Construct validity of percentage of predicted adult height and BAUS skeletal age to assess biological maturity in academy soccer

Ludwig Ruf, Sean Cumming, Sascha Härtel, Anne Hecksteden, Barry Drust & Tim Meyer

To cite this article: Ludwig Ruf, Sean Cumming, Sascha Härtel, Anne Hecksteden, Barry Drust & Tim Meyer (2021) Construct validity of percentage of predicted adult height and BAUS skeletal age to assess biological maturity in academy soccer, Annals of Human Biology, 48:2, 101-109, DOI: 10.1080/03014460.2021.1913224

To link to this article: https://doi.org/10.1080/03014460.2021.1913224
Construct validity of percentage of predicted adult height and BAUS skeletal age to assess biological maturity in academy soccer

Ludwig Ruf, Sean Cumming, Sascha Härte, Anne Heckstedt, Barry Drust and Tim Meyer

Abstract

Background: The assessment of biological maturity status plays an important role in talent identification and development programs.

Aim: To compare age at predicted adult height and BAUS skeletal age as indicators of biological maturity status in youth soccer players using a construct-validity approach.

Subjects and methods: Participants were 114 players from the U12 to U17 age groups of a professional youth soccer academy. Maturity status was determined via percentage of predicted adult height based upon the Khamis-Roche method (somatic maturity) and assessed via the SonicBone BAUS® system (skeletal maturity). Convergent and known-groups validity were evaluated between maturity assessment methods and by comparing maturity-related selection biases across age groups.

Results: Although maturity status indicators were largely interrelated (r = 0.94, 95% CL 0.91–0.96), concordance (κ = 0.31 to 0.39) and Spearman’s rank-order correlations (p = 0.45–0.52) of classification methods were moderate. A selection bias towards early maturing players emerged in the U14 age group which remained relatively consistent through to the U17 age group.

Conclusions: Results confirm the construct-validity of both methods to assess biological maturity status although further validation relative to established indicators of biological maturity is needed. Furthermore, caution is also warranted when interpreting maturity status classification methods interchangeably given the poor concordance between classification methods.

Introduction

Several multidimensional and comprehensive talent identification models have been proposed, detailing the inter-relationship of potential predictors for future sports performance (Vaejens et al. 2008). Such models typically take growth and maturation characteristics of youth athletes into consideration offering a more strategic approach to talent identification, selection and development (Cumming et al. 2017). Biological maturation refers to the progress towards maturity within each biological system and can be assessed in terms of status (level of maturation at the chronological age (CA) of observation), timing (CA at which specific maturational events occur), and tempo (rate at which maturation progresses) (Malina et al. 2015). Children and adolescents of the same CA can vary substantially in terms of both maturity status and timing (Coelho-e-Silva et al. 2010; Malina et al. 2010). Objective and valid protocols to assess biological maturity are therefore required to inform and guide key stakeholders upon talent identification and selection strategies.

The assessment of biological maturity typically includes indicators of the skeletal, sexual, somatic and dental system (Malina et al. 2015). As all tissues and organ systems mature at different times and rates, biological maturity varies with the respective system considered (Beunen et al. 2006). Although skeletal maturity is widely recognised as the best indicator of maturity status (Acheson 1966), no single method can be ascribed as the “gold standard” of maturity (Malina et al. 2015). All indicators of the construct maturity status are, on average, highly interrelated during adolescence (i.e. about 10 through 16 years of age) allowing for comparisons being made between different maturity status indicators (Bielicki et al. 1984; Beunen et al. 2006). In practice, these indicators reflecting the construct of maturity status have to be both acceptable with regard to costs and ethical issues and scientifically sound in terms of measurement properties.

