Data Modeling for Inter- and Intra-individual Stability of Young Swimmers' Performance: A Longitudinal Cluster Analysis

Jorge E. Morais, Pedro Forte, Antonio J. Silva, Tiago M. Barbosa & Daniel A. Marinho

To cite this article: Jorge E. Morais, Pedro Forte, Antonio J. Silva, Tiago M. Barbosa & Daniel A. Marinho (2020): Data Modeling for Inter- and Intra-individual Stability of Young Swimmers' Performance: A Longitudinal Cluster Analysis, Research Quarterly for Exercise and Sport, DOI: 10.1080/02701367.2019.1708235

To link to this article: https://doi.org/10.1080/02701367.2019.1708235

Published online: 06 Mar 2020.
The aims of this study were to classify, identify and follow-up young swimmers’ performance and its biomechanical determinants during two competitive seasons (in seven different moments of assessment—M), and analyze the individual variations of each swimmer. Thirty young swimmers (14 boys: 12.70 ± 0.63 years-old; 16 girls: 11.72 ± 0.71 years-old) were recruited. A set of anthropometric, kinematic, efficiency, hydrodynamic and mechanical power variables were assessed. Results: The cluster solution (i.e., number of ideal clusters for this sample) resulted in three clusters, which were named as: cluster 1 (“talented”), cluster 2 (“proficient”), and cluster 3 (“non-proficient”). The performance improved between moments of assessment in all clusters (cluster 1—M1: 68.07 ± 6.62s vs M7: 61.46 ± 3.43s; cluster 2—M1: 73.14 ± 4.87s vs M7: 65.33 ± 2.97s; cluster 3—M1: 82.60 ± 4.18s vs M7: 70.09 ± 3.48s). Anthropometric features also increased between moments of assessment, and remaining biomechanical variables (kinematic, efficiency, hydrodynamic and mechanical power) also increased between M1 and M7, in all clusters. Cluster 1 increased their swimmer’s membership between M1 and M7 (4 to 11), cluster 2 decreased (12 to 5), and cluster 3 maintained (14). Conclusion: It can be concluded that the cluster formation depends on different determinant factors during two competitive seasons, and young swimmers are prone to change from one cluster to another over this period of time.
Given the existence of this variability, other studies have verified that it is still possible to group young swimmers according to their anthropometric and/or biomechanical similarities at a given moment (Barbosa et al., 2014; Figueiredo et al., 2016), or at various moments throughout a competitive season (Morais, Silva, Marinho, Seifert, & Barbosa, 2015). Nevertheless, the information on this topic of study is still scarce since the morphological and technical changes that young swimmers depict over growth and maturation processes may interfere in their performance enhancement. A study analyzed the group membership (variations between sub-groups) of an age-group of young swimmers during one single competitive season (Morais et al., 2015). However, young swimmers keep growing up and the transition between competitive seasons may trigger other determinants that were previously not responsible for performance (Latt et al., 2009; Morais et al., 2015). Moreover, the addition of other main performance determinants, such as variables related to mechanical power (Barbosa et al., 2015a), are not yet commonly assessed in a longitudinal research design. Nonetheless, a recent study indicated that fastest swimmers are more likely to produce higher values of mechanical power in comparison to swimmers in lower tiers (Barbosa, Bartolomeu, Morais, & Costa, 2019).

In this sense, this study aimed to: (i) classify, identify and follow-up young swimmers into sub-groups (clusters), according to the performance and its biomechanical determinants, during two competitive seasons; (ii) analyze the individual variations (hypothetical changes between sub-groups) of each swimmer, at each moment of assessment. It was hypothesized that the determinant factors responsible for the sub-group formation vary from one moment of assessment to another, and that swimmers may change between clusters over time.

**Method**

**Participants**

Thirty young swimmers (14 boys: 12.70 ± 0.63 years-old, Fina points: 234.86 ± 69.76 points; 16 girls: 11.72 ± 0.71 years-old, Fina points: 288.75 ± 67.01 points, at the beginning of the study) were recruited for analysis. Swimmers were in a pre-pubertal stage (Tanner stages 1–2). So, the test for a sex effect (one-way ANOVA, significance set at p ≤ 0.05) showed no significance. In this sense boys and girls were pooled together, as reported elsewhere (Barbosa et al., 2015b). The sample included national record holders, age-group national champions and other swimmers under a talent identification program. Parents or guardians and the swimmers themselves signed an informed consent form. All procedures were in accordance to the Declaration of Helsinki regarding human research, and the University Ethics Board approved the research design.

**Study design**

This framework was based on a longitudinal follow-up design. Participants were evaluated at seven different moments (M_i) over two consecutive seasons. The swimming seasons were planned based on a traditional three peak performance, and each moment of assessment corresponded to one of these peaks (end of each macrocycle). Figure 1 depicts the moments of assessment scheme during the two seasons. Such moments were different in each season according to coaches’ advice (i.e., defined according to the training program and the competitive calendar in each season). The level of evidence based on the United States Preventive Services Task Force (USPSTF) and Level II-3: Evidence obtained from multiple data groups over time with or without intervention, were used.

**Performance**

The official 100m freestyle event (short course meter swimming pool, i.e., 25m length) was selected as the performance outcome. This was chosen because it is the most popular event among these age-group of swimmers. The time gap
between the data collection and performance was no longer than 15 days.

**Anthropometrics**

The body mass (BM) was measured on a digital scale (TANITA, BC-730, Amsterdam, Netherlands). Height (H) was measured as the distance between the vertex to the floor (with the swimmers in the orthostatic position) with a digital stadiometer.

