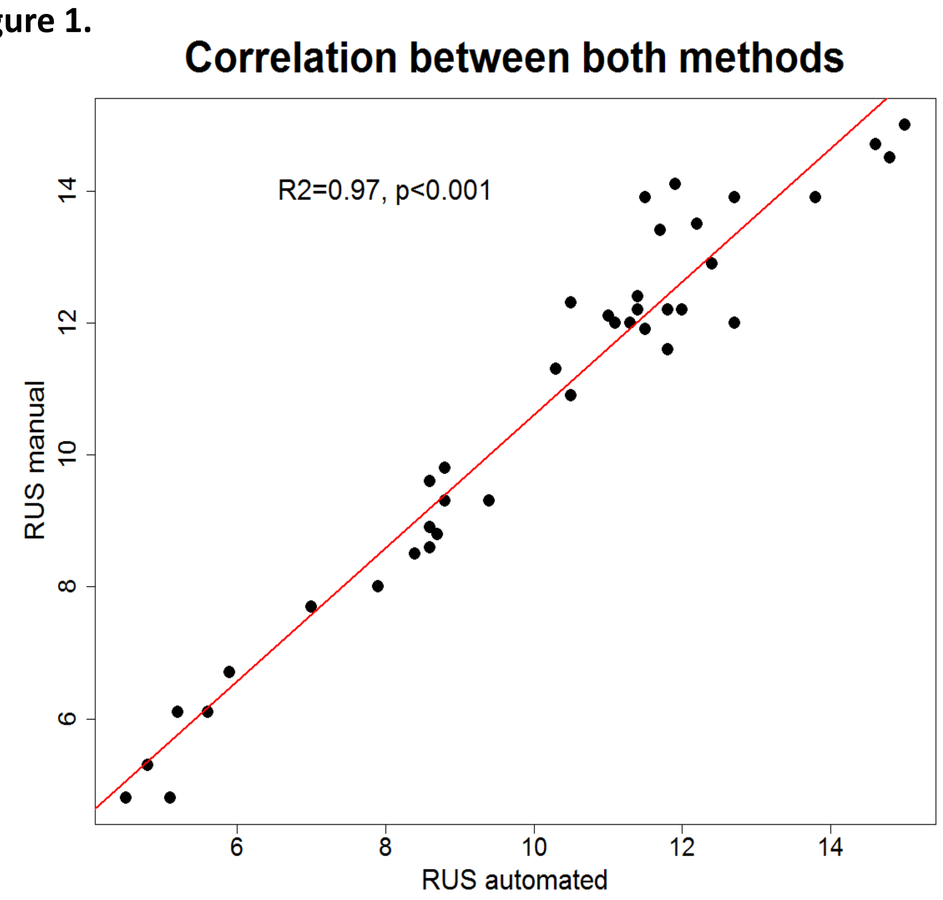
Background:
Bone age evaluation is a **basic tool** to manage the girls with Turner syndrome (TS). The **current standard** of care is to **involve an experienced medical staff** to use the Tanner Whitehouse 3 (TW3) or Greulich-Pyle (GP) method for **manual evaluation of the bone age**. As this is time consuming and may be influenced by the evaluator's skills, **automated systems may prove more efficient**.

Table 1. Anthropometric characteristics in girls with Turner syndrome.

Turner syndrome (N=39)	Mean (SD)	Median (min; max)
Age [year]	10.9 (3.0)	11.8 (4.9; 16.2)
RUS manual [year]	10.7 (2.9)	11.9 (4.8; 15.0)
RUS automated [year]	10.1 (2.8)	11.0 (4.5; 15.0)
RUS difference [year]	0.6 (0.7)	0.5 (-0.7; 2.4)
Height [cm]	138.1 (15.1)	138.7 (102.7; 166.7)
Weight [kg]	38.3 (14.9)	38.1 (14.8; 76.0)

There was no difference between the two bone age analysis methods (t-test, p=0.39).
RUS - Radius Ulna Short bones score