Commonly used and clinically established indicators of maturity status are skeletal age (SA) and secondary sex characteristics (breasts, genitals, pubic hair) (Malina et al. 2015). SA is determined by evaluating a standard radiograph of the hand-wrist bones by qualified personnel (Malina et al. 2015). Exposure to a low dose of radiation and logistical difficulties
through bone of high frequency waves of a short ultrasound
puutes the skeletal age by analysing the speed of propagation
ogy to estimate SA. Briefly, this device automatically com-
calculation of predicted maturity offset before age at peak
height velocity (PHV) (Mirwald et al. 2002) and the calcula-
status through somatic maturity indicators have therefore
environments. Non-invasive methods to estimate maturity
application (Malina et al. 2015). Several equations to predict
final adult height without an estimate of SA have been vali-
duced previously, including the Khamis-Roche method
(Khamis and Roche 1994; Coelho-e-Silva et al. 2010), Roche-
percentage of predicted adult height based on the Khamis-Roche method showed reasonable concordance with maturity status classifications based on SA in youth American football (Malina, Dompier, et al. 2007), soccer (Malina et al. 2012), roller hockey (Coelho-
and tennis players (Myburgh et al. 2019), demonstrating the construct validity of percentage of pre-
dicted adult height as an indicator of maturity status. As SA
and percentage of predicted adult height reflect different domains, yet related, aspects of the construct biological
status (skeletal vs. somatic), the adoption of both types of indicators might provide additional insight into the
maturity profile of an athlete. This highlights the need for
more practical protocols to assess skeletal maturity for coaches and researchers alike to complement the already
established and commonly applied somatic indicators (e.g.
percentage of predicted adult height based on the Khamis-
Roche method).

To address the above-mentioned limitations associated with the traditional SA assessment protocols, the non-inva-
sive portable BAUSTM system (SonicBone Medical Ltd., Israel),
SonicBone Medical Ltd., Israel, was developed using quantitative ultrasonographic technol-
ogy to estimate SA. Briefly, this device automatically com-
putes the skeletal age by analysing the speed of propagation through bone of high frequency waves of a short ultrasound
pulse and the reduction in amplitude of this ultrasound
pulse at three sites of the left hand (Rachmiel et al. 2017). To
our knowledge, the device has only been validated in com-
parison to SA derived from standard radiographs of the
hand-wrist bones using the Greulich and Pyle method in a
group of 150 male and female 4 to 17 year old patients
(10.6±3.3 years, standing height and body mass were not
provided) of a paediatric endocrinology clinic (Rachmiel et al.
2017). They found no significant bias in skeletal age between
assessment methods. In order to apply this device with confi-
dence in elite youth soccer, it is necessary to investigate the
convergent validity of this method against previously estab-
lished practical and commonly used indicators of maturity in
this population.

Previous research has shown that a selection gradient
towards early maturing athletes is evident in youth soccer.
This bias starts to emerge from 12 to 13 years of age and
tends to increase with CA (Figueiredo et al. 2009; Johnson
et al. 2017; Cumming et al. 2018; Hill et al. 2020). Children
and adolescents of the same CA can vary substantially in
terms of both maturity status and timing (Coelho-e-Silva
et al. 2010; Malina et al. 2010). Athletes advanced in bio-
logical maturity for their age tend to be, on average, taller
and heavier, show superior athletic capabilities (i.e. greater
size, strength, speed and power) and are more likely to be
selected and recruited into professional academies (Meylan
et al. 2010; Johnson et al. 2017). The ability to discriminate
between a group of individuals known to differ in a particu-
lar characteristic (i.e. known-group validity) is an important
quality criteria for measurement properties of construct valid-
ity (Prinsen et al. 2016). Therefore, replicating the same pat-
tern and magnitude for maturity-related selection biases in a
different sample of youth soccer players (Malina, Dompier,
et al. 2007; Malina et al. 2012; Myburgh et al. 2019; Hill et al.
2020) would further provide support for the construct valid-
ity of the BAUSTM system and percentage of predicted adult
height to assess biological maturity status.

In light of the previous discussion, the purpose of this
study was to establish the construct validity of the BAUSTM
system and percentage of predicted adult height based on
the Khamis-Roche method to assess the construct of bio-
logical maturity status. Our first aim was to evaluate the con-
vergent validity of the BAUSTM system to estimate the
maturity status relative to percentage of predicted adult
height derived from the Khamis-Roche method in a German
professional soccer academy. Based on observations in youth
tennis (Myburgh et al. 2019) and soccer (Malina et al. 2012)
using standard radiographs as an indicator of skeletal matur-
ity, it was hypothesised that despite large agreement for
maturity status estimates, concordance of maturity timing
classifications would be poor to moderate. Our second aim
was to determine the known-group validity of both methods
by examining the magnitude of maturity-related selection
biases across age groups. It was hypothesised that maturity-
related biases coincide with the onset of puberty and
increase with chronological age for both methods.

