Empirical Percentile Growth Curves with Z-scores Considering Seasonal Compensatory Growths for Japanese Thoroughbred Horses

Tomoaki ONODA1, Ryuta YAMAMOTO2, Kyohei SAWAMURA3, Harutaka MURASE4, Yasuo NAMBO4, Yoshinobu INOUE4, Akira MATSUI5, Takeshi MIYAKE1* and Nobuhiro HIRAI1

1Comparative Agricultural Sciences, Graduate School of Agriculture, Kyoto University, Kyoto 606-8502, Japan
2The Japan Bloodhorse Breeders’ Association, 4-5-4 Shinbashī, Minato-ku, Tokyo 105-0004, Japan
3JRA Facilities Co. Ltd., 4-5-4 Shinbashī, Minato-ku, Tokyo 105-0004, Japan
4Hidaka Training and Research Center, Japan Racing Association, 535-13 Nissha, Urakawa-cho, Hidaka, Hokkaido 057-0171, Japan
5Equine Research Institute, Japan Racing Association, 321-4 Tokami-cho, Utsunomiya, Tochigi 320-0856, Japan

Percentile growth curves are often used as a clinical indicator to evaluate variations of children’s growth status. In this study, we propose empirical percentile growth curves using Z-scores adapted for Japanese Thoroughbred horses, with considerations of the seasonal compensatory growth that is a typical characteristic of seasonal breeding animals. We previously developed new growth curve equations for Japanese Thoroughbreds adjusting for compensatory growth. Individual horses and residual effects were included as random effects in the growth curve equation model and their variance components were estimated. Based on the Z-scores of the estimated variance components, empirical percentile growth curves were constructed. A total of 5,594 and 5,680 body weight and age measurements of male and female Thoroughbreds, respectively, and 3,770 withers height and age measurements were used in the analyses. The developed empirical percentile growth curves using Z-scores are computationally feasible and useful for monitoring individual growth parameters of body weight and withers height of young Thoroughbred horses, especially during compensatory growth periods.

Key words: body weight, percentile growth curves, Thoroughbred horses, withers height

Thoroughbred horses, well-known animals in horseracing worldwide, are seasonal breeding animals. The young foals are born in spring and generally show seasonal compensatory growth (CG) patterns in which their growth rates [1, 2, 10, 20] decline in the winter and dramatically increase in the following spring, due to the coldness of winter season.

The seasonal CG pattern is clearly seen in the change in body weight or average daily gain of yearling Thoroughbreds, especially in northern regions or countries [1]. Many studies have investigated the seasonal changes of horses’ growth rates [2, 10, 16, 17, 20, 21]. Mohammed [14] reported that Thoroughbred yearlings were at a high risk of developmental disorders (e.g. osteochondrosis and orthopaedic diseases). Due to the dramatic change of the growth rate during the winter and spring seasons, the seasonal CG may be an additional risk factor for the physical development of yearling Thoroughbred foals.

Percentile growth curves, which track the values of anthropometric indices visually with percentages, are useful as a clinical indicator for monitoring growth and health conditions, and are used to evaluate variations in children’s growth status. Percentile growth curves have been widely used in many medical and health care areas for humans [9, 12, 15, 22, 24, 25] and animals [4, 13]. The construction of percentile growth curves has been investigated for Thoroughbred horses [5, 11], where percentile body weight charts for Thoroughbreds were available. The percentile...
growth curves directly or empirically considering CG patterns, however, have not yet been estimated.

Concerning body weights for young Thoroughbreds, we have proposed mathematical equations for empirical growth curves considering seasonal CG [16, 17]. These growth curve equations would be useful for the establishment of percentile growth curves considering seasonal CG, because the equations are continuous single variable functions of the age of horses, and the variances of the growth data can be used as the empirical percentiles with Z-scores under the assumption of normal distributions of data.