Trunk transverse surface area (TTSA) was measured by digital photogrammetry. The swimmers were photographed by a digital camera (Alpha 6000, Sony, Tokyo, Japan) in the transverse plane (downwards view, i.e., from above) on land simulating the streamlined position (Morais et al., 2011). This position is characterized by the upper limbs being fully extended above the head, one hand over the other, fingers also extended close together and head in neutral position. They wore their regular textile swimsuit, cap and goggles. In the camera shooting field, there was also a calibration pole with 0.945m length at the height of the xiphoid process (Morais et al., 2011). Afterward, the TTSA was measured from each swimmer’s digital photo on dedicated software (Udruler, AVPSOft, USA). For the hand surface area (HSA) and feet surface area (FSA), swimmers placed their dominant hand and foot on the scan copy of a machine, and the file was exported to a PC. The scan surface was also fitted with a 2D calibration frame (Morais et al., 2012). The HSA and FSA were measured using the same procedure described for TTSA. The arm span (AS) was also measured by digital photogrammetry. The AS with the swimmers placed in an orthostatic position, with both arms in lateral abduction at a 90° angle with the trunk. Both arms and fingers were fully extended. The distance between the tip of each third finger was measured.

**Hydrodynamics**

The active drag ($D_a$), and the active drag coefficient ($C_{Da}$) were computed based on the velocity perturbation method (Kolmogorov & Duplishcheva, 1992). Swimmers performed two all-out trials of 25m at front crawl with a push-off start. One trial was carried out towing a hydrodynamic body (i.e., a perturbation device) and one other without towing it (Kolmogorov & Duplishcheva, 1992). A camera (Sony x3000, Tokyo, Japan) was used to record the swimmer’s displacement time between the 11th and 24th meter. The swimming velocity was calculated as: $v = \frac{d}{t}$. The $D_a$ was computed as:

$$D_a = \frac{D_b v_b v^2}{\sqrt{v^3 - v_b^3}}$$  \hspace{1cm} (1)

where $D_a$ is the swimmers’ active drag at maximal velocity (N), $D_b$ is the resistance of the hydrodynamic body computed from the manufacturer’s calibration of the buoy-drag characteristics and its velocity (N), $v_b$ and $v$ are the swimming velocities with and without the perturbation device ($\text{m}\cdot\text{s}^{-1}$). The $C_{Da}$ was computed as:

$$C_{Da} = \frac{2 \cdot D_a}{\rho \cdot TTSA \cdot v^2}$$ \hspace{1cm} (2)

where $C_{Da}$ is the active drag coefficient (dimensionless), $D_a$ is the active drag (N), $\rho$ is the density of the water ($1000 \text{ kg} \cdot \text{m}^{-3}$), TTSA is the trunk transverse surface area ($\text{m}^2$), and $v$ the swimming velocity ($\text{m} \cdot \text{s}^{-1}$).

**Mechanical power**

The power to overcome drag ($P_d$) was computed as (Kolmogorov & Duplishcheva, 1992):

$$P_d = D_a \cdot v$$ \hspace{1cm} (3)

where $P_d$ is the power to overcome drag (W), $D_a$ is the swimmers’ active drag at maximal velocity, and $v$ is the swimming velocity ($\text{m} \cdot \text{s}^{-1}$).

The external mechanical power ($P_{ext}$) and the mechanical power to transfer kinetic energy to water ($P_k$) were computed respectively as (Zamparo, Pendergast, Mollendorf, Termin, & Minetti, 2005):

$$P_{ext} = \frac{P_d}{\eta_f}$$ \hspace{1cm} (4)

where $P_{ext}$ is the external mechanical power (W), $P_d$ is the power to overcome drag (W), and $\eta_f$ is the Froude efficiency (dimensionless, described in the kinematics/efficiency section).

$$P_k = P_{ext} - P_d$$ \hspace{1cm} (5)

where $P_k$ is the mechanical power to transfer kinetic energy to water (W), $P_{ext}$ is the external mechanical power (W), and $P_d$ is the power to overcome drag (W).

**Kinematics/efficiency**

The kinematics and efficiency variables were collected during a 25m maximal front-crawl trial. Swimmers were attached to a speedometer cable (Swim Speedometer, Swimsportec, Hildesheim, Germany) (Barbosa et al., 2015b). A customized software interface in LabVIEW (National Instruments, v. 2009, Austin, TX, USA) was selected to acquire ($f = 50 \text{ Hz}$), display, and process speed-time data online during the trial.
A 12-bit resolution acquisition card (USB-6008; National Instruments, Austin, TX, USA) was used to transfer the data from the mechanical apparatus to the software application. Thereafter, data were exported to a signal processing software (AcqKnowledge v. 3.9.0, Biopac Systems, Santa Barbara, CA, USA) and filtered with a 5 Hz cutoff low-pass fourth-order Butterworth. The intra-cyclic variation of the swimming velocity (dv) was computed as (Barbosa et al., 2010):

\[ dv = \sqrt{\frac{\sum_{i=1}^{n} (v_i - \bar{v}) F_i}{\sum_{i=1}^{n} v_i F_i}} \]  

(6)

where dv is the intra-cyclic variation of the swimming velocity (dimensionless), \( v \) is the mean swimming velocity (m·s\(^{-1}\)), \( v_i \) is the instant swimming velocity (m·s\(^{-1}\)), \( F_i \) is the acquisition frequency, and \( n \) is the number of observations. The mean value of three consecutive full strokes between during the middle 15m were considered for analysis. The stroke frequency (SF) was measured by calculating the number of cycles per unit of time from the time it takes to complete one full cycle (f = 1/t, the mean of three consecutive full stroke cycles was used for analysis), and afterward converted to Hz. The stroke length (SL) was computed as (Craig & Pendergast, 1979): \( SL = v/\text{SF} \). Where SL is the stroke length (m), \( v \) is the swimming velocity (m·s\(^{-1}\)), and SF is the stroke frequency (Hz).

The stroke index (SI) was computed as (Costill et al., 1985): \( SI = v \cdot SL \). Where SI is the stroke index (m\(^2\)·s\(^{-1}\)), \( v \) is the swimming velocity (m·s\(^{-1}\)), and SL the stroke length (m). The Froude efficiency (\( \eta_F \)) was computed as (Zamparo et al., 2005):

\[ \eta_F = \left( \frac{v \cdot 0.9}{2\pi \cdot SF \cdot l} \right) \cdot 2 \]  

(7)

where \( \eta_F \) is the Froude efficiency (dimensionless), \( v \) is the swimming velocity (m·s\(^{-1}\)), SF is the stroke frequency (Hz), and \( l \) is the shoulder to hand average distance (in meter).

