Relative age effect among French swimmers

Audrey Difernand, Quentin de Larochelambert, Robin Pla, Juliana Antero, Jean-François Toussaint, Adrien Sedeaud

RELATIVE AGE EFFECT AMONG FRENCH SWIMMERS

Audrey Difernand[1], Quentin De Larochelambert[1], Robin Pla[1], Juliana Antero[1], Jean-François Toussaint[1] & Adrien Sedeaud[1]

[1] *Institut de Recherche bio-Médicale et d'Epidémiologie du Sport (IRMES), Institut National du Sport, de l'Expertise et de la Performance (INSEP), 11 avenue du Tremblay, 75012 Paris*
audrey.difernand@insep.fr; qdelarochelambert@ffs.fr; robin.pla@ffnatation.fr; juliana.antero@insep.fr; jean-francois.toussaint@aphp.fr; adrien.sedeaud@insep.fr

Résumé. L'objectif de cette étude était de révéler la présence de l'effet de l'âge relatif parmi les nageurs Français et de proposer une méthode de rééquilibrage afin de mieux apprécier le potentiel de l'athlète en fonction de sa catégorie et de sa discipline. 62 610 nageurs entre 10 et 16 ans sur la discipline du 100m nage libre en bassin de 50m sont considérés pour cette étude. Parmi eux, moins d'un nageur sur cinq entre 13 et 16 ans est né dans le dernier trimestre de l'année. Pour éviter l'abandon ou la perte de vue des nageurs, nous avons mis en place une méthode de rééquilibrage basée sur la performance du nageur, son âge exact au moment de la compétition et le coefficient de régression entre la performance et l'âge au sein de la catégorie considérée. Après application de la méthode, nous avons remarqué aucune différence significative entre les performances rééquilibrées et les performances réalisées a posteriori excepté pour la catégorie des 13 ans.

Mots-clés. Age relatif, natation, détection, méthodes de rééquilibrage

Abstract. The aim of this study was to highlight the presence of the relative age effect among French swimmers and to propose a method of rebalancing in order to better appreciate the potential of each athlete according to his category and discipline. 62 610 males' swimmers between the ages of 10 and 16 in the 100m freestyle in a 50m pool are considered for this study. Less than one in five swimmers aged 13-16 was born in the last quarter of the year. We implemented a rebalancing method based on the swimmer's performance, his exact age at the time of the competition and the regression coefficient between performance and age within the category considered. After applying the method, we found no significant differences between the rebalanced performances and the performances achieved except for the 13 years old category.

Keywords. Relative age, swimming, talent identification, rebalancing method

1. Introduction

The issue of talent identification in sport is very complex. Indeed, the objective is to identify, detect and choose the future Olympic champions 10 years in advance. Age is one of the key factors in this Olympic quest. The younger the athlete is identified, the more time he or she will have to build up in high performance structures. Relative age is defined by the exact age as opposed to the integer age which does not consider the day and month of birth (Helsen, Starkes, et Van Winckel 1998). Thus, within the same age category, there can be up to 365 days difference between two athletes. This age difference can also be translated into differences in height, weight, strength, experience and agility. For this reason, older athletes within the same age group may be favored in selections compared to younger ones: this is called relative age effect (RAE) (Musch et Grondin 2001). A study highlights the over-representation of British rugby players with 73% born in the first half of the year compared to 27% in the second half (McCarthy et Collins 2014). The RAE is also present in individual sports such as swimming (Abbott et al. 2020). After having shown the presence of RAE, they aim to find a method to remove this effect (Cobley et al. 2019). In our study, the objective was not only to show that a relative age effect exists among French swimmers using the example of male 100m freestyle, but also to test a method that allows us to reveal the potential of a swimmer with respect to his age, performance and age category.