The establishment of percentile growth curves would give us useful indicators for feeding management in order to achieve sound musculoskeletal development or desirable body compositions in growing horses. In this study, we propose a method for constructing empirical percentile curves for body weight and withers height of Japanese Thoroughbred horses using Z-scores, considering the CG periods that are typical characteristics of seasonal breeding animals.

**Materials and Methods**

**Data description**

A total of 5,594 and 5,680 body weight (BW; kg) and age (day) measurements of 271 Thoroughbred colts and 237 Thoroughbred fillies, were collected by Hidaka Training and Research Center, Japan Racing Association (JRA) and the Japan Bloodhorse Breeders’ Association (JBBA) between 1999 and 2009. In addition, a total of 3,770 withers height (WH; cm) and age (day) measurements of 422 Thoroughbred colts and fillies was also collected. The maximum age in the BW data was about 1,100 days, an age which covers the two winter seasons with CG periods before foals’ debut in horseracing. The maximum age in the WH data was about 800 days, which covers only the first winter season with a CG period.

The data collection was mainly organized in the Hidaka region of Hokkaido, the northern island of Japan. The Hidaka region is famous for the intensive production of racehorses, and it has about 815 racehorse stud farms, corresponding to 82% of all Japanese racehorse stud farms (Hidaka Subprefectural Bureau, 2012). The ground is covered by snow in Hidaka in winter, and green pastures only begin to grow in the middle of May. Due to the coldness in the winter, the typical CG phenomenon appears in the growth of the Thoroughbred foals raised in the Hidaka region.

**Growth curves for body weight**

For body weights of colt and filly Thoroughbreds, the following mixed model equations (Equations 1 and 2) were used based on the growth curves derived by Onoda et al. [16, 17], in which the traditional Richards’ growth curve equation [18] was modified to incorporate CG effects with sigmoid sub-functions \( f(t) \) and \( f'(t) \). The sigmoid sub-functions \( f(t) \) and \( f'(t) \) adjust the first and second year CGs at 432 and 797 (=432 + 365) days of age, respectively, and were adapted to the biological parameters responsible for maturity in the Richards equation [18]. The age 432 days was determined by the crossing points of the actual data averages and the value of the traditional Richards’ equation (see Onoda et al. [16]).

In the mixed model equations (Equations 1 and 2), both individual horses \((a_i)\) and residuals \((e_{ij})\) for the individual \(i\) and age \(ij\) were included as random effects with their variance components, \(\sigma^2_{a}\) and \(\sigma^2_{e}\). The population averages of \(a_i\) and \(e_{ij}\) were both zero and they were independent of each other.

The individual weight factor \(a_i,\) was added to the maturity weight, 575.0 kg. Thus, the maturity weight of each individual was 575.0 + \(a_i\), indicating that the final maturity weight differs according to each individual. Based on Onoda et al. [16, 17], the equations for BW for male and female Thoroughbreds are:

\[
BW_{\text{male}}_i = \frac{575.0 + a_i}{\left(1.0 + \left(-0.94513 + 0.3582 \times f(t) + 0.7466 \times f'(t)\right) e^{-0.00215 \frac{t}{365}}\right)^{\frac{1}{2}}} + e_{ij} \tag{Equation 1}
\]

\[
BW_{\text{female}}_i = \frac{575.0 + a_i}{\left(1.0 + \left(-0.94513 + 0.3582 \times f(t) + 0.6705 \times f'(t)\right) e^{-0.00204 \frac{t}{365}}\right)^{\frac{1}{2}}} + e_{ij} \tag{Equation 2}
\]

where \(f(t) = \frac{1.0}{1.0 + e^{0.00215 \frac{t}{365}}} - \frac{1.0}{1.0 + e^{0.00204 \frac{t}{365}}},\)

\(f'(t) = \frac{1.0}{1.0 + e^{0.00215 \frac{t}{365}}} - \frac{1.0}{1.0 + e^{0.00204 \frac{t}{365}}},\)

\(a_i \sim N(0, \sigma^2_{a}),\)