Subject and methods

Participants

In total, a convenience sample of 114 male youth soccer
players (age: 14.2±1.7 years, standing height: 166.3±13.0 cm,
body mass: 57.4±14.5 kg, of European ancestry n = 95, Afri-
can n = 10, Middle Eastern n = 9) from an accredited elite
youth soccer academy in Germany agreed to participate in
this study. Players were selected by the academy based on
current sport-specific qualities and future potential in terms
of technical, tactical, social and physical skills. Participants
were selected from the Under 12 (U12) to U17 age groups as
these age groups included players from different stages of
biological maturity (i.e. pre-pubescent; pubescent, post-
pubescent). This sample is representative of adolescent ath-
letes involved in youth sports as previous research indicated
advanced maturity levels across several other sports, consist-
ent with data in youth soccer (Malina 2011). Data were
collected before training sessions over the course of four weeks during the first half of the season as part of the regular anthropometry assessment (October/November 2019). Upon enrolment of each player, parents/guardians signed contracts providing consent and assent confirming that data arising as a condition of regular player monitoring procedures can be used for research purposes. The study was approved by the institutional ethics committee and was conducted according to the Declaration of Helsinki.

**Procedures**

**Biological maturity**

Standing height (±0.1 cm; KERN MPE, KERN & SOHN GmbH, Balingen, Germany) and body mass (± 0.1 kg; KERN MPE, KERN & SOHN GmbH, Balingen, Germany) were measured by accredited academy physiotherapists using standardised procedures according to the International Society for the Advancement of Kinanthropometry (ISAK) guidelines (Stewart & Marfell-Jones 2011). Percentage of predicted adult height attained at the time of the observation was used as a measure of somatic maturity status (Roche et al. 1983). Owing to the invasiveness and logistical constraints associated with the assessment of standard radiographs, percentage of predicted adult height has been selected as this method has been adopted by practitioners across various sports (Cumming et al. 2017). Although this prevents following a criterion-based validity approach, it allows the comparison of practically used methods to assess the biological maturity in applied sport settings. The Khamis-Roche method was used to predict the adult height of each participant using the participant’s standing height, body mass, chronological age at observation and self-reported mid-parental standing height (Khamis and Roche 1994).

Participant’s standing height, body mass, chronological age at observation and self-reported mid-parental standing height were used to apply the Khamis-Roche method (Khamis and Roche 1994). The associated mean error ± standard deviation at the 50th percentile of this method was 2.2 ± 0.6 cm between actual and predicted height in young males aged 4 to 18 years (Khamis and Roche 1994). Standing heights of the biological parents of each participant were self-reported and subsequently adjusted for overestimation using the equations provided by Epstein et al. (Epstein et al. 1995). Although measured parental heights might improve the prediction and strength the concordance between maturity indicators, mean-adjusted paternal and maternal heights (178.8 ± 6.5 cm and 166.9 ± 6.3 cm, respectively) were similar to parent heights in an earlier study (178.2 ± 6.6 cm and 164.9 ± 6.4 cm, respectively) (Malina et al. 2005). These mean parental values fell approximately midway between sex-specific medians and 75th percentiles for United States adults 30–39 years of age (Fryar et al. 2016) and were similar to sex-specific averages for German adults 30–40 years of age (Statistisches Bundesamt (Destatis)). Biological maturity status was then expressed as a z-score relative to age-specific means and standard deviations for boys in the Berkeley Growth Study (Bayer and Bailey 1959). Three sets of criteria were applied using the z-score to classify players: i) on time, z-score between −1.00 and +1.00; late, z-score below −1.00; early, z-score greater than +1.00, ii) on time, z-score between −0.75 and +0.75; late, z-score below −0.75; early, z-score greater than +0.75, and iii) on time, z-score between −0.50 and +0.50; late, z-score below −0.50; early, z-score greater than +0.50. Given the variety of previously used criteria, these three cut-offs were used to assess the performance of each criterion in relation to the ±1.0 year band in SA (Myburgh et al. 2019; Hill et al. 2020). Finally, percentage of predicted adult height for each participant was compared to the age- and sex-specific UK 1990 growth data (Cole et al. 1998) to derive an index of maturity status, labelled biological age (BA) to allow for direct comparison of the two maturity estimates (Myburgh et al. 2020).