2. Methods

All performances from local to international level were collected by the French Swimming Federation between 2000 and 2020. In this study, we were interested in males' 100m freestyle event in the Olympic pool. The database was composed of the date of birth, the date of performance and the performance in seconds of each swimmer.
In the first part, we kept only the season best performance for each swimmer. To better visualize the differences between swimmers, we defined the relative age gap. It was the time between the last birthday and the competition date. Swimmers were divided into birth quarters (Q1: January, February, March, Q2: April, May, June, Q3: July, August, September, Q4: October, November, December). To test for a significant difference between the birth quarters, chi-square test, odds ratios and 95% confidence intervals were calculated.
Within each age category, we related the relative age gap to performance in order to reveal the nature of the relationship between these two variables. Then, using the regression coefficient, we calculated the rebalanced performance:

$$T_r = T_0 - c * m$$

where:
T_r: rebalanced performance time
T_0: initial performance time
c: slope coefficient
m: time until the next birthday

In the second part, in order to validate our rebalancing method, we used a database composed of swimmers having achieved 2 performances in the season with a minimum of 6 months difference. The rebalancing method is applied to the first performance considering the difference with the second performance.

$$T_r = T_1 - c * (T_2 - T_1)$$

where:
T_r: rebalanced performance time
T_1: initial performance time
T_2: second performance time
c: slope coefficient

Thus, it has been possible to compare the rebalanced performance with the (second) performance achieved.

3. Results

The dataset consisted of 62 610 swimmers between the ages of 10 and 16 who have competed at least once in the last 20 years. The distribution in birth quarters is presented in table 1. At 15, for example, we were twice as likely to find a swimmer born in the first quarter as in the last quarter (OR = 2.13, $p<0.05$). Indeed, 32.2% of them were born in the first quarter compared to only 18.2% in the last quarter. Also, we noted that among the top 10% swimmers, the proportion born in the first quarter increases with age in contrast to the proportion born in the last quarter. Among the 14 years old, 19.1% of swimmers were born in the last quarter for all swimmers combined, compared to 16.1% in the top 10%.

Age	Top %	Total	Q1 (%)	Q2 (%)	Q3 (%)	Q4 (%)	χ^2	p-value	OR	p	95% CI
10	100	3767	897 (23.81)	871 (23.12)	1087 (28.86)	912 (24.21)	30.78	***	0.98	**	[0.97 ; 0.99]
11	100	6982	1706 (24.43)	1738 (24.89)	1941 (27.8)	1597 (22.87)	35.46	***	1.09	**	[1.06 ; 1.12]
12	100	9807	2661 (27.13)	2498 (25.47)	2554 (26.04)	2094 (21.35)	75.20	***	1.37	**	[1.28 ; 1.47]
13	100	12787	3655 (28.58)	3426 (26.79)	3152 (24.65)	2554 (19.97)	211.99	***	1.60	**	[1.5 ; 1.7]
14	100	15140	4559 (30.11)	4147 (27.39)	3540 (23.38)	2894 (19.11)	418.50	***	1.82	**	[1.72 ; 1.93]
15	100	14127	4547 (32.19)	3976 (28.14)	3034 (21.48)	2570 (18.19)	679.78	***	2.13	**	[2.01 ; 2.25]
10	10	376	48 (12.77)	92 (24.47)	134 (35.64)	102 (27.13)	40.26	***	0.39	**	[0.29 ; 0.52]
11	10	698	91 (13.04)	179 (25.64)	267 (38.25)	161 (23.07)	90.15	***	0.50	**	[0.43 ; 0.58]
12	10	980	156 (15.92)	262 (26.73)	365 (37.24)	197 (20.1)	101.69	***	0.75	**	[0.71 ; 0.79]
13	10	1278	281 (21.99)	386 (30.2)	381 (29.81)	230 (18.0)	55.39	***	1.28	**	[1.2 ; 1.37]
14	10	1514	428 (28.27)	450 (29.72)	392 (25.89)	244 (16.12)	68.26	***	2.05	**	[1.73 ; 2.43]
15	10	1412	426 (30.17)	416 (29.46)	342 (24.22)	228 (16.15)	70.95	***	2.24	**	[1.86 ; 2.7]

Table 1: *Distribution by birth quarters with respect to age groups and level of competitiveness among male's 100m Freestyle in 50m pool (top %: top percentage, Total: number of observations, Q1-4: birth quarters 1 to 4, χ^2: chi-square adjustment value, ***: $p<0.01$, **: $p<0.05$, OR: odds-ratios between Q1 and Q4, 95% CI: 95% Confidence Interval.)*

Linear regressions were run in each age category (*Figure 1*). The relationship between the relative age gap and the average performance was negative and the absolute value of the coefficient decreased with age. At 12, on average, one month more of relative age gap gained 0.826 seconds, which over a year gave us more than 9 seconds.