**Data analysis**

The normality and homoscedasticity assumptions were analyzed with Kolmogorov-Smirnov and Levene tests, respectively. The mean plus one standard deviation, and the 95% confidence interval (95CI) were calculated as descriptive statistics. The cluster modeling was performed based in the k-means approach (nonhierarchical), which permitted the prior definition of several clusters to be used. The k-means defines a centroid (i.e., the mean of a group of points/subjects), based on their similarities (Rein, Button, Davids, & Summers, 2010). To ensure a coherent comparison of data sets with different magnitudes and/or units, standardized z-scores of all variables were computed.

The one-way ANOVA was used to identify the main determinants responsible for the cluster formation in each moment of assessment (\( p \leq 0.05 \)). A discriminant analysis was performed to validate the cluster formation. The total eta square (\( \eta^2 \)) was selected as effect size index, and deemed as: (i) without effect if \( 0 < \eta^2 \leq 0.04 \); (ii) minimum if \( 0.04 < \eta^2 \leq 0.25 \); (iii) moderate if \( 0.25 < \eta^2 \leq 0.64 \) and; (iv) strong if \( \eta^2 > 0.64 \) (Ferguson, 2009). The swimmers’ changes between clusters were analyzed by visual inspection.

**Results**

The model defined 3 different clusters. Cluster 1 includes the fastest swimmers according to their performance in the 100m freestyle, cluster 2 includes the intermediate swimmers, and cluster 3 includes the slowest swimmers. Thus, cluster 1 was labeled as “talented”, cluster 2 as “proficient”, and cluster 3 as “non-proficient”. The performance improved over time in all clusters (cluster 1—M1: 68.07 ± 6.62s vs M7: 61.46 ± 3.43s, cluster 2—M1: 73.14 ± 4.87s vs M7: 65.33 ± 2.97s, cluster 3—M1: 82.60 ± 4.18s vs M7: 70.09 ± 3.48s). Between the end of season 1 and beginning of season 2 (M3 vs M4: detraining), swimmers showed a performance impairment (Table 1).

All anthropometric variables increased between the first (M1) and last (M7) moment of assessment (Table 1). Overall, the kinematic variables (SF, SL, \( v \), and dv), mechanical power (\( P_{\text{int}} \), \( P_{\text{ext}} \) and \( P_d \)) and efficiency (\( \eta_F \) and SI), improved between the first (M1) and last (M7) moment of assessment. The hydrodynamic variables (\( D_a \) and \( C_{D_a} \)) increased (i.e., negative effect to the performance), but this fact may be related to the increase in swimming velocity, which is directly related to the drag (Table 1).

The determinant factors that better discriminated (based on higher F-ratio and standardized \( \eta^2 \) effect) the cluster solutions differed between moments of assessment (M1: CP and BM; M2: \( P_d \) and \( P_{\text{ext}} \); M3: H and AS; M4: BM and CP; M5: H and \( P_k \); M6: CP and BM; M7: FSA and BM) (Table 1).

The determinant factors that characterized each cluster (responsible for the swimmers’ similarities) in each moment of assessment, are presented in Table 1, and summarized in Table 2. For example, in the first moment of assessment (M1), cluster 1 was characterized by high \( P_{\text{ext}} \), \( P_d \), and \( P_k \) (power), cluster 2 by large BM, H and CP (anthropometrics), and cluster 3 by
Table 1. Descriptive statistics (mean ± one standard deviation), and 95CI for all parameters analyzed in each cluster, at each moment of assessment. The F-ratios are also presented, showing the determinant factors responsible for the cluster formation (and correspondent effect size: eta square—$\eta^2$).

### M1

| Parameter | Clusters | Mean ± 1SD (95CI) | Z | f | p | $\eta^2$ |
|-----------|----------|-------------------|---|---|---|---------|
|          | Cluster 1 (N = 4) |                      |   |   |   |         |
|          |          | Mean ± 1SD (95CI) |   |   |   |         |
|           |          | 11.32 ± 0.42 (12.43; 13.80) | 1.1260 |   |   |         |
|            |          | 21.75 ± 0.68 (21.75; 21.75) | 0.1730 |   |   |         |
|            |          | 165.6 ± 6.24 (155.5; 175.4) | 1.0416 |   |   |         |
|            |          | 88.5 ± 0.70 (85.5; 91.5) | 0.1260 |   |   |         |
|            |          | 38.4 ± 0.50 (36.9; 40.0) | 0.1260 |   |   |         |
|            |          | 105.7 ± 6.85 (100.0; 111.4) | 0.1260 |   |   |         |
|            |          | 3.1 ± 0.05 (3.0; 3.2) | 0.1260 |   |   |         |
|           | Cluster 2 (N = 12) |                      |   |   |   |         |
|           |          | 12.31 ± 0.42 (11.79; 12.84) | 0.1633 |   |   |         |
|           |          | 21.75 ± 0.68 (21.75; 21.75) | 0.1730 |   |   |         |
|           |          | 165.6 ± 6.24 (155.5; 175.4) | 1.0416 |   |   |         |
|           |          | 88.5 ± 0.70 (85.5; 91.5) | 0.1260 |   |   |         |
|           |          | 38.4 ± 0.50 (36.9; 40.0) | 0.1260 |   |   |         |
|           |          | 105.7 ± 6.85 (100.0; 111.4) | 0.1260 |   |   |         |
|           |          | 3.1 ± 0.05 (3.0; 3.2) | 0.1260 |   |   |         |
|           | Cluster 3 (N = 14) |                      |   |   |   |         |
|           |          | 11.32 ± 0.42 (11.79; 12.84) | 0.1633 |   |   |         |
|           |          | 21.75 ± 0.68 (21.75; 21.75) | 0.1730 |   |   |         |
|           |          | 165.6 ± 6.24 (155.5; 175.4) | 1.0416 |   |   |         |
|           |          | 88.5 ± 0.70 (85.5; 91.5) | 0.1260 |   |   |         |
|           |          | 38.4 ± 0.50 (36.9; 40.0) | 0.1260 |   |   |         |
|           |          | 105.7 ± 6.85 (100.0; 111.4) | 0.1260 |   |   |         |
|           |          | 3.1 ± 0.05 (3.0; 3.2) | 0.1260 |   |   |         |