The BAUS™ system (SonicBone Medical Ltd., Israel) was used to assess skeletal maturity through measuring bone density at three sites of the left hand: wrist (distal radius and ulna’s secondary ossification centres of the epiphyses), metacarpals (distal metacarpal epiphyses), and phalanx (proximal third phalanx shaft of middle finger). SA was automatically determined after each measurement by the manufacturers’ proprietary software (BAUS, v 1.0.0.12). Details of this technique have been reported elsewhere (Rachmiel et al. 2017). Briefly, the device measures two parameters: the speed of propagation through bone of high frequency waves of a short ultrasound pulse and the reduction in amplitude of this ultrasound pulse as a function of distance through the bone. Based on these parameters the SonicBone software automatically computes skeletal age based on the TW2 method (formula is protected by a nondisclosure statement) (Tanner et al. 1983). Measurements were performed by the same trained examiner (LR). Intra-examiner reliability for the three sites was assessed from 39 samples prior to data collection (radius, standardised mean difference (ES): 0.05; coefficients of variation (CV): 0.6%; intra-class correlation coefficient (ICC): 0.97; meta-carpals, ES: 0.02; CV: 0.6%; ICC: 0.97; phalanx, ES: 0.02; CV: 0.3%; ICC: 0.99). A single measurement takes between 4–6 minutes. Predicted adult height (based on the TW2 method (Tanner et al. 1983)) was also derived from the software. The difference between SA and chronological age (CA) was then used to classify players whereby an SA younger than CA by >1.0 years defined late maturity status, an SA within ±1.0 years of CA defined on-time maturity status and an SA older than CA by >1.0 years defined early maturity status (Malina 2011). The band of ±1.0 year approximated standard deviations of SA within single-year CA groups of boys aged 5–17 years in both general and athletic populations (Malina et al. 2018).

**Statistical analyses**

Normal distribution was verified before statistical analysis using by the Shapiro-Wilks test (p > .05) and visual inspection of Q-Q plots. All data are presented either as mean with standard deviations (SD) or 95% confidence limits (CL). A range of analyses were performed to assess the convergent validity of predicted absolute and percentage adult heights.
and biological age as derived from the BAUS software and the Khamis-Roche method. First, descriptive analyses were conducted to assess the mean, standard deviation, and maximum and minimum differences between both methods. Agreement and the presence of a fixed bias between both methods were assessed by calculating Pearson’s correlation coefficients ($r$) and standardised differences or effect sizes (ES, based on Cohen’s effect size principle using pooled SD) (Vet et al. 2011). Absolute and relative technical errors of measurement (TEM) as measures of error variability were also calculated to determine the absolute and relative error margin between the two methods (Ulijaszek et al. 1999). Further, cross-tabulations of maturity status classifications based on SA (BAUS™) and z-scores for the percentage of predicted adult height (Khamis-Roche method) were calculated. As in previous studies (Malina et al. 2012; Myburgh et al. 2019), percentage agreement, Cohen’s unweighted kappa coefficient ($\kappa$), and Spearman rank-order correlations ($\rho$) were computed to evaluate the concordance of maturity classifications methods. Furthermore, scatterplots and Pearson’s correlation coefficients ($r$) were calculated to evaluate the association between SA-CA differences and z-scores for the percentage of predicted adult height. The following scale was used to interpret $\kappa$: $<0.2$ as slight, $0.2–0.4$ as fair, $0.4–0.6$ as moderate, $0.6–0.8$ as substantial, and $0.8–1.0$ as almost perfect agreement (Landis and Koch 1977). Magnitudes for ES values were interpreted as follows: $<0.2$ as trivial, $0.2–0.6$ as small, $0.6–1.2$ as moderate, and $>1.2$ as large (Hopkins et al. 2009). The following scale was adopted to interpret the magnitude of $r$ and $\rho$: $<0.1$ as trivial, $0.1–0.3$ as small, $0.3–0.5$ as moderate, $0.5–0.7$ as large, $0.7–0.9$ as very large, and $>0.9$ as almost perfect association (Hopkins et al. 2009). All analyses were performed using the RVAideMemoire (version 0.9–73) and stats package (version 3.4.2) with R (version 3.6.2, R Foundation for Statistical Computing, Vienna, Austria).