\(e_{ij} \sim N(0, \sigma^2_{e}),\)

where \(BW_{\text{male}}_i\) and \(BW_{\text{female}}_i\) are the body weights of the individual \(i\) measured at age (time) \(t_{ij}\). The unknown variance components (i.e., \(\sigma^2_{a}\) and \(\sigma^2_{e}\)) were estimated for each sex by the SAS NLMIXED procedure [19].
Growth curve for withers height

For WH of Thoroughbreds, the growth curve equation was constructed by following the method of the growth curve constructions for BW. We combined both sexes in the WH analysis, because the difference of WH between sexes was small in our preliminary analyses (about 0.5 cm at the most), which corresponded to the small gender difference of WH of 2 cm at the most [6], and the total number of data available for WH was small. Furthermore, only the WH data of less than about 800 days which corresponds to the first seasonal CG period was available. Thus, the sub-function \( f(t) \) handling the first CG was only included in the equation (Equation 3). The \( f(t) \) was identical to that used in Equations 1 and 2 assuming that the CG affects WH at the same time as BW. The equation for WH is:

\[
WH_y = \frac{161.0 + a_e}{1.0 + (-0.8452 + 0.4304 \times f(t))} e^{-0.00209 \times t} + \epsilon_y
\]

where \( f(t) = \frac{1.0}{1 + 0.4304 \times e^{-268.49 t}} - \frac{1.0}{1 + 0.4304 \times e^{-268.49 t}} \)

\( a_e \sim N(0, \sigma^2_a) \), and \( \epsilon_y \sim N(0, \sigma^2_y) \).

The unknown variance components were also estimated by the SAS NLMIXED procedure [19].

Empirical percentile growth curves using Z-scores

In the definition of ‘percentile’, the value of 50 percentile is the median of the data, and not the average. If the data is not normally distributed, the data value of the 50 percentile is not the mean of the data, and if the data is normally distributed, the data value of the 50 percentile is identical to the mean of the data. To compute the exact 50 percentile curve for body weight, for example, numerical medians of the bodyweight data must be computed at all ages, according to Kocher and Staniar [11]. For the computation of other percentile curves (e.g. 25 percentile curves), the same manner of computation is necessary. Even though this approach to get exact percentile curves is useful for the illustration of data variations, it is generally difficult to obtain continuous single variable mathematical growth curve function of age (time).

Using Z-scores is an alternative way to obtain empirical percentile growth curves. The Z-score is a standardized score that indicates how many standard deviations a data point is apart from the population mean, based on the assumption of normal distributions. Based on the Z-scores of the estimated variance components, the empirical percentile growth curves are easily constructed. Combined with the Equations 1, 2 and 3 and the information of the estimated variance components and their Z-scores, the equations for the empirical percentile growth curves for BW (Equations 4 and 5) and WH (Equation 6) were constructed as follows:

\[
BW_{male\_percentile} = \frac{575.0 + Z_2 \sigma^2_w}{\left(1 + (0.94513 + 0.3582 \times f'(t) + 0.7466 \times f''(t)) e^{-0.00313 \times t}\right)^{1/2}} + Z_2 \sigma^2_w
\]

(Equation 4),

\[
BW_{female\_percentile} = \frac{575.0 + Z_2 \sigma^2_w}{\left(1 + (0.94880 + 0.3582 \times f'(t) + 0.6705 \times f''(t)) e^{-0.00401 \times t}\right)^{1/2}} + Z_2 \sigma^2_w
\]

(Equation 5),

and

\[
WH_{percentile} = \frac{161.0 + Z_2 \sigma^2_w}{\left(1 + (-0.8452 + 0.4304 \times f(t)) e^{-0.00209 \times t}\right)^{1/2}} + Z_2 \sigma^2_w
\]

(Equation 6).