### M2

| Parameter | Clusters | Mean ± 1SD (95CI) | Z | f | p | $\eta^2$ |
|-----------|----------|-------------------|---|---|---|---------|
|          | Cluster 1 (N = 5) |                      |   |   |   |         |
|          |          | Mean ± 1SD (95CI) |   |   |   |         |
|           |          | 13.42 ± 0.79 (12.65; 14.18) | 1.0400 |   |   |         |
|            |          | 25.6 ± 0.79 (24.8; 26.4) | 1.0400 |   |   |         |
|            |          | 165.6 ± 6.24 (155.5; 175.4) | 1.0416 |   |   |         |
|            |          | 88.5 ± 0.70 (85.5; 91.5) | 0.1260 |   |   |         |
|            |          | 38.4 ± 0.50 (36.9; 40.0) | 0.1260 |   |   |         |
|            |          | 105.7 ± 6.85 (100.0; 111.4) | 0.1260 |   |   |         |
|            |          | 3.1 ± 0.05 (3.0; 3.2) | 0.1260 |   |   |         |
|           | Cluster 2 (N = 8) |                      |   |   |   |         |
|           |          | 13.42 ± 0.79 (12.65; 14.18) | 1.0400 |   |   |         |
|           |          | 25.6 ± 0.79 (24.8; 26.4) | 1.0400 |   |   |         |
|           |          | 165.6 ± 6.24 (155.5; 175.4) | 1.0416 |   |   |         |
|           |          | 88.5 ± 0.70 (85.5; 91.5) | 0.1260 |   |   |         |
|           |          | 38.4 ± 0.50 (36.9; 40.0) | 0.1260 |   |   |         |
|           |          | 105.7 ± 6.85 (100.0; 111.4) | 0.1260 |   |   |         |
|           |          | 3.1 ± 0.05 (3.0; 3.2) | 0.1260 |   |   |         |
|           | Cluster 3 (N = 17) |                      |   |   |   |         |
|           |          | 13.42 ± 0.79 (12.65; 14.18) | 1.0400 |   |   |         |
|           |          | 25.6 ± 0.79 (24.8; 26.4) | 1.0400 |   |   |         |
|           |          | 165.6 ± 6.24 (155.5; 175.4) | 1.0416 |   |   |         |
|           |          | 88.5 ± 0.70 (85.5; 91.5) | 0.1260 |   |   |         |
|           |          | 38.4 ± 0.50 (36.9; 40.0) | 0.1260 |   |   |         |
|           |          | 105.7 ± 6.85 (100.0; 111.4) | 0.1260 |   |   |         |
|           |          | 3.1 ± 0.05 (3.0; 3.2) | 0.1260 |   |   |         |

(Continued)
| Clusters | Mean ± 1SD | z   | p  | n²  |
|----------|-----------|-----|----|-----|
| Cluster 1 (N = 5) | 13.94 ± 0.41 (13.27; 14.61) | 1.3225 | 0.1974 |
| Cluster 2 (N = 8) | 13.00 ± 0.65 (12.66; 13.35) | 0.1974 | 0.1974 |
| Cluster 3 (N = 17) | 12.13 ± 0.57 (11.72; 12.55) | -0.8448 | 13.99 <0.001 0.51 |
| Mean ± 1SD | 0.8615 | 0.0212 |
| z   | 1.2325 | 0.1974 |
| p  | 0.1974 | 0.1974 |
| n²  | 0.1974 | 0.1974 |
| Cluster 1 (N = 4) | 17.5 ± 0.13 (15.2; 19.17) | 1.3680 | 0.0163 |
| Cluster 2 (N = 16) | 14.0 ± 0.17 (13.15; 15.4) | 1.2470 | 0.0238 |
| Cluster 3 (N = 10) | 13.9 ± 0.15 (13.02; 14.25) | 1.1315 | 0.005 0.40 |
| Mean ± 1SD | 0.0615 | 0.0212 |
| z   | 1.3680 | 0.0163 |
| p  | 0.0163 | 0.0163 |
| n²  | 0.0163 | 0.0163 |
| Cluster 1 (N = 6) | 14.00 ± 0.53 (13.44; 14.56) | 1.0595 | 0.0864 |
| Cluster 2 (N = 17) | 13.16 ± 0.61 (12.84; 13.47) | 0.0864 | 0.0864 |
| Cluster 3 (N = 7) | 12.13 ± 0.42 (11.73; 12.52) | -1.1494 | 18.02 <0.001 0.57 |
| Mean ± 1SD | 0.0615 | 0.0212 |
| z   | 1.0595 | 0.0864 |
| p  | 0.0864 | 0.0864 |
| n²  | 0.0864 | 0.0864 |

(Continued)
## Table 1 (Continued)

### M4

| Clusters | N | Mean ± 1SD | z | F | p | \( \eta^2 \) |
|----------|---|------------|---|---|---|------------|
| Cluster 1 (N = 6) | | 15.908 | 152.92 ± 41.64 (131.51; 174.34) | -0.4156 |
| Cluster 2 (N = 17) | | 162.50 ± 60.27 (106.75; 218.24) | -0.3540 | 25.64 | <0.001 | 0.65 |
| Cluster 3 (N = 7) | | 117.20 ± 43.00 (77.42; 156.97) | -0.3201 | 26.11 | <0.001 | 0.66 |
| | | 145.88 ± 5.94 (151.65; 158.11) | 0.0000 | 24.74 | <0.001 | 0.65 |
| Cluster 4 (N = 13) | | 140.58 ± 21.04 (119.94; 161.22) | 0.0000 | 24.74 | <0.001 | 0.65 |