### Results

Age, standing height, body mass and maturity-related characteristics for each age group are summarised in Table 1, while the cross-tabulations of maturity status classifications based on SA and percentage of predicted adult height are summarised in Table 2. Overall, percentage agreement between classification methods was 68% (95% CL 58–76%) for the total sample. The Kappa coefficient $\kappa$ was 0.37 (95% CL 0.22–0.53) for the total sample, indicating fair agreement between both maturity status classifications. Spearman rank-order correlation between maturity classification methods was moderate for the total sample, $\rho = 0.48$ (95% CL 0.32–0.62). Figure 1 shows the scatter-plot of the absolute difference between SA minus CA relative to z-scores of the percentage of predicted adult height. It also highlights the respective cut-off points for defining the maturity status for both methods and wrongly classified players. Pearson’s correlation was 0.66 (95% CL 0.54–0.76), indicating a large association between both absolute values. Corresponding analyses

<table>
<thead>
<tr>
<th>Variable</th>
<th>U12 (n = 18)</th>
<th>U13 (n = 15)</th>
<th>U14 (n = 17)</th>
<th>U15 (n = 20)</th>
<th>U16 (n = 22)</th>
<th>U17 (n = 20)</th>
<th>Total (n = 114)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA (years)</td>
<td>11.4 ± 0.3</td>
<td>12.6 ± 0.3</td>
<td>13.5 ± 0.2</td>
<td>14.6 ± 0.3</td>
<td>15.5 ± 0.4</td>
<td>16.5 ± 0.4</td>
<td>14.2 ± 1.8</td>
</tr>
<tr>
<td>Standing height (cm)</td>
<td>146.4 ± 6.2</td>
<td>153.5 ± 6.6</td>
<td>167.0 ± 8.3</td>
<td>171.1 ± 5.8</td>
<td>177.9 ± 6.6</td>
<td>174.7 ± 6.6</td>
<td>166.3 ± 1.0</td>
</tr>
<tr>
<td>Body mass (kg)</td>
<td>37.3 ± 6.2</td>
<td>41.3 ± 5.1</td>
<td>56.6 ± 9.4</td>
<td>61.2 ± 7.4</td>
<td>69.0 ± 7.5</td>
<td>70.3 ± 7.3</td>
<td>57.4 ± 14.5</td>
</tr>
<tr>
<td>SA (years)</td>
<td>11.3 ± 1.1</td>
<td>12.8 ± 1.0</td>
<td>14.9 ± 1.4</td>
<td>15.8 ± 1.0</td>
<td>17.2 ± 1.1</td>
<td>17.0 ± 0.9</td>
<td>15.1 ± 2.4</td>
</tr>
<tr>
<td>SA-CA (years)</td>
<td>-0.1 ± 1.1</td>
<td>0.2 ± 1.1</td>
<td>1.4 ± 1.3</td>
<td>1.2 ± 1.1</td>
<td>1.7 ± 0.9</td>
<td>0.5 ± 1.0</td>
<td>0.9 ± 1.2</td>
</tr>
<tr>
<td>Predicted adult height (cm) SonicBone™</td>
<td>178.7 ± 4.7</td>
<td>177.5 ± 4.2</td>
<td>181.2 ± 3.4</td>
<td>176.8 ± 2.8</td>
<td>179.9 ± 4.4</td>
<td>177.7 ± 4.4</td>
<td>178.6 ± 4.2</td>
</tr>
<tr>
<td>Predicted adult height (cm) Khamis-Roche</td>
<td>176.4 ± 6.7</td>
<td>177.4 ± 6.2</td>
<td>180.8 ± 6.4</td>
<td>179.8 ± 4.9</td>
<td>181.8 ± 5.4</td>
<td>177.3 ± 7.0</td>
<td>179.4 ± 6.3</td>
</tr>
<tr>
<td>Percentage predicted adult height (cm)</td>
<td>82.0 ± 1.7</td>
<td>86.4 ± 1.9</td>
<td>92.1 ± 3.3</td>
<td>96.8 ± 2.6</td>
<td>98.9 ± 1.9</td>
<td>98.3 ± 1.4</td>
<td>93.1 ± 6.7</td>
</tr>
</tbody>
</table>