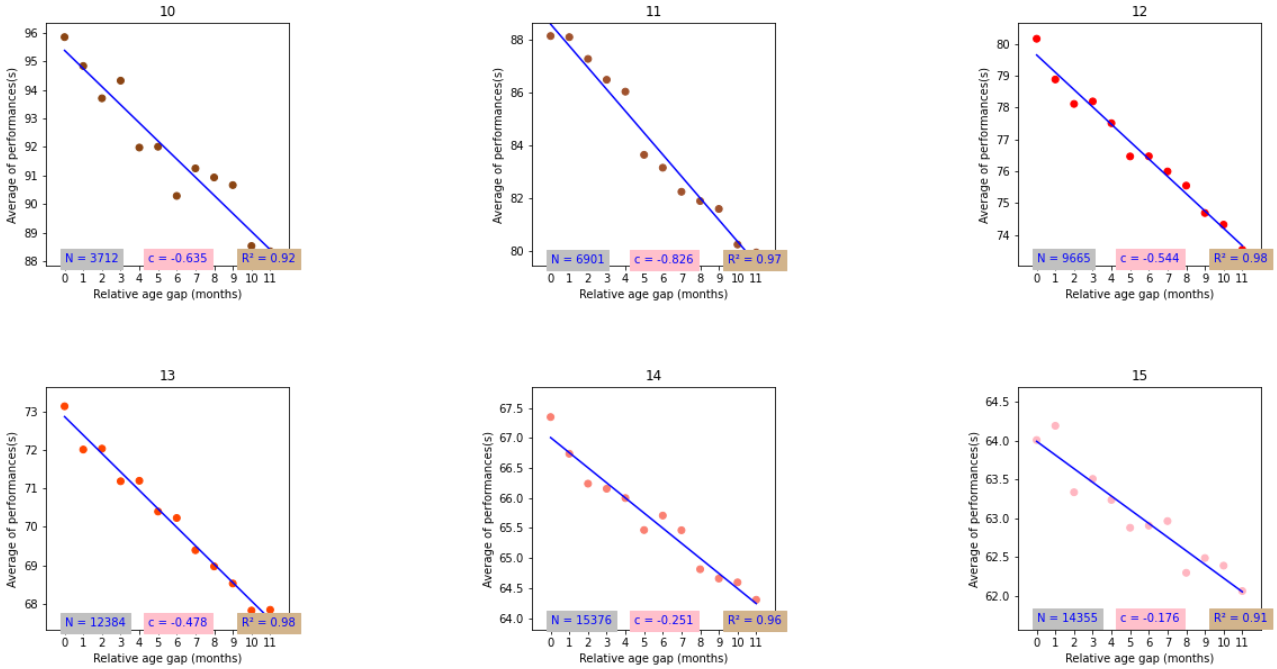

***Figure 1**: Average of performances (in seconds) by relative age gap (in months) among male's 100m Freestyle in 50m pool (N: number of observations, c: slope regression coefficient, R²: coefficient of determination)*

The rebalancing method is illustrated in *figure 2*. We observed that the best performances before rebalancing were mostly achieved by relatively old swimmers within the age category considered. After rebalancing, the top 10 best performances was renewed and new potentials emerged. Indeed, 132, 118, 187, 65, 63 and 42 swimmers had rebalanced performances that were better than the threshold of the pre-rebalancing top 10 best performances. The rate of stability on the other hand tended to be higher with age.