### M5

| Clusters | N | Mean ± 1SD | z | F | p | \( \eta^2 \) |
|----------|---|------------|---|---|---|------------|
| Cluster 1 (N = 5) | | 0.29 ± 0.02 (0.26; 0.32) | 0.31 ± 0.02 (0.30; 0.32) | 0.3415 |
| Cluster 2 (N = 12) | | 0.31 ± 0.02 (0.30; 0.32) | 0.3415 |
| Cluster 3 (N = 13) | | 0.31 ± 0.02 (0.30; 0.32) | 0.3415 |

### M6

| Clusters | N | Mean ± 1SD | z | F | p | \( \eta^2 \) |
|----------|---|------------|---|---|---|------------|
| Cluster 1 (N = 8) | | 0.29 ± 0.02 (0.26; 0.32) | 0.31 ± 0.02 (0.30; 0.32) | 0.3415 |
| Cluster 2 (N = 15) | | 0.31 ± 0.02 (0.30; 0.32) | 0.3415 |
| Cluster 3 (N = 7) | | 0.31 ± 0.02 (0.30; 0.32) | 0.3415 |

(Continued)
Table 1. (Continued).

|                | Cluster 1 (N = 8) | Mean ± 1SD | z  | F     | p     | η2  |
|----------------|------------------|------------|----|-------|-------|-----|
| Pext [N]       | 534.02 ± 221.29  | 391.18 ± 115.57 | 0.0512 |
| P0 [N]         | 370.32 ± 149.55  | 263.76 ± 82.00  | 0.0378 |
| nη [%]         | 0.30 ± 0.02      | 0.33 ± 0.04 (0.30; 0.35) | -0.0570 |
| SI [m s⁻¹]     | 2.81 ± 0.40      | 2.47 ± 0.20 (2.36; 2.58) | -0.1248 |
| Perf [s]       | 60.36 ± 2.09     | 66.95 ± 3.72 (64.89; 69.02) | 0.0960 |

|                | Cluster 2 (N = 15) | Mean ± 1SD | z  | F     | p     | η2  |
|----------------|--------------------|------------|----|-------|-------|-----|
| Pext [N]       | 168.4 ± 91.89     | 186.5 ± 115.57 | 0.0512 |
| P0 [N]         | 120.46 ± 66.19    | 120.46 ± 66.19 | 0.0378 |
| nη [%]         | 0.36 ± 0.04       | 0.36 ± 0.04 (0.33; 0.40) | 0.8095 |
| SI [m s⁻¹]     | 2.25 ± 0.18       | 2.25 ± 0.18 (2.08; 2.42) | -0.7773 |
| Perf [s]       | 72.30 ± 3.11      | 72.30 ± 3.11 (69.42; 75.18) | 1.0989 |

|                | Cluster 3 (N = 7) | Mean ± 1SD | z  | F     | p     | η2  |
|----------------|------------------|------------|----|-------|-------|-----|
| Pext [N]       | 186.4 ± 91.89    | 186.4 ± 91.89 | 0.0512 |
| P0 [N]         | 120.46 ± 66.19   | 120.46 ± 66.19 | 0.0378 |
| nη [%]         | 0.36 ± 0.04      | 0.36 ± 0.04 (0.33; 0.40) | 0.8095 |
| SI [m s⁻¹]     | 2.25 ± 0.18      | 2.25 ± 0.18 (2.08; 2.42) | -0.7773 |
| Perf [s]       | 72.30 ± 3.11     | 72.30 ± 3.11 (69.42; 75.18) | 1.0989 |

Note. BM = body mass; H = height; AS = arm span; TSSA = trunk transverse surface area; HSA = hand surface area; FSA = feet surface area; SI = stroke frequency; SL = stroke length; Vel = swimming velocity; dv = intra-cyclic swimming velocity; D0 = drag coefficient; P0 = power to overcome drag; Pext = external mechanical power; nη = Froude efficiency; SI = stroke index; Perf = performance; z = standardized coefficient; F = F-ratios; η2 = standardized eta square; * and # = identify the main determinants of each cluster of each moment of assessment (* = high; # = low).
small BM, H and CP (anthropometrics). In the last moment of assessment (M7), cluster 1 by high BM, HAS and FSA (anthropometrics), cluster 2 by high $D_a$, $P_d$, and low dv (hydrodynamics, power and kinematics), and cluster 3 by low BM, H and v (anthropometrics and kinematics).

For a qualitative analysis, the discriminant analysis showed a very good compactness/separation at all moments of assessment, with a correct classification of the original groups varying between 93.3% (M2) and 100% (M3, M4, M6, and M7) (Figure 2).

Table 2. Summary of key-features characterizing each cluster over two seasons.

| Cluster | M1               | M2               | M3               | M4               | M5               | M6               | M7               |
|---------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|
| Cluster 1 | Power (+)        | Power (+)        | Power (+)        | Power (+)        | Power (+)        | Anthropometrics (+) | Anthropometrics (+) |
|         | Kinematics (+)   | Anthropometrics (+) | Power (+)        | Kinematics (+)   | Power (+)        | Anthropometrics (+) | Kinematics (+)   |
| Cluster 2 | Anthropometrics (+) | Anthropometrics (+) | Anthropometrics (+) | Kinematics (+) | Power (+)        | Efficiency (+)   | Anthropometrics (+) |
|         |                 |                 |                 | Power (+)        |                 | Kinematics (+) | Power (+)        |
| Cluster 3 | Anthropometrics (−) | Anthropometrics (−) | Anthropometrics (−) | Kinematics (−) | Efficiency (−) | Power (−)        | Anthropometrics (−) |
|         |                 |                 |                 | Kinematics (−) |                 | Power (−)        | Kinematics (−) |

Note. (+) = high values; (−) = low values.