**Table 1.** Descriptive data (mean ± SD) regarding biological maturity for each age group ($n = 114$).

<table>
<thead>
<tr>
<th>Maturity status based on percentage of predicted adult height</th>
<th>Late</th>
<th>On-time</th>
<th>Early</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>$z = 0.50$ Late</td>
<td>2</td>
<td>7</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>On-time</td>
<td>3</td>
<td>34</td>
<td>19</td>
<td>48</td>
</tr>
<tr>
<td>Early</td>
<td>0</td>
<td>19</td>
<td>38</td>
<td>57</td>
</tr>
<tr>
<td>Total</td>
<td>5</td>
<td>60</td>
<td>49</td>
<td>114</td>
</tr>
<tr>
<td>$z = 0.75$ Late</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>On-time</td>
<td>3</td>
<td>53</td>
<td>26</td>
<td>82</td>
</tr>
<tr>
<td>Early</td>
<td>0</td>
<td>5</td>
<td>23</td>
<td>28</td>
</tr>
<tr>
<td>Total</td>
<td>5</td>
<td>60</td>
<td>49</td>
<td>114</td>
</tr>
<tr>
<td>$z = 1.00$ Late</td>
<td>1</td>
<td>5</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>On-time</td>
<td>4</td>
<td>58</td>
<td>33</td>
<td>95</td>
</tr>
<tr>
<td>Early</td>
<td>0</td>
<td>1</td>
<td>16</td>
<td>17</td>
</tr>
<tr>
<td>Total</td>
<td>5</td>
<td>60</td>
<td>49</td>
<td>114</td>
</tr>
</tbody>
</table>

**Table 2.** Crosstabulation of maturity status classifications based on z-scores for percentage of predicted adult height and absolute differences between skeletal age (SA) and chronological age (CA) ($n = 114$).
for all players categorised in their respective age groups are summarised in the Supplementary file.

Agreement between predicted absolute and percentage adult heights as derived from the BAUS system and Khamis-Roche method can be considered acceptable (Table 3). Predicted adult heights from the Khamis-Roche method were slightly larger than from the BAUS system, however, Pearson’s correlation coefficients were very large. Similarly, effect sizes denoting the standardised difference between the predicted adult height from the Khamis-Roche method and the BAUS system were trivial and TEM were fair (Table 3). Finally, there was an almost perfect correlation between SA derived from the BAUS system and percentage of predicted adult height derived from the Khamis-Roche method ($r = .94, 95\%\text{CL} 0.91$ to $0.96$) (see Figure 2), although magnitudes were slightly worse when analysed by age groups (see Supplementary file).

Figure 3 illustrates the distributions of z-scores of percentage of predicted adult height (panel A) and absolute differences between SA and CA (panel B) for each age group and the total sample. SA only approximated CA in the U12 and U13 age group with the older age groups demonstrating advanced skeletal maturity status. More specifically, 78% and 65% of U12 and U13 players, respectively, were classified as on-time in maturity status. In contrast, only 42%, 40% and 27% were classified as on-time with the remaining 58%, 60% and 73% being classified as early for the U14, U15 and U16 age groups. A similar pattern was observed for z-scores of percentage of predicted adult height.

Table 3. Agreement between the BAUS software and Khamis-Roche method for predicted absolute adult height, percentage adult height and biological age ($n = 114$).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Absolute predicted adult height (cm)</th>
<th>Percentage predicted adult height (%)</th>
<th>Biological age (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean difference ± SD</td>
<td>−0.73 ± 3.4</td>
<td>0.37 ± 1.8</td>
<td>0.06 ± 1.6</td>
</tr>
<tr>
<td>Minimum and maximum difference</td>
<td>−8.7; 8.1</td>
<td>−4.8; 4.6</td>
<td>−2.7; 6.2</td>
</tr>
<tr>
<td>Pearson’s correlation coefficient (95% CL)</td>
<td>0.86 (0.81 to 0.90)</td>
<td>0.96 (0.95 to 0.98)</td>
<td>0.80 (0.72 to 0.86)</td>
</tr>
<tr>
<td>Effect size (95% CL)</td>
<td>−0.14 (−0.25 to −0.02)</td>
<td>0.06 (0.01 to 0.11)</td>
<td>0.02 (−0.09 to 0.14)</td>
</tr>
<tr>
<td>Upper and lower limits of agreement</td>
<td>−7.30 to 5.85</td>
<td>−3.08 to 3.82</td>
<td>−3.10 to 3.22</td>
</tr>
<tr>
<td>Absolute and relative (%) technical error of measurement</td>
<td>2.4; 13%</td>
<td>1.3; 1.4%</td>
<td>1.1; 1.2%</td>
</tr>
</tbody>
</table>

