



OPEN

Comparing cross-sectional and longitudinal tracking to establish percentile data and assess performance progression in swimmers

Dennis-Peter Born^{1,2✉}, Eva Rüeger², C. Martyn Beaven³ & Michael Romann²

To provide percentile curves for short-course swimming events, including 5 swimming strokes, 6 race distances, and both sexes, as well as to compare differences in race times between cross-sectional analysis and longitudinal tracking, a total of 31,645,621 race times of male and female swimmers were analyzed. Two percentile datasets were established from individual swimmers' annual best times and a two-way analysis of variance (ANOVA) was used to determine differences between cross-sectional analysis and longitudinal tracking. A software-based percentile calculator was provided to extract the exact percentile for a given race time. Longitudinal tracking reduced the number of annual best times that were included in the percentiles by 98.35% to 262,071 and showed faster mean race times ($P < 0.05$) compared to the cross-sectional analysis. This difference was found in the lower percentiles (1st to 20th) across all age categories ($P < 0.05$); however, in the upper percentiles (80th to 99th), longitudinal tracking showed faster race times during early and late junior age only ($P < 0.05$), after which race times approximated cross-sectional tracking. The percentile calculator provides quick and easy data access to facilitate practical application of percentiles in training or competition. Longitudinal tracking that accounts for drop-out may predict performance progression towards elite age, particularly for high-performance swimmers.

Abbreviations

BU	Butterfly
BA	Backstroke
BR	Breaststroke
FR	Freestyle
IM	Individual medley

Previous studies have suggested longitudinal tracking of race times to assess talent development and performance trajectories¹. Longitudinal tracking involves the retrospective analysis of performance of successful swimmers from age of peak performance (21–26 years of age) back to their adolescent performance². As such, longitudinal tracking accounts for early drop-outs when predicting elite age success³. This is of particular importance, as success at junior age is only a poor predictor for success at elite age^{4,5}. Of the successful swimmers under the age of 15 years, less than one third, i.e. 18 out of 60, were reselected for the senior team (aged ≥ 19 years)⁴. Furthermore, transition rate from junior to senior success was even as low as 10% for the top 10 male swimmers of the age category⁵. Although, transition rate improves with age (14 vs. 17 years of age)⁵, swimmers are typically selected at an early age^{1,6}. Therefore, deselected talents are irreversibly lost at a timepoint with low transition rate. To account for these early drop-outs, longitudinal tracking appears to be the best choice when establishing reference values for talent development.

On the other hand, such longitudinal reference values may be biased by flaws in the current talent identification programs. As such, the relative age effect favors swimmers born early in the year¹. At 13 years of age,

¹Swiss Swimming Federation, Section for High-Performance Sports, Bern, Switzerland. ²Department for Elite Sport, Swiss Federal Institute of Sport Magglingen, Hauptstrasse 247, 2532 Magglingen, Switzerland. ³University of Waikato, Te Huataki Waiora School of Health, Tauranga, New Zealand. ✉email: dennis.born@swiss-aquatics.ch

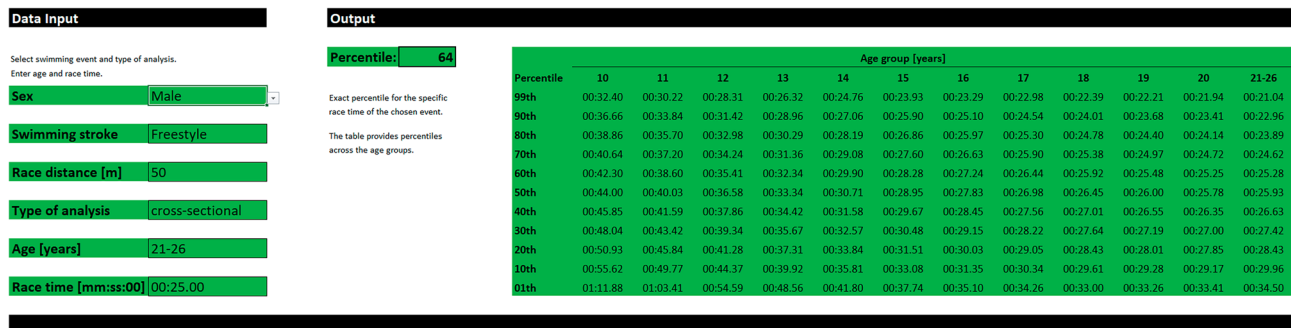


Figure 1. Screenshot of the percentile calculator that displays percentile data for a given race time of a particular swimming event. The software-based tool can be retrieved from the Supplementary Material.