Figure 3 (panel a) presents the cluster membership in each moment of assessment. Between the first (M1) and last (M7) moment of assessment, cluster 1 increased their membership (from four to 11 swimmers), cluster 2 decreased (from 12 to five swimmers), and cluster 3 maintained (14 swimmers). A deeper insight of the swimmers’ shifts is portrayed in Figure 3 (panel b). The initial four swimmers included in cluster 1 in M1, maintained their membership in M7. From the 12 swimmers included in cluster 2, only four kept their membership at M7. Thus, remaining
eight shifted to other clusters. On the other hand, out of the 14 swimmers assigned to cluster 3 at M1, 11 kept the same membership at M7 (only three changed to another cluster). Altogether, out of 30 swimmers, only eight did not shift the cluster membership at any moment. The remaining 22 changed their cluster membership at least once over two consecutive seasons.

Discussion

The main aims of this study were to classify, identify and follow-up young swimmers’ performance and determinant factors, gathered into sub-groups during two competitive seasons, and analyzing the variations in cluster membership over time. It was verified that the selected variables increased and/or improved between the first and the last moment of assessment. In each moment of assessment, the determinant factors responsible for the characterization of the cluster were different. It was found that most swimmers shift clusters over time.

The performance improved over the two seasons, and the determinant factors presented an overall trend to increase/enhance. Previous studies monitoring young swimmers’ have shown an increase of the performance and its determinant factors between moments of assessment (Latt et al., 2009; Morais et al., 2014; Zacca et al., 2019a). In the present study, performance showed the same trend (except between the end of the first season and the beginning of the second).

However, if more moments of assessment are included (as in this study), some of the determinant factors included; kinematic, efficiency, hydrodynamic and mechanical power variables, may show some slight and circumstantial decreases/increases (i.e., sinusoidal profile) between moments of assessment (Table 1). These increases and/or decreases do not occur
concurrently in all determinant factors (e.g., between M1 and M2: the v and SL decreased, and the D₃ increased). Moreover, the non-linear trend may not occur simultaneously in all clusters as well. For example, between M5 and M6 the SF increased in the swimmers included in cluster 1 and 2, but decreased in cluster 3.

Most studies with young swimmers seek to determine/predict which are the performance determinants, but designing cross-sectional studies (Garrido et al., 2012; Saavedra, Escalante, & Rodriguez, 2010). These cross-sectional studies do not give insight into the variations that occur in the performance determinants over time. The data of the present study shows that young swimmers are prone to change (increase/decrease) their stroke mechanics (kinematic and efficiency), hydrodynamics and mechanical power at least three times in each competitive season. As sub-groups were formed, a deeper insight shows that such variations may not occur at the time same in all clusters.

The best solution to classify and identify the swimmers during two competitive seasons consisted in three clusters. Thus, it was decided to use the labels already found in the literature (Barbosa et al., 2014; Morais et al., 2015). This cluster formation is due to the similarity/difference between swimmers. Throughout the seven moments of assessment, different determinant factors explained the clusters’ formation, indicating that young swimmers’ performance is a holistic phenomenon (Morais et al., 2017) (Table 1). For swimmers of a similar age-group, and included in the same competition level, it was noted that the main determinant factors that contributed to grouping of swimmers (the main factor of differentiation) varied at different moments of assessment (e.g., in M1 it was the CP and BM; and at M2 the Pₜ and Pₑₓₜ) (Table 1). This information is relevant for coaches and swimmers, since it is possible to state that within the same age-group and/or competitive level, it is likely to group swimmers according to their anthropometric and/or technical characteristics (Silva et al., 2019b). This suggests that coaches should not put all swimmers of the same age-group and/or competitive level under the same training regime as if one size would fit all. Coaches should design the training and developing programs in tandem to anthropometric and biomechanical characteristics of each of the sub-group of swimmers (Table 2).

Overall, it was suggested that the evolution of an age-group of swimmers is under a process of natural selection. For example, during a competitive season, the number of “talented” swimmers will decrease, and likewise the number of swimmers assigned to the “non-proficient” cluster will increase (Morais et al., 2015). However, this is just an overall analysis of the swimmers’ stability in each cluster. Indeed, it is not possible to present a deeper approach understanding the constituent factors of the number of swimmers who change cluster.

For example, a cluster can keep the same number of swimmers from one moment of assessment to another even if new swimmers are assigned to this cluster and others leave. The stability remains the same, despite not giving insight into individual variation (Morais et al., 2015). This issue of individuality seems to be extremely important, since some studies suggested that each athlete should be seen as “unique” and that there are several ways to improve performance (Durand-Bush & Salmela, 2002; Morais et al., 2014). Thus, a visual inspection was used to understand the individual variation of each swimmer (Figure 3b). It is possible to note that the natural selection argument may not be clear. Data from this study shows that in the first season the number of “talented” swimmers was the same in M1 and M3 (4 swimmers), “proficient” increased (from 12 to 16), and “non-proficient” swimmers decreased (from 14 to 10). However, immediately at the beginning of the second competitive season (M4), the “talented” cluster increased their membership (six swimmers).

Interestingly, in M4 (baseline of the second season) the determinant factors responsible for the cluster formation were the BM, CP, and H (anthropometric features). Swimmers are growing over the break when they are not undergoing training, which will may promote significant changes in their technical profile, enhancing their performance (Moreira et al., 2014; Zacca et al., 2019b). Thus, assessing young swimmers during a longer timeframe (including the non-training periods) will permit a deeper understanding of the swimmers’ individual changes. Moreover, it is also possible to note that from the four swimmers initially assigned to “talented” cluster, three changed (all dropped one level, from cluster 1 to cluster 2), and hence only one kept his/her membership during the entire two seasons (S7; Figure 3b). On the other hand, five swimmers that were included in cluster 2 (“proficient”) were promoted to the “talented” cluster (S1, S10, S11, S14, and S24), and two swimmers included in the “non-proficient” cluster at M1 were able to change their membership to the “talented” one at M7 (S21 and S30). In this sense, it seems that training does have a major effect on performance enhancement at young ages, simultaneously with the anthropometric features.
As main limitations it can be pointed out: (i) the determinant factors responsible for the cluster formation in each moment of assessment may be only reliable for short-distance events (i.e., 100m freestyle); (ii) in further studies the relationship between the external workload and the main determinants of each cluster could be checked.