These equations are the continuous single variable function of age (t). In these equations, Z is the Z-score that corresponds to an empirical percentile. We assigned the empirical percentile curves for body weight and withers height of Thoroughbreds as 3, 10, 25, 50, 75, 90 and 97%. Each empirical percentile corresponds to Z-scores of \(-1.881, -1.282, -0.675, 0.000, 0.675, 1.282, 1.881\), respectively. A Z-score of zero corresponds to the empirical 50 percentile curve, the Equations 4 and 5 become identical to the equations proposed by Onoda et al. [17].

Validation of the estimated percentile growth curves

For the validation of the constructed empirical percentile growth curves, the data percentages between these percentile curves were counted. For the validation analyses, the age data periods were from zero in multiples of 180 days. The percentages of the data positioned between each percentile curve intervals (i.e. 0 to 3, 3 to 10% and so on) were counted from 0 to each age period (i.e. 0 to 180, 360, 540, 720, 900 and all days).
Results

Based on Equations 1 and 2, the estimates of the variance components for the body weights of Thoroughbred colts are $\sigma^2_a = 1431.14$ and $\sigma^2_e = 191.51$, and those for the body weights of Thoroughbred fillies are $\sigma^2_a = 1431.01$ and $\sigma^2_e = 228.95$. For withers heights of both sexes, based on Equation 3, the variance components are $\sigma^2_a = 16.14$ and $\sigma^2_e = 3.52$. All variance components were significantly different from zero ($P<0.0001$). The square root of these values of the estimated variances were introduced to Equations 4, 5 and 6 for the construction of the empirical percentile growth curves.

The scatterplots of the weight- and height-age data and the estimated empirical percentile growth curves are shown in Figs. 1 and 2. These percentile curves express both the changes of the growth rate in the first and second CG periods, and the distribution of the data in the analyzed population. In these figures, the tendency of the seasonal CG can be recognized around 432 and 797 days, corresponding to the first and second CG periods, respectively. Our former studies provided clear evidence of the presence of both the first and second CG for body weight, based on model comparison of with and without CG [16, 17]. Concerning the withers height, we also confirmed the presence of the CG in the same manner, using a model comparison procedure.

The results of validation analyses for estimated percentile growth curves are shown in Tables 1, 2 and 3 for colt and filly body weights and the withers height for both sexes. In these tables, the majority of the data approximated to the over 50% curves at all ages and for all traits.

Discussion

The growth curve equation models used in this study were based on our former studies [16, 17]. In the present study the data from both JBBA and Hidaka Training and Research Center were combined in order to estimate the variance components with a larger data set. Interestingly, the
Fig. 2. The scatterplot of the height-age data and the empirical percentile growth curves for withers height of Japanese Thoroughbred horses of both sexes. The height-age data is shown with gray dots and the percentile curves are shown with black lines: from the bottom line, 3, 10, 25, 50, 75, 90 and 97 percentile curves are shown. Dashed line indicates the center of the first CG period (432 days).

Table 1. Data distributions of Thoroughbred colts’ body weights at different periods of age

<table>
<thead>
<tr>
<th>Age (days)</th>
<th>0–3</th>
<th>3–10</th>
<th>10–25</th>
<th>25–50</th>
<th>50–75</th>
<th>75–90</th>
<th>90–97</th>
<th>97–100</th>
</tr>
</thead>
<tbody>
<tr>
<td>180</td>
<td>0.06</td>
<td>0.64</td>
<td>6.07</td>
<td>28.58</td>
<td>43.41</td>
<td>17.07</td>
<td>4.16</td>
<td>0.00</td>
</tr>
<tr>
<td>360</td>
<td>0.09</td>
<td>0.55</td>
<td>5.41</td>
<td>27.10</td>
<td>40.65</td>
<td>20.43</td>
<td>5.41</td>
<td>0.37</td>
</tr>
<tr>
<td>540</td>
<td>0.07</td>
<td>0.40</td>
<td>5.36</td>
<td>25.26</td>
<td>40.38</td>
<td>22.18</td>
<td>5.70</td>
<td>0.67</td>
</tr>
<tr>
<td>720</td>
<td>0.06</td>
<td>0.39</td>
<td>4.97</td>
<td>25.16</td>
<td>40.09</td>
<td>22.33</td>
<td>5.91</td>
<td>1.10</td>
</tr>
<tr>
<td>900</td>
<td>0.06</td>
<td>0.41</td>
<td>4.85</td>
<td>25.36</td>
<td>39.15</td>
<td>22.75</td>
<td>6.23</td>
<td>1.16</td>
</tr>
<tr>
<td>All</td>
<td>0.07</td>
<td>0.39</td>
<td>4.74</td>
<td>25.01</td>
<td>39.01</td>
<td>22.99</td>
<td>6.42</td>
<td>1.29</td>
</tr>
</tbody>
</table>