56% of Australian's top national age-group swimmers were born in the first quarter of the year and only 4% in the last¹. Additionally, at age of peak height velocity, i.e. 13.8 and 12.0 years for males and females, respectively, large inter-individual variation in maturation results in a biological age difference of up to 5 years^{7,8}. Thus, early maturing swimmers are promoted for talent development programs, and potentially highly talented but late maturing swimmers can be deselected and irreversibly lost from the talent program⁹. The dominance of early maturing swimmers and those born early in the year may result in faster race times during adolescence¹⁰, especially in simultaneous strokes and sprint events¹¹. Thus, expectations for talented, but normal or late maturing age-group swimmers would be overestimated. To reduce bias of current talent identification programs, previous studies analyzed race times using a cross-sectional approach and simply included all available data related to the research question¹².

A recent study has established reference values based on percentile curves for Olympic swimming long-course events (50 m pool length)³. However, during the winter season in the northern hemisphere, races are held as short-course events. Competing in the same swimming strokes and over the same race distances, number of turns are increased due to the 25 m pool length. As such, at the beginning of each lap, swimmers achieve velocities far beyond the actual (clean) swimming velocity due to the push-off from the pool wall during each turn¹³. The different pacing pattern and velocity distributions improve short-course performance by $2.0 \pm 0.6\%$ for freestyle (FR) and $4.3 \pm 3.2\%$ for individual medley (IM) compared to the long-course events^{14,15}. Therefore, specific reference values for short-course races are required, as they cannot be compared to long-course races.

Previous studies have analyzed performance development for a specific swimming stroke or race distance^{5,16–18}. However, coaches and swimmers require reference values for both sexes, i.e. males and females, over all swimming strokes, i.e. butterfly (BU)–backstroke (BA)–breaststroke (BR)–FR–IM and all race distances, i.e. 50–100–200–400–800–1500 m. As such, race times across each age group (junior to elite age) differ for each of the aforementioned 34 swimming events and require specific reference values.

Modern technology enables us to gather the required data for both longitudinal and cross-sectional analyses¹⁹. Specific percentile curves can be established for each swimming event, using the multiple million race results of the database of the European swimming federation (LEN)²⁰. These percentiles provide a relative measure of race times enabling comparison between various race distances and swimming strokes²¹ across a wide range of performance levels and age groups²². However, to utilize the data in practice, coaches and other users need a software solution to avoid a time-consuming search of the required percentile values from multiple tables and charts.

Thus, the aims of the present study were to (1) establish reference values for competitive swimmers based on percentiles for swimming short-course (25 m pool length) races for both sexes over all swimming strokes and all race distances; (2) compare percentiles established by cross-sectional and longitudinal tracking; and (3) provide a software-based percentile calculator to enhance practical utility of the reference values for talent development.

Results

Percentiles for the cross-sectional dataset were based on 15,928,723 annual best times. Longitudinal tracking reduced number of annual best times by 98.35% to 262,071 annual best times. Percentile data, including tables and figures for all 68 swimming events for both sexes, all swimming strokes, and all race distances, can be retrieved from the percentile calculator and tables in the Supplementary Material. Figure 1 illustrates the percentile calculator showing the exact percentile for a specific race time of the chosen swimming event.

Cross-sectional analysis resulted in slower mean race times compared to longitudinal tracking. This difference was significant in the early junior and late junior age category for upper (80–99th) percentiles ($P < 0.05$), in the early junior, late junior, and sub-elite age categories for medium (40–59th) percentiles ($P < 0.05$), and in all age categories for lower (1st–20th) percentiles ($P < 0.05$) for both male (Table 1) and female swimmers (Table 2). In summary, the lower the performance level, the more higher age categories were affected by the difference between cross-sectional analysis and longitudinal tracking.

Regarding cross-sectional analysis, race times significantly ($P < 0.05$) improved up to elite age for the high (80–99th) percentiles, compared to a plateauing in race times at sub-elite age when using longitudinal tracking. However, with both cross-sectional analysis and longitudinal tracking for male and female swimmers, race times of the low (1st–20th) percentiles became significantly faster ($P < 0.05$) up to sub-elite age but significantly slower again at elite age.