Discussion

The aim of the present study was to assess the construct validity of the BAUS™ system and percentage of predicted adult height based on the Khamis-Roche method to assess biological maturity status in a German professional soccer academy. Our findings show that SA as derived from the BAUS™ system and percentage of predicted adult height were almost perfectly interrelated. In contrast, concordance of maturity status classifications was relatively poor to moderate between both methods confirming our initial hypotheses. In partial support of our second aim, there was a selection bias towards players advanced in biological maturation emerging in the U14 age group which remained relatively consistent through to the U17 age group. Taken together, these findings demonstrate that both methods are able to quantify the construct of biological maturity status. However, caution is warranted when comparing the classification of athletes as early, on-time or late maturing from different maturity status classification methods given the poor concordance between methods due to the associated limitations of arbitrarily dichotomising continuous variables. It is therefore recommended to use pre-defined cut-off values as

![Figure 1. Scatterplot on the absolute difference between skeletal age (SA) and chronological age (CA) and z-scores for the percentage of predicted adult height. Three sets of criteria were applied using the z-score to classify early, on-time and late maturity status: z-score $\pm 0.50$ (upper panel, A), z-score $\pm 0.75$ (middle panel, B), z-score $\pm 1.00$ (lower panel, C). Note that black circles indicate the same maturity classification with each method, while grey triangles indicate disagreement between maturity classification methods.](image-url)
a guide and interpret differences in maturity status in conjunction with the measurement error of the respective system.

Results of the convergent validity analysis using SA derived from the ultrasound based BAUSTM system and percentage of predicted adult height derived from the Khamis-Roche method were generally comparable with studies in community-level American football (Malina, Dompier, et al. 2007), club-level soccer (Malina et al. 2012) and elite tennis (Hill et al. 2020) investigating the interrelationships between skeletal and somatic maturity using standard radiographs. Maturity status classifications based on percentage of predicted adult height had moderate concordance with classifications based on SA (Fels method) in these studies with kappa coefficients ranging from 0.22 to 0.46. Kappa coefficients of our sample were of similar magnitude (κ = 0.31 to 0.39) but varied with age (see Supplementary file). This variability is likely due to the differential tempo of maturation of different biological systems (Ratel and Williams 2017) and the application of one uniform cut-off threshold across all age groups to classify players as early, on-time and late maturing (see below). Overall agreement between classifications ranged between 57 and 70% and Spearman’s rank-order correlations were moderate (ρ = 0.45–0.52). These data are again consistent with the reported data (agreement: 57–70%, ρ = 0.27–0.47) from the previously mentioned studies (Malina, Dompier, et al. 2007; Malina et al. 2012; Myburgh et al. 2019). Moreover, large agreements and small systematic errors were observed between the BAUS software and Khamis-Roche method for predicted absolute adult height,

![Figure 2. Scatterplot on skeletal age derived from the BAUS system and percentage of predicted adult height derived from the Khamis-Roche method. Note that shaded shapes indicate different age groups.](image-url)
percentage adult height and biological age (Table 3). Corresponding data in athletic populations are lacking. While the values produced by the BAUSTM system and Khamis-Roche method assess different aspects of maturation (skeletal vs. somatic) and classifications were not expected to correspond exactly, the level of agreement between these methods for assessing the construct maturity status was acceptable. Moreover, SA derived with each assessment protocol (i.e. Greulich-Pyle, Fels, TW2, TW3), though related, are not equivalent as criteria, methods and references differ among methods (Malina et al. 2018). It is unclear how these factors influenced SA assigned by the BAUSTM system and in turn the convergent validity. Therefore, further evaluation against other maturity estimates (i.e. Greulich-Pyle, Fels, TW2, TW3) in healthy youth are needed.