Table 2. Data distributions of Thoroughbred fillies’ body weights at different periods of age

<table>
<thead>
<tr>
<th>Age (days)</th>
<th>0–3</th>
<th>3–10</th>
<th>10–25</th>
<th>25–50</th>
<th>50–75</th>
<th>75–90</th>
<th>90–97</th>
<th>97–100</th>
</tr>
</thead>
<tbody>
<tr>
<td>180</td>
<td>0.07</td>
<td>0.35</td>
<td>5.10</td>
<td>34.78</td>
<td>42.88</td>
<td>12.92</td>
<td>3.84</td>
<td>0.07</td>
</tr>
<tr>
<td>360</td>
<td>0.07</td>
<td>0.77</td>
<td>7.12</td>
<td>34.42</td>
<td>41.04</td>
<td>13.11</td>
<td>3.33</td>
<td>0.10</td>
</tr>
<tr>
<td>540</td>
<td>0.07</td>
<td>0.78</td>
<td>6.34</td>
<td>32.62</td>
<td>41.27</td>
<td>14.58</td>
<td>3.85</td>
<td>0.48</td>
</tr>
<tr>
<td>720</td>
<td>0.06</td>
<td>0.76</td>
<td>6.39</td>
<td>31.59</td>
<td>41.47</td>
<td>15.15</td>
<td>3.96</td>
<td>0.61</td>
</tr>
<tr>
<td>900</td>
<td>0.06</td>
<td>0.76</td>
<td>6.87</td>
<td>31.63</td>
<td>40.91</td>
<td>14.90</td>
<td>4.24</td>
<td>0.65</td>
</tr>
<tr>
<td>All</td>
<td>0.05</td>
<td>0.08</td>
<td>7.22</td>
<td>31.37</td>
<td>40.81</td>
<td>14.82</td>
<td>4.26</td>
<td>0.69</td>
</tr>
</tbody>
</table>

Table 3. Data distributions of Thoroughbred colts’ and fillies’ withers heights at different periods of age

<table>
<thead>
<tr>
<th>Age (days)</th>
<th>0–3</th>
<th>3–10</th>
<th>10–25</th>
<th>25–50</th>
<th>50–75</th>
<th>75–90</th>
<th>90–97</th>
<th>97–100</th>
</tr>
</thead>
<tbody>
<tr>
<td>180</td>
<td>0.00</td>
<td>0.35</td>
<td>3.49</td>
<td>12.15</td>
<td>23.81</td>
<td>25.91</td>
<td>10.54</td>
<td>2.23</td>
</tr>
<tr>
<td>360</td>
<td>0.00</td>
<td>0.50</td>
<td>5.39</td>
<td>19.99</td>
<td>30.60</td>
<td>20.85</td>
<td>6.75</td>
<td>1.10</td>
</tr>
<tr>
<td>540</td>
<td>0.00</td>
<td>0.39</td>
<td>5.07</td>
<td>21.73</td>
<td>29.92</td>
<td>18.64</td>
<td>6.00</td>
<td>0.81</td>
</tr>
<tr>
<td>All</td>
<td>0.03</td>
<td>0.45</td>
<td>6.00</td>
<td>27.22</td>
<td>35.81</td>
<td>22.33</td>
<td>7.22</td>
<td>0.96</td>
</tr>
</tbody>
</table>
estimated variances of the individual horses ($\sigma^2$) for colt and filly body weights were quite similar to each other in this study (i.e. 1,431.14 vs. 1,431.01). For the application of this result to other horse populations, however, other studies will be necessary.