	Early junior 10–14 years	Late junior 15–17 years	Sub-elite 18–20 years	Elite > 21 years		F-value	P-value	$p\eta^2$
80–99th percentiles								
Cross-sectional	02:34.31 ± 12.73	02:12.20 ± 10.36 [#]	02:03.13 ± 09.02 [#]	01:59.63 ± 08.86 [#]	(a)	$F_{(1 198)} = 10$	$P = 0.002$	0.05
Longitudinal	02:27.70 ± 11.33 [*]	02:06.04 ± 09.15 ^{##}	01:59.32 ± 08.24 [#]	01:58.79 ± 08.24	(b)	$F_{(1 220)} = 13,810$	$P < 0.001$	0.99
					(c)	$F_{(1 220)} = 114$	$P < 0.001$	0.37
40–59th percentiles								
Cross-sectional	02:57.51 ± 12.99	02:28.91 ± 11.62 [#]	02:18.07 ± 10.32 [#]	02:16.83 ± 09.91 [#]	(a)	$F_{(1 198)} = 25$	$P < 0.001$	0.11
Longitudinal	02:46.43 ± 11.88 [*]	02:19.63 ± 09.98 ^{##}	02:11.70 ± 09.22 ^{##}	02:13.36 ± 09.37 [#]	(b)	$F_{(1 220)} = 19,143$	$P < 0.001$	0.99
					(c)	$F_{(1 220)} = 174$	$P < 0.001$	0.47
1st–20th percentiles								
Cross-sectional	03:30.68 ± 18.08	02:54.75 ± 16.89 [#]	02:42.92 ± 16.87 [#]	02:46.26 ± 17.88 [#]	(a)	$F_{(1 198)} = 24$	$P < 0.001$	0.11
Longitudinal	03:14.35 ± 16.36 [*]	02:42.34 ± 15.12 ^{##}	02:34.51 ± 15.82 ^{##}	02:37.59 ± 16.00 ^{##}	(b)	$F_{(1 303)} = 10,570$	$P < 0.001$	0.98
					(c)	$F_{(1 303)} = 90$	$P < 0.001$	0.31

Table 1. Comparison of percentiles based on cross-sectional vs. longitudinal analysis of *male* swimmers. Analysis of variance (ANOVA) with repeated measure and one between subject factor was used to compare upper (80–99th), medium (40–59th), and lower percentiles (1st–20th) based on mean [mm:ss.00] ± standard deviation [ss.00] across the 200 m events of all swimming strokes. (a) Main effect: type of analysis (cross-sectional vs. longitudinal). (b) Main effect: age category (early junior–late junior–sub-elite–elite). (c) Interaction effect: type of analysis × age group. *Post-hoc* comparison. ^{*}Significant difference to cross-sectional analysis. [#]Significant difference to previous age category.

	Age categories					F-value	P-value	$p\eta^2$
	Early junior 10–14 years	Late junior 15–17 years	Sub-elite 18–20 years	Elite > 21 years				
80–99th percentiles								
Cross-sectional	02:38.57 ± 13.01	02:23.78 ± 11.50 [#]	02:17.07 ± 10.66 [#]	02:13.84 ± 10.69 [#]	(a)	$F_{(1 198)} = 11$	$P < 0.001$	0.06
Longitudinal	02:29.53 ± 11.70 [*]	02:16.71 ± 10.04 ^{##}	02:12.95 ± 09.59 [#]	02:13.28 ± 09.71	(b)	$F_{(1 261)} = 10,083$	$P < 0.001$	0.98
					(c)	$F_{(1 261)} = 394$	$P < 0.001$	0.67
40–59th percentiles								
Cross-sectional	03:01.96 ± 13.11	02:42.76 ± 12.54 [#]	02:35.01 ± 11.70 [#]	02:35.45 ± 11.50	(a)	$F_{(1 198)} = 28$	$P < 0.001$	0.12
Longitudinal	02:48.28 ± 11.57 [*]	02:32.53 ± 11.01 ^{##}	02:28.25 ± 10.47 ^{##}	02:31.62 ± 10.68 [#]	(b)	$F_{(1 222)} = 7111$	$P < 0.001$	0.97
					(c)	$F_{(1 222)} = 281$	$P < 0.001$	0.59
1st–20th percentiles								
Cross-sectional	03:34.78 ± 17.78	03:11.49 ± 18.00 [#]	03:05.15 ± 18.83 [#]	03:11.06 ± 20.29 [#]	(a)	$F_{(1 198)} = 27$	$P < 0.001$	0.12
Longitudinal	03:15.87 ± 16.10 [*]	02:58.15 ± 16.84 ^{##}	02:55.09 ± 17.45 ^{##}	03:01.78 ± 18.03 ^{##}	(b)	$F_{(2 305)} = 2062$	$P < 0.001$	0.91
					(c)	$F_{(2 305)} = 80$	$P < 0.001$	0.29

Table 2. Comparison of percentiles based on cross-sectional vs. longitudinal analysis of *female* swimmers. Analysis of variance (ANOVA) with repeated measure and one between subject factor was used to compare upper (80–99th), medium (40–59th), and lower percentiles (01–20th) based on mean [mm:ss.00] ± standard deviation [ss.00] across the 200 m events of all swimming strokes. (a) Main effect: type of analysis (cross-sectional vs. longitudinal). (b) Main effect: age category (early junior–late junior–sub-elite–elite). (c) Interaction effect: type of analysis × age group. *Post-hoc* comparison. ^{*}Significant difference to cross-sectional analysis. [#]Significant difference to previous age category.