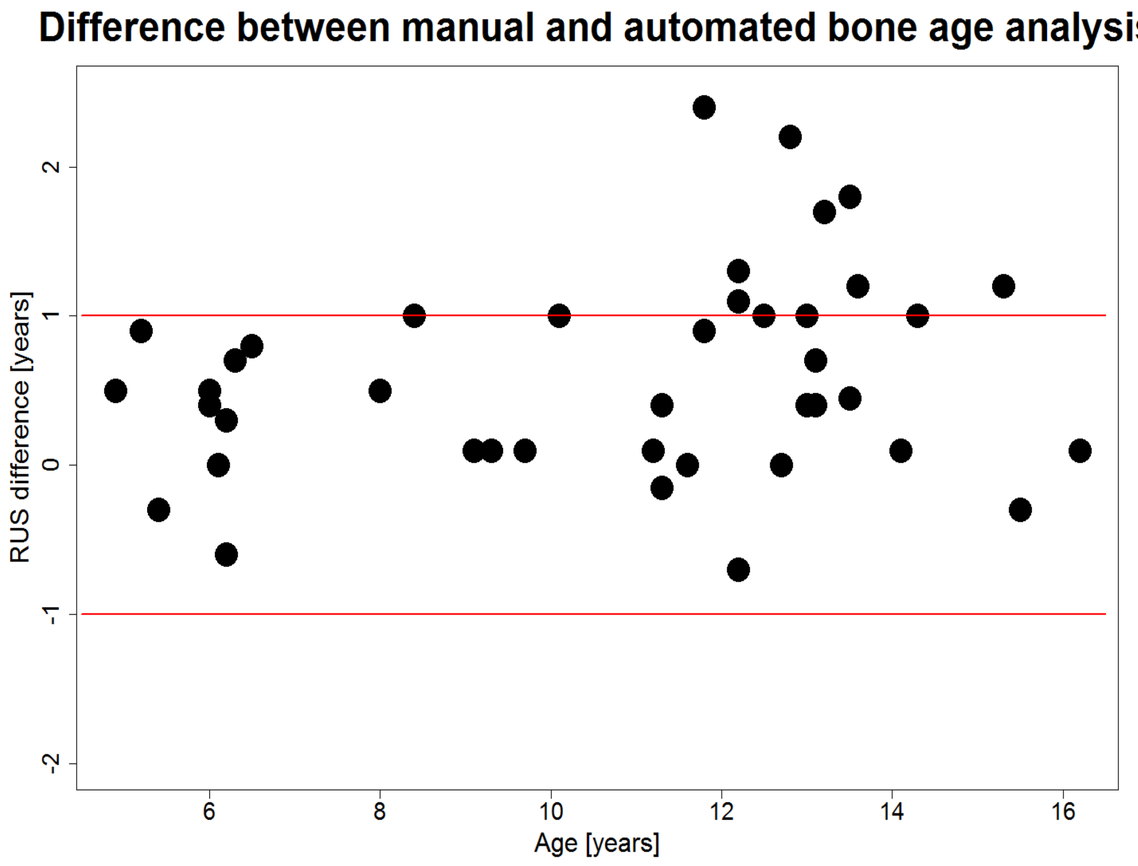
Figure 1.



Methods
The **manual bone age evaluation** was performed by an **experienced anthropologist** while using the **TW3 method**. The **BoneXpert software** (Visiana, Denmark) was used for the **automated analyses**. The **difference in the RUS** parameter between the two methods **was calculated** (t-test) and the influence of age and pubertal status was tested (multiple linear regression).

Results
Thirty nine girls with TS participated in this study. The breast stage development according to Tanner's pubertal scale was 1, 2+3 and 4+5 in 19, 12 and 10 girls, respectively. There was **no statistically significant difference between the manual and the automated RUS bone score** (Table 1, Figure 1). However, **in seven girls** (18%), the **automated software underestimated the bone age by more than 1.0 year** as compared to the manually assessed bone age values (Figure 2). **Neither the age nor the pubertal status were significantly associated with the RUS difference** in the multivariate regression model (Table 2).

Figure 2.



The RUS difference was calculated by subtracting automated from manually assessed Radius-Ulna-Short bone score.

Table 2. Multivariate regression model results.

Coefficients:	Estimate	Std Error	T value	Pr(>|t|)
(Intercept)	0.52085	0.45545	1.144	0.260
Age	-0.01549	0.05594	-0.277	0.783
Tanner (M)	0.10861	0.12150	0.894	0.377

Conclusions
The automated bone age analysis software produces similar values compared to the manual assessment. Therefore, it keeps promise for more efficiency in daily clinical routine. **However, in some girls with TS, the extent of underestimation may be of clinical concern.** We performed the **method validation on large population of different diseases** to draw the final conclusion and identify the potential pitfalls of the otherwise very convenient endocrinology tool: **please refer to poster number: P1-P033**