What does this article add?

Since the determination/formation of each cluster depended on different factors in each one of the moments of assessment, perhaps swimming organizations could become aware of the existence of sub-groups for swimmers in the same age bracket or competitive level. Moreover, those sub-groups might not be determined by the same determinant factors. So, coaches should monitor and evaluate their swimmers in order to understand in which subgroup a given swimmer is clustered and which are the main determinants responsible for the performance. This will help coaches to apply a specific training plan according to such swimmer characteristics, and understand which determinants the swimmer can improve. Additionally, coaches should be aware that each swimmer can shift his/her membership over time. Therefore, as much as possible, coaches should design training and development programs catering the needs of each sub-group of swimmers and re-assign them to the most effective program depending on their state of development (depending on the cluster membership) at any given time. Since the determinant factors that are responsible for each cluster performance are different and can change in a short timeframe, training should be specific for each cluster (group of swimmers with similarities). Coaches should monitor their swimmers regularly in order to understand when they are ready to shift to a better cluster (i.e., faster), or when they shift to a worst cluster (i.e., slower), and they are able to understand such reasoning and change/adapt the training plan.

In conclusion, it is possible to group a set of swimmers of the same age-group and/or competition level into three different sub-groups (clusters). The factors responsible for the cluster formation vary over time and the rate of cluster membership is also significant. Coaches should pay attention to the swimmer’s intra-variability, i.e., swimmer’s changes between clusters. Swimmers included in the fastest cluster may not maintain his/her membership during a timeframe, and shift to a slower cluster. The inverse may also occur, where slower swimmers may shift to fastest clusters. In summary, this research supports the understanding that performance at such early ages is a non-linear, dynamic, complex and holistic phenomenon.

Acknowledgments

We would like also to thank the support of the University of Beira Interior and Santander Universities (Bolsa BIPD/ICIFCSH-Santander Universidades-UBI/2017).

Funding

This project was supported by the National Funds through FCT - Portuguese Foundation for Science and Technology [UID/DTP/04045/2019], and the European Fund for regional development (FEDER) allocated by European Union through the COMPETE 2020 Programme [POCI-01-0145-FEDER-006969].

ORCID

Jorge E. Morais @ http://orcid.org/0000-0002-6885-0648
Pedro Forte @ http://orcid.org/0000-0003-0184-6780
Tiago M. Barbosa @ http://orcid.org/0000-0001-7071-2116

References

Amaro, N. M., Marinho, D. A., Marques, M. C., Batalha, N. P., & Morouço, P. J. (2017). Effects of dry-land strength and conditioning programs in age group swimmers. *Journal of Strength and Conditioning Research*, 31(9), 2447–2454. doi:10.1519/JSC.0000000000001709
Barbosa, T. M., Bartolomeu, R., Morais, J. E., & Costa, M. J. (2019). Skillfull swimming in age-groups is determined by anthropometrics, biomechanics and energetics. *Frontiers in Physiology*, 10, 73.
Barbosa, T. M., Bragada, J. A., Reis, V. M., Marinho, D. A., Carvalho, C., & Silva, A. J. (2010). Energetics and biomechanics as determining factors of swimming performance: Updating the state of the art. *Journal of Science and Medicine in Sport*, 13(2), 262–269. doi:10.1016/j.jsams.2009.01.003
Barbosa, T. M., Morais, J. E., Costa, M. J., Goncalves, J., Marinho, D. A., & Silva, A. J. (2014). Young swimmers’ classification based on kinematics, hydrodynamics, and anthropometrics. *Journal of Applied Biomechanics*, 30(2), 310–315. doi:10.1123/jab.2013-0038
Barbosa, T. M., Morais, J. E., Marques, M. C., Costa, M. J., & Marinho, D. A. (2015a). The power output and sprinting performance of young swimmers. *Journal of Strength Conditioning Research*, 29(2), 440–450. doi:10.1519/JSC.0000000000006626
Barbosa, T. M., Morais, J. E., Marques, M. C., Silva, A. J., Marinho, D. A., & Kee, Y. H. (2015b). Hydrodynamic profile of young swimmers: Changes over a competitive season. *Scandinavian Journal of Medicine and Science in Sports*, 25(2), e184–e196. doi:10.1111/smss.2015.25.issue-2
Costill, D. L., Kovalesski, J., Porter, D., Kirwan, J., Fielding, R., & King, D. (1985). Energy expenditure during front crawl swimming: Predicting success in middle-distance events. *International Journal of Sports Medicine*, 6(5), 266–270. doi:10.1055/s-2008-1025849
Craig, A. B., Jr., & Pendergast, D. R. (1979). Relationships of stroke rate, distance per stroke, and velocity in competitive swimming. *Medicine and Science in Sports, 11*(3), 278–283.

Durand-Bush, N., & Salmela, J. H. (2002). The development and maintenance of expert athletic performance: Perceptions of world and Olympic champions. *Journal of Applied Sport Psychology, 14*(3), 154–171. doi:10.1080/10413200290103473

Ferguson, C. J. (2009). An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice, 40*(5), 532–538. doi:10.1037/a0015808

Figueiredo, P., Pendergast, D. R., Vilas-Boas, J. P., & Fernandes, R. J. (2013). Interplay of biomechanical, energetic, coordinative, and muscular factors in a 200 m front crawl swim. *Biomedical Research International, 2013, 897232*. doi:10.1155/2013/897232

Figueiredo, P., Silva, A., Sampaio, A., Vilas-Boas, J. P., & Fernandes, R. J. (2016). Front Crawl Sprint Performance: A cluster analysis of biomechanics, energetics, coordinative, and anthropometric determinants in young Swimmers. *Motor Control, 20*(3), 209–221. doi:10.1123/mc.2014-0050