Discussion

The present study provides percentile curves for competitive swimmers for short-course race times from 10 years of age to age of peak performance (21–26 years) for both sexes, all swimming strokes, and all race distances. Longitudinal tracking reduced the number of subjects included, and resulted in faster race times compared to the cross-sectional analysis. The older the age category, the more similar became cross-sectional to longitudinal data for the high but not low percentiles. Regarding low (1st–20th) percentiles, race times became faster up to sub-elite age, before deteriorating at elite age both with cross-sectional and longitudinal tracking. High percentile (80–99th) race times improved up to elite age when cross-sectionally analyzed but plateaued at sub-elite age when using longitudinal tracking. A software-based percentile calculator was provided that enables coaches and performance analysts to determine the exact percentile for a particular race time and age group for each of 34 swimming events.

Successful elite age swimmers outperform their lower ranked peers from the age of 12 years onwards¹⁷. Therefore, longitudinal tracking, i.e. swimmers that were retrospectively tracked from peak performance age, showed faster race times during adolescence compared to cross-sectional data, which included all swimmers in each age-group regardless of later drop-outs. Interestingly, regarding high percentiles, cross-sectional data aligned with longitudinal data at sub-elite age. Due to the low transition rate from junior to senior age^{4,5}, swimmers are less likely to drop-out once they reach sub-elite age category^{5,23}. Additionally, performing within the high percentile range requires a structured training process and performance-oriented attitude²⁴. Swimmers competing for fun rather than success likely either drop from the performance level (80–99th percentiles) or out of the sport, while other swimmers accumulate valuable training time and competition experience on their way to elite age success^{25,26}. This intended or natural deselection may explain the lack of a performance difference between cross-sectional and longitudinal data at the high percentiles after sub-elite age.

It is clear that longitudinal tracking dramatically reduced number of subjects included compared to the cross-sectional analysis. Longitudinal tracking excludes data from poor performers who compete for reasons other than winning and eventually drop out of competitive swimming during or after junior age²³. Therefore, percentiles established on retrospective analysis from peak performance age are appropriate to predict performance progression towards elite age¹. This improves practical utility of the percentiles, in particular for the high percentile range, and provides “realistic data on the long-term potential”⁵ to establish development guidelines for high-performance swimmers.

Relative age and biological maturation are known to affect swimming performance during adolescence^{27,28} and may therefore bias longitudinal percentile data. For instance, the same race time, e.g. 02:22.00 [mm:ss.00] in 200 m FR, is rated on the 90th or 68th percentile for a 12 or 13 year old male swimmer, respectively. Still, early (January 1st) and late (December 31st) in the year born swimmers, which have a one-year age difference, are assessed within the same age-group. Therefore, previous studies have developed corrective adjustments for the relative age effect in swimming, which should be considered when interpreting percentiles^{1,29}. However, biological age could not be determined within the present study due to 31.6 million race data included. From a practical perspective, coaches can determine the maturity off-set using invasive or non-invasive methods^{7,8,30} and thus, use the biological, as opposed to the chronological, age of their swimmers for the present percentiles and calculator software. Additionally, bio-banding, i.e. grouping junior athletes for competitions based on biological rather than chronological age³¹, reduced physical but increased technical demand of soccer match play³². Bio-banding has not yet been scientifically evaluated in swimming. However, with the reduced physical advantage of early maturing individuals^{31,32}, bio-banding may improve technical development of early and selection chances for late maturing swimmers.

Irrespective of chronological and biological age differences, the present percentiles were established to assess progression of swimming performance. The aim was to add a variable to talent identification beyond the traditional and one-dimensional comparison of current race times with swimmers from the same age-group. As such, during adolescence, swimmers' race times continuously improve due to growth^{1,18}. Swimmers with average development are expected to follow a particular percentile over the years. However, a lower percentile ranking from one year to another would indicate underperformance despite faster race times. In contrast, some late maturing swimmers may show medium percentile race times. However, annually improved percentile rankings would indicate effective training and high trainability, e.g. the genetic capacity to adapt to training^{6,33,34}. As such, a swimmer's overall trajectory may still be worthwhile to achieve a high percentile when reaching elite age with a possible performance progression of 34–42% from 8 to 18 years of age¹⁸. Therefore, assessment of trainability, along with tight monitoring of performance progression, i.e. percentile ranking and its annual development, may help to discover potentially overlooked talents, even without knowledge of the maturation status. Still, coaches and federation officials should be aware of the relative age effect, which is larger at younger ages than covered by the present percentiles (< 10 years), and incorporate corrective adjustments^{1,28}.