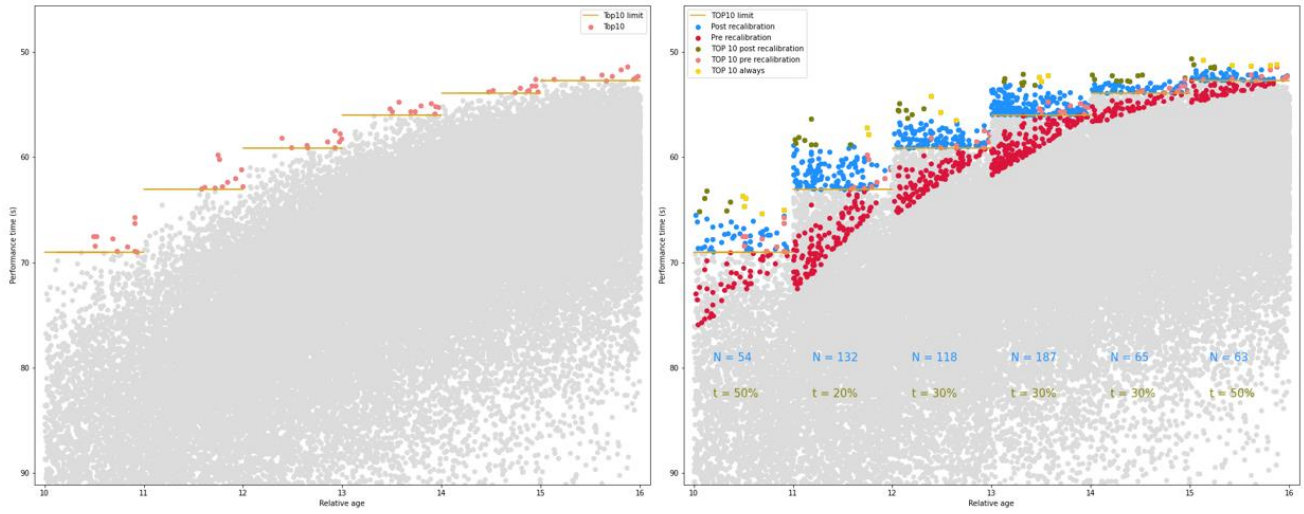

***Figure 2**: On the left, pre-rebalanced performances between 10 and 16 years old among male's 100m Freestyle in 50m pool. Horizontal bars (gold) mark the top10 best performances (pink) threshold. On the right, post-rebalanced performances are illustrated. Performances from the old and new top 10 best performances are shown in gold, those from the new top 10 best performances after rebalancing are in green and those from the old top 10 best performances before rebalancing are still in pink. Performances above the old top10 threshold are shown in blue and the red dots represent their original performances before rebalancing.*

Since the sample does not follow the normal distribution hypothesis, a non-parametric Friedman test is performed to detect a significant difference between the three types of performance. We assume the null hypothesis that the three types of performance follow the same distribution. We have for each age category a p-value smaller than 0.01, which allows us to reject the null hypothesis. Therefore, a Wilcoxon test is performed for each pair of possible performances to reveal which performance type is significantly different from the other two. *Figure 3* shows the comparisons between the original, second and rebalanced performances. Indeed, we found no significant differences on average between the actual second performances and the rebalanced performances except for the 12- and 15-years old category ($p<0.05$). Indeed, for 10, 11, 13 and 14 years-old categories, the average of rebalanced performances was 82.46, 77.25, 65.01 and 61.79 seconds and the average of realized second performances was 84.09, 77.51, 64.01 and 61.86 seconds respectively.