Little attention has been given to the commonly used maturity classification criteria for late, on-time and early maturing players. Although the band of ±1.0 year is widely adopted in studies of youth athletes as this approximates the large inter-individual variation in skeletal maturity within age groups while the z-score of percentage of predicted adult height was based on equations for youth in southwest Ohio (Fels study (Roche 2008)), while z-scores were calculated based on percentages of predicted adult height attained at different chronological ages by a relatively small sample of boys in the Berkeley study (Bayley and Pinneau 1952). While only small differences were evident between the three cut-off criteria when analysing the entire sample, it should be noted that concordance varied substantially with age group and among cut-off criteria. More specifically, the z = 0.75 cut-off performed best for the U12 and U17 age groups while the z = 1.00 criteria was slightly superior for the U13 and U14 age groups and the z = 0.50 cut-off was most sensitive for the U15 and U16 age groups (see Supplementary file). Collectively, this highlights the need to refine cut-off criteria for classification methods as arbitrarily dichotomising continuous variables impact the sensitivity of concordance analysis (Altman and Royston 2006).


dicotomising continuous variables impact the sensitivity of

Similar to previous research, this sample of elite youth soccer players was, on average, advanced in SA relative to CA (+0.9 ± 1.2 years) and z-scores of percentage of predicted adult height (+0.43 ± 0.57, Table 1). Late-, on-time-, and early-maturing players were equally represented within the U12 and U13 age group, however, a selection bias emerged in the U14 age group and remained relatively consistent through to the U17 age group favouring early- and on-time maturing players while players late in SA were underrepresented with only one player being represented in the U17 age group. These data are consistent with previous observations of male youth players aged 11 to 17 years (Malina, Chamorro, et al. 2007; Hirose 2009; Johnson et al. 2017), although differences in SA between assessment protocols (i.e. Tanner-Whitehouse, Greulich and Pyle, and Fels method (Tanner et al. 2001)) need to be acknowledged (Malina et al. 2015). More careful observation of Figure 3(b) also highlights the large inter-individual variation in skeletal maturity within age groups with observed absolute differences of up to 4.5 years. Similar large variation in SA among players of the similar CA have been previously reported in youth soccer players (Malina 2011). This demonstrates the known-group validity of the percentage of predicted adult height method and BAUSTM system whereby both measures were able to discriminate between athletes of different maturity status that are expected to differ based on the well-established evidence on maturity-related selection bias in elite youth soccer academies (Malina et al. 2010; Johnson et al. 2017; Hill et al. 2020).

**Limitations**

Several limitations of the current study must be acknowledged. First, potential bias of self-reported standing heights (average bias: 1.4 cm in men, 0.7 cm in women (Roberts 1995)) of biological parents should be noted, although parental heights were adjusted for overestimation. The lack of a clinical established indicator of SA (i.e. hand-wrist radiograph) prevents us from drawing firm conclusions regarding the criterion validity of the BAUSTM system. However, given the associated expenses, logistical constraints and the lack of qualified individuals knowledgeable of the different assessment protocols and interpretations in the sport sciences associated with standard radiographs, the current study design aimed to compare the BAUSTM system against...
practically used methods such as the percentage of predicted adult height as an established non-invasive indicator of maturity status.

Conclusions

Results of the present study suggest that both indicators considered were largely interrelated, however, agreement between maturity classifications methods were moderate at best. Results also highlight the apparent selection bias towards players who advanced in biological maturity in a professional youth soccer academy, irrespective of the method used. Taken together this demonstrates the construct validity of the BAUSTM system and percentage of predicted adult height to assess biological maturity status in healthy youth soccer players. However, there is a need for further refinement and validation of both investigated protocols (i.e. skeletal age derived from the BAUSTM system and percentage of predicted adult height derived from the Khamis-Roche method) relative to established indicators of biological maturity in youth.

Disclosure statement

There was no conflict of interest.

Funding

No funding was provided for completion of this manuscript. Ludwig Ruf is supported by the “Science and Health in Soccer” scholarship funded by the Saarland University (GradUS program).

ORCID

Ludwig Ruf http://orcid.org/0000-0001-8589-8910
Sean Cumming http://orcid.org/0000-0003-1705-9642
Barry Drust http://orcid.org/0000-0003-2092-6962
Tim Meyer http://orcid.org/0000-0003-3425-4546

References