Finally, percentiles may help identify the event in which junior swimmers have their largest potential. Junior swimmers typically compete in multiple swimming strokes and race distances for a broad technical and physiological education^{35,36}, and the individual's strongest events are typically selected based on success in regional and national championships. However, success is affected by performance of the other competitors and the overall performance level of the region and nation. Thus, comparing percentile rankings and their annual development between various swimming strokes and race distances may help identify a swimmer's strengths and weaknesses, as the present percentiles normalize race times to a standard score and provide a relative measure of swimming performance based on international race data^{21,37}.

In conclusion, the present study provides percentile curves for competitive swimmers for all swimming strokes, all race distances, and both sexes for short-course pool events. Longitudinal tracking showed significantly faster race times compared to cross-sectional data. In particular for the high percentiles, i.e. 80–99th, and high-performance swimmers, for drop-outs accounted longitudinal tracking may predict progression towards elite age. The percentile calculator facilitates quick and easy data access for practical application of percentiles in training and competition, while avoiding the inconvenience of searching the exact percentile for a given race time and age from 68 charts or tables with up to 1188 data points each. Performance analysts and coaches can use the percentiles to assess race times and establish individual performance trajectories and define realistic goals for young talented swimmers. As swimmers with average development are expected to follow a particular percentile over the years, the combination of two factors, i.e. current percentile ranking and changes over time, can help to assess trainability and identify talented swimmers during junior age.

Methods

Sample. Race times were provided by the publicly available database of the European Swimming Federation LEN (Ligue Européenne de Natation)²⁰. A total of 31,645,621 short-course (25 m pool length) race times from 2003 to 2019 were included in the study. The study was pre-approved by the internal review board of the Swiss Federal Institute of Sport Magglingen (Reg.-Nr. 124_LSP_201221_234-3.2.127) and is in accordance with the ethical charta of the World Health Organization for studies on human subjects (Helsinki Declaration). No informed consent of the subjects was required, as race times and age were retrieved from a publicly available database.

Data collection/analysis. Two datasets of percentiles were established. The first dataset involved a cross-sectional analysis based on individual annual best times of all swimmers from the database in that period of time (2003–2019) for each specific swimming event. The second dataset involved longitudinal tracking to account for drop-out during junior age. As such, swimmers ($n=8205$) were only included if they still competed at age of peak performance, which occurs between 21 and 26 years of age². For each specific swimming event, swimmers at peak performance age were identified in the 2019 dataset and their annual best times were tracked retrospectively. A minimum of two individual annual best times (one in 2019 and one in another year of the time period investigated) in the particular swimming event were required to be included in the longitudinal analysis.

Percentiles were established for both male and female swimmers across all swimming strokes and all race distances, i.e. BU (50–100–200 m), BA (50–100–200 m), BR (50–100–200 m), FR (50–100–200–400–800–1500 m), and IM (200–400 m). Swimmers typically start their talent pathway with learn-to-swim programs aged 6–10 years³⁸. Subsequently, talented swimmers transition to competitive sports with participation in regional and national competitions. During the initial years of the talent pathway, race distances typically increase with age^{39,40}. Therefore, percentiles were established from the age of 10, 11, 12, and 13 years of age for 50 m, 100 m, 200 m/400 m, 800 m/1500 m events, respectively. To compare the cross-sectional and longitudinal approach, mean race times were compared for each age category: early junior age (10–14 years), late junior age (15–17 years), sub-elite age (18–20 years), and elite age (21–26 years).

Annual best times that were slower than three times the standard deviation for a particular age group were excluded as outliers⁴¹. Following the exclusion of 968,770 outliers (3.06%) from the dataset, percentiles were calculated, with the z-score around the median³⁷. The Lamda-Mu-Sigma (LMS) method was applied to normalize data and account for potential right- and left-sided skewness^{37,42}. The LMS method corrects skewness (L) with the median (M) and coefficient of variation (S), so that the z-score is a valid indicator for the percentiles³⁷. The LMS method is particularly useful when describing non-linear percentile curves during growth and adolescence^{22,37,43}. For the diagrams, percentile curves were smoothed using the cubic spline interpolation³⁷. Percentiles and diagrams were established with RStudio (version 1.1.456, RStudio Team, Boston, United States).

