Longitudinal study of the influence of the relative age effect (rae) on Czech tennis players aged 10–12

Roman Koloničný, Michal Bozděch, Jiří Zháněl

Faculty of Sports Studies, Masaryk University, Brno

Abstract
The issue of the phenomenon known as RAE first appeared in research within the educational system. In the field of sports research, it first emerged in the 1980s. The theory of RAE is based on the assumption that athletes born in the early months of the year are usually more advanced in both mental and physical fitness. Many studies have confirmed that birthdates of players in sports such as handball, ice hockey or football are distributed asymmetrically in the respective quarters of the year. A high incidence of birthdates has been shown to exist in the first half of the year. The aim of the study was to determine the impact of RAE on Czech tennis players in the age category of 10–12. The investigated sample consisted of players (n = 1 500) who ranked among the top one hundred in the official ranking of the Czech Tennis Association from 2003 to 2017. In order to verify the conformity of the distribution of the frequency of the players’ birthdates in the sample and the frequency of birthdates among the same population (n = 1 506 504) the Chi-Square test ($\chi^2$) was used. The results of the study have shown a significant influence of RAE in certain years in the reference age category of 10–12, and in thirteen of the fifteen years surveyed. The overall results in the reference period 2003–2017 have shown that most of the players were born in Q1 (n = 551) and Q2 (n = 442). Therefore, 2/3 of the players (66%) were born in the first half of the year, while only 1/3 (34%) were born in the second half. A longitudinal study of tennis players aged 10–12 showed the influence of RAE both in specific years (with the exception of 2011 and 2014) and in the reference period 2003–2017 taken as a whole. The results of our research correspond to the results of similar studies performed by other authors. Tennis coaches are advised to address the issue of RAE when working with tennis youth and when selecting players to the sports centers, representations, etc.

Keywords: birthdate, Chi-Square test, athletic talent, tennis.

ACKNOWLEDGMENTS
This publication was written at Masaryk University as part of the project “Diagnostics of the level of sport-specific motor preconditions in the context of age, somatic, gender and lateral asymmetries in sport” number MUNI/A/1087/2017 with the support of the Specific University Research Grant, as provided by the Ministry of Education, Youth and Sports of the Czech Republic in the year 2018.

INTRODUCTION
Already in the 1960s, the first research of the influence of birthdates on study results of primary school pupils (Green & Simmons, 1962; King, 1955) appeared. These studies showed that children born early in the calendar year did better at school than children born later. Researchers in the educational system have referred to a phenomenon which was termed „Relative Age Effect“ (RAE), or more precisely Birth Date Effect, Birth Quarter or Age Effect (Agricola et al., 2013).
The first research in the field of sport appeared in the 1980s and was focused – as in later years – mainly on collective sports such as ice hockey (Barnsley, Thompson & Barnsley, 1985; Barnsley & Thompson, 1988; Romann & Fuchslocher, 2009), football (Auguste & Lames, 2011; Barnsley, Thompson & Legault, 1992; Bäumler, 2000), baseball (Thompson, Barnsley & Stebelsky, 1991), basketball (Delorme & Raspau, 2009) and volleyball (Delorme, Boiche & Raspau, 2009). The results of those studies have shown that the influence of RAE was evident particularly in the junior categories, both the boys’ and girls’ ones. According to Filipčič (2001), this influence seems to be slightly stronger in boys’ categories. Other studies have shown that a larger effect is demonstrable in sports competitions at the level of national representations and higher national competitions. The said results imply, that in a given age category, individuals born in the first months of the year enjoy a certain advantage (mainly as regards physical and mental) compared to individuals born later because the age differences can reach up to twelve months. Sports successes achieved with the help of this fact might be mistaken for athletic talent. Such successful individuals may be selected to represent while real talents may be left unnoticed (Perič, 2010). In some cases, the influence of RAE may result in unwanted discrimination of players born later in a given year (Simon & Paull, 2001). According to the authors, Edgar and O’Donoghue (2005) a player who is about eight months older can be physically and mentally up to 5.3% more mature than the younger one.