Slightly higher deviations of the data variation from the estimated percentile growth curves can be recognized (Figs. 1 and 2). In Tables 1, 2 and 3, the higher deviations are also clearly recognized. Our growth curves considering CG [16, 17] were based on accumulated knowledge of Japanese Thoroughbreds, and the assumptions of a mature weight of 575.0 kg and a mature withers height of 161.0 cm [6]. The higher deviations of body weight and withers height in the current Thoroughbred population suggest the possibility that the body size of recent colts and fillies are larger than those of former Japanese Thoroughbred populations on The Japanese Feeding Standard for Horses [6].

The potential problems of using the current underestimated percentile curves obtained in this study would relate to misunderstanding of the absolute superiority of the superior growing foals. In the recent Japanese Thoroughbred population, the 50–75% percentile intervals seem to be ordinary (see Tables 1, 2 and 3). Racehorse managers in the Hidaka region should recognize this information when using these underestimated percentile curves for the practical management of their horses. Even though the percentile curves proposed by this study are currently underestimated, the personal growing profile each foal can be precisely obtained and it can be used for the safe feeding management of a foal during CG periods. As more data of the mature weights and heights of the recent Japanese Thoroughbred population becomes available, more accurate growth curve models and variances will be estimated.

A simple way for the construction of the exact percentile growth curves is to compute the data summarizing statistics (e.g. ordered data percentages or medians) at all ages and to plot them on a graph visually, as noted. This approach gives us percentile graphs that are ‘data dependent’, and the obtained graphs tended to have sharp notches (i.e. not smooth lines). A variety of empirical methods have been used to develop smoothing curves using the actual data plots [4, 7]. Methods for smoothing include cubic splines, kernel regression, locally weighted regression and the lambda-mu-sigma (LMS) method [3]. The LMS method is often used for estimation of smoothed percentiles [4, 8, 22]. These methods are alternatives to deriving other empirical percentile growth curve equations for Thoroughbreds. As discussed by Wang and Chen [23], the use of Z-scores for the development of empirical percentile growth curves is still useful because of its ability to quantify the extreme data values based on parametric statistical estimations and hypothesis testing. Based on the assumption of normal distribution of the analyzed data, the estimated empirical percentile growth curves using Z-scores give us standardized scales for the analyzed data distributions.

Practical control of withers height would be a challenging task compared with the practical control of body weight. The estimated empirical percentile curves of withers height, however, will still be useful as a reference for the development of young horses when combined with the percentile information of body weight. The integrated percentile information of many growth traits gives us valuable references for the sound development of young Thoroughbred horses.

A useful property of our approach is the introduction of continuous single variable growth curve equations considering CG, which can be easily used to obtain empirical percentile curve equations, as shown in this study. The developed empirical percentile growth curves using Z-scores are computationally feasible and have value for understanding the aspect of body weight and withers height distributions and for managing young Thoroughbred horses, especially during compensatory growth periods.

Acknowledgments

We would like to thank the Racehorse Breeding Promotion Program (2005–2010) organized by the National Association of Racing, through which the Japan Bloodhorse Breeders’ Association was able to conduct its equine nutrition consulting study program as part of the Bloodstock Industry Business Management Training Program, and also the Japan Racing Association Hidaka Training and Research Center for providing growth data of horses as study material for this research. We also would like to thank Dr. Malcolm Fitz-Earle, Emeritus Professor of Capilano University, Canada, for revisions to this manuscript.

References