Percentile calculator. A software-based percentile calculator was provided to allow easy access to the percentiles based on the large dataset, which includes 34 swimming pool events. Based on the =VLOOKUP function in Microsoft Excel (Microsoft Corporation, Redmond, WA), the percentile for a particular race time is displayed based on the chosen sex, swimming stroke, and race distance. Additionally, a table of the selected swimming event is displayed to provide an overview of the percentiles across all age categories.

Statistical analysis. A two-way analysis of variance (2×4 ANOVA) with repeated measures and one between-subject factor: type of analysis (cross-sectional vs. longitudinal) \times age category (early junior – late junior – sub-elite – elite) was used to compare mean values \pm standard deviation of the upper (80th–99th), medium (40th–59th), and lower percentiles (1st–20th) with a Tukey's *post-hoc* test, where partial eta-squares indicate a small (0.01), medium (0.06), and large (0.14) effect⁴⁴. If variances were not equal based on the Levene's test, a Bonferroni's *post-hoc* test was applied⁴¹. The Greenhouse–Geisser correction was applied to the main effects with unequal variances of the within-subject factors based on an $\epsilon < 0.75$ in Mauchly's test of sphericity⁴¹. As the 200 m events are the only common race distances across all swimming strokes for Olympic swimming events⁴⁵, the 200 m race times across all swimming strokes were used for the statistical analysis. Normality was confirmed by Gaussian distribution in the histogram and standardized residuals showing a diagonal straight line in the Q–Q plot⁴¹. An alpha-level of 0.05 indicated statistical significance. Statistical analyses were performed using the JASP statistical software package version 0.14 (JASP-Team, University of Amsterdam, Amsterdam, The Netherlands).

Data availability

All data generated and analyzed during the study are included in the published article and its supplementary information file.

Received: 24 November 2021; Accepted: 13 May 2022

Published online: 18 June 2022

References

- Cobley, S. *et al.* Removing relative age effects from youth swimming: The development and testing of corrective adjustment procedures. *J Sci. Med. Sport* **22**, 735–740. <https://doi.org/10.1016/j.jsams.2018.12.013> (2019).
- Allen, S. V., Vandenbogaerde, T. J. & Hopkins, W. G. Career performance trajectories of Olympic swimmers: Benchmarks for talent development. *Eur. J. Sport Sci.* **14**, 643–651. <https://doi.org/10.1080/17461391.2014.893020> (2014).