An interesting thing are the findings of authors Albuquerque et al. (2012) that in combat sports no asymmetric distribution of the frequency of birthdates was found to exist since weight in addition to the age of the contestants is used for purposes of division into combat competition categories.

The research in tennis provides insight into an asymmetric distribution of birthdates of players and demonstrates that the most successful junior players were born in the first half of the year (Agricola, 2013; Agricola et al., 2013; Agricola et al., 2012; Ulbricht et al., 2015). Baxter – Jones (1995) found that 85% of elite British junior players were born in the first half of the year. Half of the investigated sample of Danish junior tennis players ranked at the top of the ranking list were born in the first three months of the year (Dudink, 1994). In a study on the division of male and female tennis players into tennis sports centers for youth in the Czech Republic, the authors reported that in 2008 there had been seven Czech female tennis players who were ranked among the first hundred on the WTA ranking list. All these tennis players were born in the first quarter of the calendar year, and all of them were selected to the centers as juniors (Suchý & Pecha, 2009).

It follows from the above synthesis that the impact of RAE in sport is significant and attention should be paid to it since RAE may influence the sport performance. In the hunt for, identification and selection of sport talents, or when making selections to junior representations – not only the calendar age but also the biological one should be taken into account. In youth sport individuals are divided into age categories in order to avoid unequal sports competitions and to take account of the differences in physical and mental maturity of these individuals. Nevertheless, even within these age categories, big age differences among individuals may arise. Individuals born early in the defined period have a big advantage over their team – or school – mates who were born towards the end of the period. Differences come to light both in anthropometric parameters (height, weight, muscular growth, etc.), as well as in the physical development of an individual. There are many research studies which point out to the facts termed „Relative Age Effect.“ We focused our research on the longitudinal monitoring of the influence of RAE on Czech tennis players aged 10–12.
**METHODOLOGY**

For the purposes of the research, groups of Czech junior tennis players (n = 1500) who ranked up to the 100th place in the official ranking list of the Czech Tennis Association (CTA) were chosen on purpose. They were from the younger pupils’ category (aged 10–12). The research data (birthdate and ranking) were obtained from the official website of the CTA. Of course, the legal requirement of personal data protection was complied with. The players’ names in connection with their birthdates were not used in the evaluation. The so-called „cut-off date“ (the date that is decisive for their inclusion in age categories) is always the first of January according to applicable rules of the CTA. The birthdates of those included in the research were analysed as follows: Q1 (1. 1.–31. 3.), Q2 (1. 4.–30. 6.), Q3 (1. 7.–30. 9.), Q4 (1. 10.–31. 12.), and S1 (1. 1.–30. 6.) and S2 (1. 7.–31. 12.) respectively.

The research data were analyzed using the methods of descriptive (relative and absolute frequency) and inferential statistics (Chi-Square, a test of good conformity, $\chi^2$). The Chi-Square test was used for the assessment of conformity of the distribution of expected and acquired frequency (Hendl, 2012). The acquired frequency, well then the frequency of the same age population in the respective quarters were set by the data of the Czech Statistical Office for years 2003–2016 as follows: Q1 = 24%, Q2 = 26%, Q3 = 27%, Q4 = 23%.

In light of above – mentioned research problems the following research questions were formulated:
1. Is it possible to prove the influence of RAE in specific years of the period 2003–2017?
2. Is it possible to prove the influence of RAE throughout the period 2003–2017?

**RESULTS**

In table 1 there is an absolute (and also relative) frequency of birthdates by quarters in specific years, and according to the value of $\chi^2$, the influence of RAE was assessed.