Garrido, N. D., Silva, A. J., Fernandes, R. J., Barbosa, T. M., Costa, A. M., Marinho, D. A., & Marques, M. C. (2012). High level swimming performance and its relation to non-specific parameters: A cross-sectional study on maximum handgrip isometric strength. *Perceptual Motor Skills, 114*(3), 936–948. doi:10.2466/05.25.PMS.114.3.936-948

Kolmogorov, S. V., & Dupilishcheva, O. A. (1992). Active drag, useful mechanical power output and hydrodynamic force coefficient in different swimming strokes at maximal velocity. *Journal of Biomechanics, 25*(3), 311–318. doi:10.1016/0021-9290(92)90028-Y

Latt, E., Jurimae, J., Haljaste, K., Cicchella, A., Purge, P., & Jurimae, T. (2009). Longitudinal development of physical and performance parameters during biological maturation of young male swimmers. *Perceptual Motor Skills, 108*(1), 297–307. doi:10.2466/pms.108.1.297-307

Marinho, D. A., Barbosa, T. M., Costa, M. J., Figueiredo, C., Reis, V. M., Silva, A. J., & Marques, M. C. (2010). Can 8-weeks of training affect active drag in young swimmers? *Journal of Sports Science and Medicine, 9*(1), 71–78.

Moraes, J. E., Costa, M. J., Meijas, E. J., Marinho, D. A., Silva, A. J., & Barbosa, T. M. (2011). Morphometric study for estimation and validation of trunk transverse surface area to assess human drag force on water. *Journal of Human Kinetics, 28*, 5–13. doi:10.2478/v10078-011-0017-x

Moraes, J. E., Jesus, S., Lopes, V., Garrido, N., Silva, A., Marinho, D., & Barbosa, T. M. (2012). Linking selected kinematic, anthropometric and hydrodynamic variables to young swimmer performance. *Pediatric Exercise Science, 24*(4), 649–664. doi:10.1123/pcs.24.4.649

Moraes, J. E., Marques, M. C., Marinho, D. A., Silva, A. J., & Barbosa, T. M. (2014). Longitudinal modeling in sports: Young swimmers’ performance and biomechanics profile. *Human Movement Science, 37*, 111–122. doi:10.1016/j.humov.2014.07.005

Morais, J. E., Silva, A. J., Marinho, D. A., Lopes, V. P., & Barbosa, T. M. (2017). Determinant factors of long-term performance development in young swimmers. *International Journal of Sports Physiology and Performance, 12*(2), 198–205. doi:10.1123/ijsspp.2015-0420

Morais, J. E., Silva, A. J., Marinho, D. A., Marques, M. C., Batalha, N., & Barbosa, T. M. (2016). Modelling the relationship between biomechanics and performance of young sprinting swimmers. *European Journal of Sport Science, 16*(6), 661–668. doi:10.1080/17461391.2016.1149227

Morais, J. E., Silva, A. J., Marinho, D. A., Seifert, L., & Barbosa, T. M. (2015). Cluster stability as a new method to assess changes in performance and its determinant factors over a season in young swimmers. *International Journal of Sports Physiology and Performance, 10*(2), 261–268. doi:10.1123/ijsspp.2013-0533

Moreira, M. F., Morais, J. E., Marinho, D. A., Silva, A. J., Barbosa, T. M., & Costa, M. J. (2014). Growth influences biomechanical profile of talented swimmers during the summer break. *Sports Biomechanics, 13*(1), 62–74. doi:10.1080/14763141.2013.865139

Rein, R., Button, C., Davids, K., & Summers, J. (2010). Cluster analysis of movement patterns in multiarticular actions: A tutorial. *Motor Control, 14*(2), 211–239. doi:10.1123/mc.14.2.211

Saavedra, J. M., Escalante, Y., & Rodriguez, F. A. (2010). A multivariate analysis of performance in young swimmers. *Pediatric Exercise Science, 22*(1), 135–151. doi:10.1123/pes.22.1.135

Sammoud, S., Nevill, A. M., Negra, Y., Bouguezzi, R., Helmi, C., & Hachana, Y. (2019). Key somatic variables in young backstroke swimmers. *Journal of Sports Sciences, 37*(10), 1162–1167.

Silva, A. F., Figueiredo, P., Morais, S., Vilas-Boas, J. P., Fernandes, R. J., & Seifert, L. (2019b). Task constraints and coordinate flexibility in young swimmers. *Motor Control, 23*, 535–552. [Epub ahead of print]. doi:10.1123/mc.2018-0070

Silva, A. F., Figueiredo, P., Ribeiro, J., Alves, F., Vilas-Boas, J. P., Seifert, L., & Fernandes, R. J. (2019a). Integrated analysis of young swimmers’ sprint performance. *Motor Control, 23*(3), 354–364. doi:10.1123/mc.2018-0014

Zacca, R., Azevedo, R., Chainok, P., Vilas-Boas, J. P., Castro, F. A. S., Pyne, D. B., & Fernandes, R. J. (2018). Monitoring age-group swimmers over a training macrocycle: Energetics, technique, and anthropometrics. *Journal of Strength and Conditioning Research, 1*. [Epub ahead of print]. doi:10.1519/JSC.0000000000002762

Zacca, R., Azevedo, R., Ramos, V. R., Jr., Abraldes, J. A., Vilas-Boas, J. P., Castro, F. A. S., ... Fernandes, R. J. (2019a). Biophysical follow-up of age-group swimmers during a traditional three-peak preparation program. *Journal of Strength and Conditioning Research, 1*. [Epub ahead of print]. doi:10.1519/JSC.0000000000002964

Zacca, R., Toubekis, A., Freitas, L., Silva, A. F., Azevedo, R., Vilas-Boas, J. P., ... Fernandes, R. J. (2019b). Effects of detraining in age-group swimmers’ performance, energetics and kinematics. *Journal of Sports Sciences, 37*(13), 1490–1498. doi:10.1080/02660414.2019.1572434

Zamparo, P., Pendergast, D. R., Mollendorf, J., Termin, A., & Minetti, A. E. (2005). An energy balance of front crawl. *European Journal of Applied Physiology, 94*(1–2), 134–144. doi:10.1007/s00421-004-1281-4