3. Born, D. P., Lomax, I., Rueger, E. & Romann, M. Normative data and percentile curves for long-term athlete development in swimming. *J. Sci. Med. Sport* **25**, 266–271. <https://doi.org/10.1016/j.jsams.2021.10.002> (2022).
4. Barreiros, A., Cote, J. & Fonseca, A. M. From early to adult sport success: Analysing athletes' progression in national squads. *Eur. J. Sport Sci.* **14**(Suppl 1), S178–182. <https://doi.org/10.1080/17461391.2012.671368> (2014).
5. Brustio, P. R. *et al.* Being a top swimmer during the early career is not a prerequisite for success: A study on sprinter strokes. *J. Sci. Med. Sport* <https://doi.org/10.1016/j.jsams.2021.05.015> (2021).
6. Rees, T. *et al.* The Great British Medalists project: A review of current knowledge on the development of the World's Best Sporting Talent. *Sports Med.* **46**, 1041–1058. <https://doi.org/10.1007/s40279-016-0476-2> (2016).
7. Malina, R. M., Coelho, E. S. M. J., Figueiredo, A. J., Carling, C. & Beunen, G. P. Interrelationships among invasive and non-invasive indicators of biological maturation in adolescent male soccer players. *J. Sports Sci.* **30**, 1705–1717. <https://doi.org/10.1080/02640414.2011.639382> (2012).
8. Malina, R. M., Bouchard, C. & Bar-Or, O. *Growth, Maturation, and Physical Activity* 2nd edn. (Human Kinetics, 2004).
9. Abbott, S. *et al.* Maturity-related developmental inequalities in age-group swimming: The testing of "Mat-CAPs" for their removal. *J. Sci. Med. Sport* **24**, 397–404. <https://doi.org/10.1016/j.jsams.2020.10.003> (2021).
10. Sokolowski, K. *et al.* Biological age in relation to somatic, physiological, and swimming kinematic indices as predictors of 100 m front crawl performance in young female swimmers. *Int. J. Environ. Res. Public Health* **18**, 25. <https://doi.org/10.3390/ijerph18116062> (2021).
11. Lorenzo-Calvo, J. *et al.* Prevalence and impact of the relative age effect on competition performance in swimming: A systematic review. *Int. J. Environ. Res. Public Health* **18**, 25. <https://doi.org/10.3390/ijerph182010561> (2021).
12. Moesch, K., Elbe, A. M., Hauge, M. L. & Wikman, J. M. Late specialization: The key to success in centimeters, grams, or seconds (CGS) sports. *Scand. J. Med. Sci. Sports* **21**, e282–290. <https://doi.org/10.1111/j.1600-0838.2010.01280.x> (2011).
13. Olstad, B. H., Wathne, H. & Gonjo, T. Key factors related to short course 100 m breaststroke performance. *Int. J. Environ. Res. Public Health* **17**, 25. <https://doi.org/10.3390/ijerph17176257> (2020).
14. Wolfrum, M., Knechtle, B., Rust, C. A., Rosemann, T. & Lepers, R. The effects of course length on freestyle swimming speed in elite female and male swimmers—a comparison of swimmers at national and international level. *Springerplus* **2**, 643. <https://doi.org/10.1186/2193-1801-2-643> (2013).
15. Wolfrum, M., Rust, C. A., Rosemann, T., Lepers, R. & Knechtle, B. The effect of course length on individual medley swimming performance in national and international athletes. *J. Hum. Kinet.* **42**, 187–200. <https://doi.org/10.2478/hukin-2014-0073> (2014).
16. Costa, M. J., Marinho, D. A., Bragada, J. A., Silva, A. J. & Barbosa, T. M. Stability of elite freestyle performance from childhood to adulthood. *J. Sports Sci.* **29**, 1183–1189. <https://doi.org/10.1080/02640414.2011.587196> (2011).
17. Post, A. K., Koning, R. H., Visscher, C. & Elferink-Gemser, M. T. Multigenerational performance development of male and female top-elite swimmers—a global study of the 100 m freestyle event. *Scand. J. Med. Sci. Sports* **30**, 564–571. <https://doi.org/10.1111/sms.13599> (2020).
18. Alshdokhi, K., Petersen, C. & Clarke, J. Improvement and variability of adolescent backstroke swimming performance by age. *Front. Sports Act. Liv.* **2**, 46. <https://doi.org/10.3389/fspor.2020.00046> (2020).
19. de Koning, J. J. Data, more data, big data. *Int. J. Sports Physiol. Perform.* **12**, 1. <https://doi.org/10.1123/IJSP.2017-0012> (2017).
20. Swimrankings.net. <https://www.swimrankings.net/index.php?&language=us> (2021). Accessed 2 Aug 2021.
21. Dobosz, J., Mayorga-Vega, D. & Vicianá, J. Percentile values of physical fitness levels among Polish children aged 7 to 19 years—a population-based study. *Cent. Eur. J. Public Health* **23**, 340–351. <https://doi.org/10.21101/cejph.a4153> (2015).
22. Sandercock, G., Voss, C., Cohen, D., Taylor, M. & Stasinopoulos, D. M. Centile curves and normative values for the twenty metre shuttle-run test in English schoolchildren. *J. Sports Sci.* **30**, 679–687. <https://doi.org/10.1080/02640414.2012.660185> (2012).
23. Moulds, K. *et al.* Sink or Swim? A survival analysis of sport dropout in Australian youth swimmers. *Scand. J. Med. Sci. Sports* **30**, 2222–2233. <https://doi.org/10.1111/sms.13771> (2020).
24. Pollock, S. *et al.* Training regimes and recovery monitoring practices of Elite British swimmers. *J. Sports Sci. Med.* **18**, 577–585 (2019).