**Tab. 1:** Frequency of birthdates by quarters in specific years

<table>
<thead>
<tr>
<th>Year</th>
<th>Q1 (%)</th>
<th>Q2 (%)</th>
<th>Q3 (%)</th>
<th>Q4 (%)</th>
<th>$\chi^2$</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003</td>
<td>40</td>
<td>28</td>
<td>17</td>
<td>15</td>
<td>17,307</td>
<td>&lt;0,001</td>
</tr>
<tr>
<td>2004</td>
<td>32</td>
<td>38</td>
<td>20</td>
<td>10</td>
<td>17,368</td>
<td>&lt;0,001</td>
</tr>
<tr>
<td>2005</td>
<td>39</td>
<td>28</td>
<td>21</td>
<td>12</td>
<td>16,123</td>
<td>&lt;0,001</td>
</tr>
<tr>
<td>2006</td>
<td>46</td>
<td>27</td>
<td>12</td>
<td>15</td>
<td>31,321</td>
<td>&lt;0,001</td>
</tr>
<tr>
<td>2007</td>
<td>40</td>
<td>29</td>
<td>17</td>
<td>14</td>
<td>18,238</td>
<td>&lt;0,001</td>
</tr>
<tr>
<td>2008</td>
<td>37</td>
<td>27</td>
<td>21</td>
<td>15</td>
<td>11,196</td>
<td>0,011</td>
</tr>
<tr>
<td>2009</td>
<td>38</td>
<td>30</td>
<td>20</td>
<td>12</td>
<td>15,858</td>
<td>0,001</td>
</tr>
<tr>
<td>2010</td>
<td>37</td>
<td>27</td>
<td>21</td>
<td>15</td>
<td>11,196</td>
<td>0,011</td>
</tr>
<tr>
<td>2011</td>
<td>34</td>
<td>24</td>
<td>22</td>
<td>20</td>
<td>5,638</td>
<td>0,131</td>
</tr>
<tr>
<td>2012</td>
<td>39</td>
<td>25</td>
<td>24</td>
<td>12</td>
<td>15,008</td>
<td>0,002</td>
</tr>
<tr>
<td>2013</td>
<td>37</td>
<td>30</td>
<td>17</td>
<td>16</td>
<td>13,491</td>
<td>0,004</td>
</tr>
<tr>
<td>2014</td>
<td>28</td>
<td>31</td>
<td>25</td>
<td>16</td>
<td>3,907</td>
<td>0,272</td>
</tr>
<tr>
<td>2015</td>
<td>31</td>
<td>37</td>
<td>15</td>
<td>17</td>
<td>13,594</td>
<td>0,004</td>
</tr>
<tr>
<td>2016</td>
<td>33</td>
<td>36</td>
<td>20</td>
<td>11</td>
<td>15,297</td>
<td>0,002</td>
</tr>
<tr>
<td>2017</td>
<td>40</td>
<td>25</td>
<td>25</td>
<td>10</td>
<td>18,201</td>
<td>&lt;0,001</td>
</tr>
</tbody>
</table>
In the search for critical values by quarters, 3 degrees of freedom were used \((k - 1)\), for 5\% level of significance which was chosen, the critical value is \(\chi^2 = 7,8153\). The results show that in Q1 of each monitored year (with the exception of 2004, 2014, 2015, 2016) most players were born. The maximum value \(\chi^2\) was found in 2006; there were the biggest differences between observed and expected frequencies of birthdates within specific quarters \((\chi^2 = 31,321, p<0,001)\). The minimum value \(\chi^2\) (less than the critical value) was found in 2014 \((\chi^2 = 3,907, p=0,272)\); the differences between observed and expected frequencies of birthdates within specific quarters were the least (unimportant). On the basis of the calculated values \(\chi^2\) we can conclude that almost within all monitored years (with the exception of 2011 and 2014), significant differences between observed and expected frequencies (table 1) were found because the values \(\chi^2\) exceed the critical value \((\chi^2 = 7,8153)\). The impact of RAE may be considered proven on the basis of these results (with the exception 2011 and 2014).

**Fig. 1:** Graphical representation of the frequency of birthdates in specific years

Figure 1 shows graphically (bar graph) the division of frequencies of birthdates in significant years; it specifies the information which is included in table 1 and documents significantly declining trends in specific quarters from Q1 to Q4.