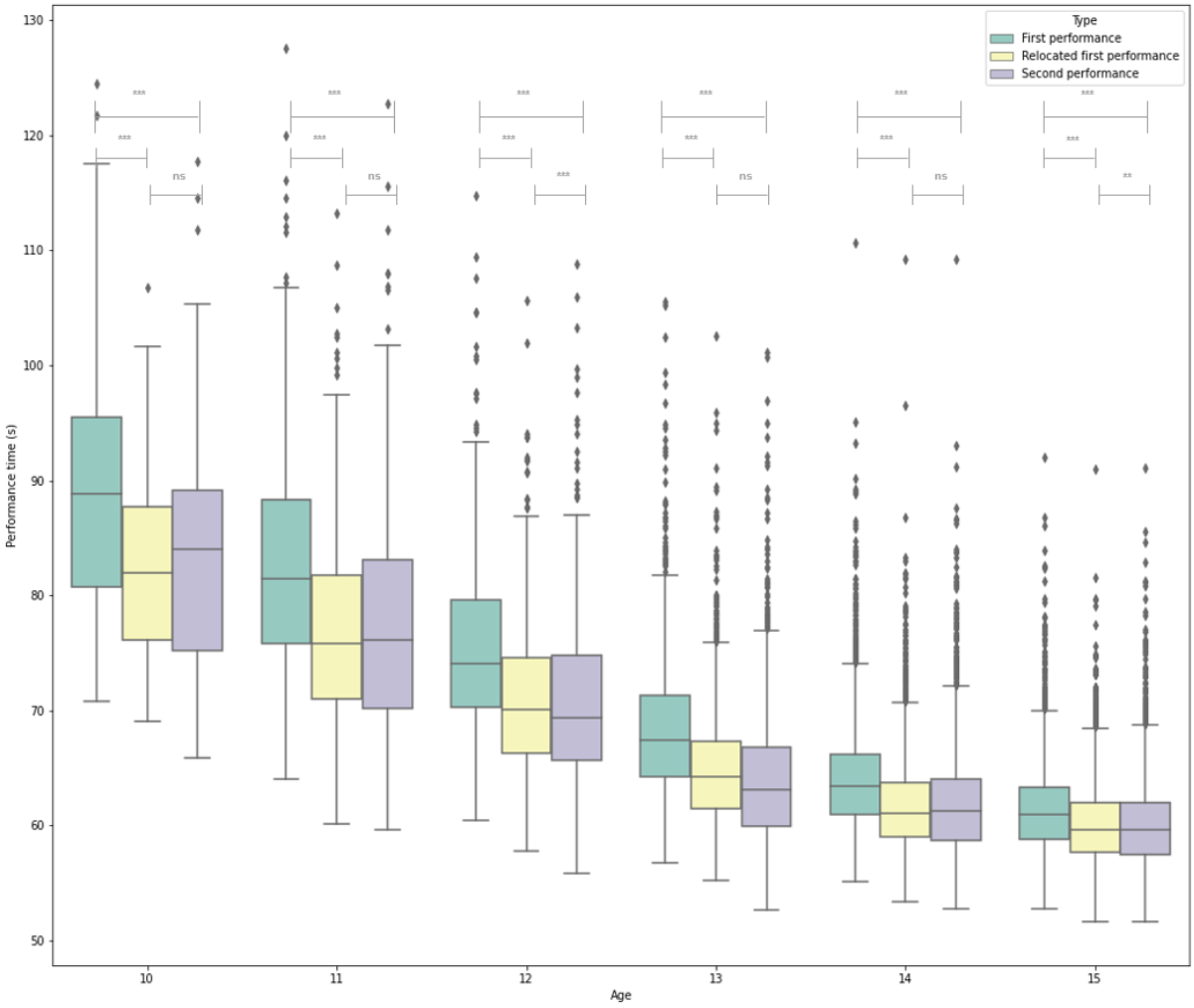

***Figure 3**: Boxplots of first, rebalanced and second performances by age between 10 and 15 years old among male's 100m Freestyle in 50m pool.*

4. Conclusion

This study is the first to demonstrate the relative age effect among French swimmers, particularly among males' 100m Freestyle in 50m pools. It is worth noticing that this effect is more important as the level of competitiveness increases. Also, thanks to the rebalancing method, it is possible to objectify swimmers' performances in relation to their exact age and thus better appreciate their potential. As performance is multifactorial and complex to predict, further studies should be considered, including other variables as maturity for example.

Bibliographie

Abbott, Shaun et al. 2020. « Testing the Application of Corrective Adjustment Procedures for Removal of Relative Age Effects in Female Youth Swimming ». *Journal of Sports Sciences* 38(10): 1077-84.

Cobley, Stephen et al. 2019. « Removing Relative Age Effects from Youth Swimming: The Development and Testing of Corrective Adjustment Procedures ». *Journal of Science and Medicine in Sport* 22(6): 735-40.

Helsen, Werner F., Janet L. Starkes, et Jan Van Winckel. 1998. « The Influence of Relative Age on Success and Dropout in Male Soccer Players ». *American Journal of Human Biology: The Official Journal of the Human Biology Council* 10(6): 791-98.

McCarthy, Neil, et Dave Collins. 2014. « Initial Identification & Selection Bias versus the Eventual Confirmation of Talent: Evidence for the Benefits of a Rocky Road? » *Journal of Sports Sciences* 32(17): 1604-10.

Musch, Jochen, et Simon Grondin. 2001. « Unequal Competition as an Impediment to Personal Development: A Review of the Relative Age Effect in Sport ». *Developmental Review* 21(2): 147-67.