25. Born, D. P., Stacker, I., Romann, M. & Stoggl, T. Competition age: Does it matter for swimmers?. *BMC Res. Notes* **15**, 82. <https://doi.org/10.1186/s13104-022-05969-6> (2022).
26. Born, D. P. *et al.* Competition-based success factors during the talent pathway of Elite male swimmers. *Front. Sports Act. Liv.* **2**, 25. <https://doi.org/10.3389/fspor.2020.589938> (2020).
27. Oliveira, M. *et al.* Anthropometric variables, propulsive force and biological maturation: A mediation analysis in young swimmers. *Eur. J. Sport Sci.* **21**, 507–514. <https://doi.org/10.1080/17461391.2020.1754468> (2021).
28. Copley, S. *et al.* Transient relative age effects across annual age groups in national level Australian swimming. *J. Sci. Med. Sport* **21**, 839–845. <https://doi.org/10.1016/j.jsams.2017.12.008> (2018).
29. Abbott, S. *et al.* Testing the application of corrective adjustment procedures for removal of relative age effects in female youth swimming. *J. Sports Sci.* **20**, 1–8. <https://doi.org/10.1080/02640414.2020.1741956> (2020).
30. Mirwald, R. L., Baxter-Jones, A. D., Bailey, D. A. & Beunen, G. P. An assessment of maturity from anthropometric measurements. *Med. Sci. Sports Exerc.* **34**, 689–694 (2002).
31. Cumming, S. P., Lloyd, R. S., Oliver, J. L., Eisenmann, J. C. & Malina, R. M. Bio-banding in sport: Applications to competition, talent identification, and strength and conditioning of youth athletes. *Strength Cond. J.* **39**, 34–47. <https://doi.org/10.1519/SSC.000000000000281> (2017).
32. Romann, M., Ludin, D. & Born, D. P. Bio-banding in junior soccer players: A pilot study. *BMC Res. Notes* **13**, 240. <https://doi.org/10.1186/s13104-020-05083-5> (2020).
33. Malina, R. M., Rogol, A. D., Cumming, S. P., Coelho e Silva, M. J. & Figueiredo, A. J. Biological maturation of youth athletes: Assessment and implications. *Br. J. Sports Med.* **49**, 852–859. <https://doi.org/10.1136/bjsports-2015-094623> (2015).
34. Tucker, R. & Collins, M. What makes champions? A review of the relative contribution of genes and training to sporting success. *Br. J. Sports Med.* **46**, 555–561. <https://doi.org/10.1136/bjsports-2011-090548> (2012).
35. Gullich, A. & Emrich, E. Considering long-term sustainability in the development of world class success. *Eur. J. Sport Sci.* **14**(Suppl 1), S383–397. <https://doi.org/10.1080/17461391.2012.706320> (2014).
36. Gullich, A., Macnamara, B. N. & Hambrick, D. Z. What makes a champion? Early multidisciplinary practice, not early specialization, predicts world-class performance. *Perspect. Psychol. Sci.* **17**, 6–29. <https://doi.org/10.1177/1745691620974772> (2022).
37. Indrayan, A. Demystifying LMS and BCPE methods of centile estimation for growth and other health parameters. *Indian Pediatr.* **51**, 37–43. <https://doi.org/10.1007/s13312-014-0310-6> (2014).
38. Olstad, B. H., Berg, P. R. & Kjendlie, P. L. Outsourcing swimming education-experiences and challenges. *Int. J. Environ. Res. Public Health* **18**, 25. <https://doi.org/10.3390/ijerph18010006> (2020).
39. Swimming Australia. *Australian Swimming Framework*. <https://www.swimming.org.au/community-1/coaches/australian-swimming-framework> (2021). Accessed 06 Jun 2021.
40. USA Swimming. *The Pyramid*. <https://www.usaswimming.org/about-usas/organization/overview> (2020). Accessed 06 Jun 2021.
41. Field, A. *Discovering Statistics using IBM SPSS Statistics* 58–375 (Sage Publications Ltd., 2013).
42. Cole, T. J. The LMS method for constructing normalized growth standards. *Eur. J. Clin. Nutr.* **44**, 45–60 (1990).
43. Golle, K., Muehlbauer, T., Wick, D. & Granacher, U. Physical fitness percentiles of German children aged 9–12 years: Findings from a longitudinal study. *PLoS One* **10**, e0142393. <https://doi.org/10.1371/journal.pone.0142393> (2015).

44. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences* 2nd edn, 407 (L. Erlbaum Associates, 1988).
45. IOC. *Swimming events*. <https://olympics.com/tokyo-2020/en/schedule/swimming-schedule> (2021). Accessed 06 Jun 2021.

Author contributions

Conceptualization, D.-P.B. and M.R.; methodology, D.-P.B. and M.R.; software, D.-P.B. and M.R.; validation, D.-P.B., E.R., C.M.B., and M.R.; formal analysis, D.-P.B., E.R., C.M.B., and M.R.; investigation, D.-P.B., E.R., C.M.B., and M.R.; resources, D.-P.B. and M.R.; data curation, D.-P.B., E.R., C.M.B., and M.R.; writing-original draft preparation, D.-P.B.; writing-review and editing, E.R., C.M.B., and M.R.; visualization, D.-P.B. and M.R.; supervision, M.R.; project administration, D.-P.B.; funding acquisition, D.-P.B. and M.R. All authors have read and agreed to the published version of the manuscript.

Funding

The study was financed by the research fund of Swiss Olympic (National Sport Association). However, the funder had no role in the study methodology, data analysis, data interpretation, or decision to publish the article.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-13837-3>.

Correspondence and requests for materials should be addressed to D.-P.B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022