**Tab. 2:** The absolute frequency of birthdates within the periods 2003–2017 by quarters

<table>
<thead>
<tr>
<th>Quarters ((n = 1500))</th>
<th>Q₁ (%)</th>
<th>Q₂ (%)</th>
<th>Q₃ (%)</th>
<th>Q₄ (%)</th>
<th>(\chi^2)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>551</td>
<td>442</td>
<td>297</td>
<td>210</td>
<td>189,896</td>
<td>&lt;0,001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Semester ((n = 1500))</th>
<th>S₁ (%)</th>
<th>S₂ (%)</th>
<th>(\chi^2)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>993</td>
<td>507</td>
<td>157,464</td>
<td>&lt;0,001</td>
</tr>
</tbody>
</table>
Table 2 includes the total and relative frequencies of birthdates within the period of 2003–2017 by quarters, by half-years and the value $\chi^2$. According to the division of frequencies of birthdates, it is possible to examine – analogously to previous findings – significant differences between observed and expected frequencies of birthdates within specific quarters ($\chi^2 = 157,464$, $p < 0.001$). Total frequencies of data by half-years show that in $S_1$ (66%) the frequency of birthdates is almost 2/3 bigger than in $S_2$ (34%). Significant differences between $S_1$ and $S_2$ were found here as well ($\chi^2 = 157,464$, $p < 0.001$). These findings confirm the impact of RAE.

DISCUSSION

A significant influence of RAE was found in almost all years of the monitored period of 2003–2017 except 2011 and 2014. With the exception of 2004, 2014, 2015 and 2016, the values of the frequency of data were of maximum in the first quarter of the year. The half – year’s comparison within the whole period of 2003–2017 shows, that 2/3 of players ranked in the first hundred were born in the first half-year. The comparison of our results to the results of other authors shows a similar nature of the division of frequencies in specific quarters. Filipič (2001) presented a very similar division of frequencies of birthdates in tennis players aged 12 as in our case ($Q1 = 30\%$, $Q2 = 20\%$, $Q3 = 18\%$, $Q4 = 12\%$), Ulbricht et al. (2005) found the influence of RAE on tennis players aged 12 ($Q1 = 38\%$, $Q2 = 22\%$, $Q3 = 20\%$, $Q4 = 9\%$) as well and he pointed at a certain fall of this influence on tennis players aged 18 ($Q1 = 21\%$, $Q2 = 20\%$, $Q3 = 13\%$, $Q4 = 8\%$), therefore the weakening of the influence of RAE is influenced by increasing age. Similar results of research on World Junior Tennis Finals (WJTF) in 2007–2011 were published by Agricola et al. (2012) who were finding out the impact of RAE on the best world tennis players aged 13–14 ($Q1 = 40.8\%$, $Q2 = 26.3\%$, $Q3 = 22.1\%$, $Q4 = 10.8\%$). On the basis of this research, we can conclude that the results of our studies correspond to the research performed by other authors who confirm the existence of the influence of RAE on junior tennis players. Junior tennis players born in Q1, resp. in Q2 may be more advanced in both mental and physical fitness. These facts should be respected when identifying sports talents.
CONCLUSIONS

The longitudinal study has shown the influence of RAE on Czech tennis players aged 10–12 (n = 1500). The Chi-Square test results ($\chi^2$) have shown the absence of good harmony between observed and expected division of frequencies of birthdates almost in all periods of 2003–2017 with the exception of 2010 and 2014. The impact of RAE was proven within the whole monitored period both in specific quarters and a half - years as well. Our found results correspond to results performed by other authors. The issue of the influence of RAE should be respected by trainers when working with tennis youth, e.g., selecting players to sports centers, representation teams and when identifying sports talents.

There are several ways to prevent the influence of RAE in tennis:
- age categories adjustment (e.g., half-year age category);
- quotas establishing (e.g., 25 % players born in a quarter);
- to sort teams according to anthropometric parameters (e.g., height and weight);
- rotation of cut-off dates.

However, it is questionable whether the umbrella tennis organizations (CTA, ITF, TE) will be favorably inclined to these changes.”

References

Longitudinal study of the influence of the relative age effect (rae) on Czech tennis players aged 10–12


Corresponding author:
Roman Koloničný,
E-mail: roma.kolo@gmail.com
Tel: 602 152 000,
Address: Dřevčice 147, 250 01, CZ