B4-2 ID15

LONGITUDINAL COMPARISON OF YOUNG SWIMMERS' ANTHROPOMETRICS, HYDRODYNAMICS, KINEMATICS AND EFFICIENCY BY CLUSTER ANALYSIS

Jorge E Morais[1,5], Pedro Forte[2], Mário J Costa[5], Daniel A Marinho[3,5], António J Silva[1,5], Tiago M Barbosa[4,5]

[1]University of Trás-os-Montes and Alto Douro, Vila Real, Portugal
[2]Polytechnic Institute of Bragança, Bragança, Portugal
[3]University of Beira Interior, Covilhã, Portugal
[4]National Institute of Education, Nanyang Technological University, Singapore
[5]Research Centre in Sports, Health and Human Development, Vila Real, Portugal

The aim of this study was to identify the variables that better discriminate young swimmers' biomechanical profile during a competitive season by a cluster analysis. Fifteen boys and eighteen girls were evaluated three times throughout a competitive season. Arm span, chest perimeter, stroke length, velocity, speed fluctuation, coefficient of active drag, propelling efficiency and stroke index were selected as variables. Cluster and discriminant analysis were computed, and MANOVA used to verify the gender and performance effects. Swimmers' classification is mainly determined by anthropometric, kinematic/efficiency and hydrodynamic features. Throughout the season the changes in the clustering solution suggests moderate-high stability in their biomechanical profile.

INTRODUCTION: Research about young swimmers' biomechanical profile is scarce in comparison to the body of knowledge regarding adult/elite counterparts. A longitudinal research design allows understanding which domains and variables are more determinants for the performance enhancement in different moments of a competitive season. One way to assess the stability of the swimmers' biomechanical profile throughout a time period is their classification based on clustering solutions. Such approach allows verifying if the swimmer remains or changes from cluster group in each evaluation moment. To the best of our knowledge this research design was never attempted in competitive swimming. The aim of this study was to identify the variables that better discriminate young swimmers' biomechanical profile in three different time points (TP) during a competitive season and assess its stability.

METHODS: 33 young swimmers (15 boys and 18 girls: 11.81 ± 0.75 years old and Tanner stages 1-2 by self-evaluation) participating on regular basis in regional and national level competitions were assessed. Swimmers were evaluated in three different TP's: (i) October (TP1), close to the season's first competition; (ii) March (TP2), xlose to the winter peak competition and; (iii): June (TP3), close to the summer peak competition.
Swimming performance was taken from the 100 m freestyle time lists event of official short course (i.e. 25 m swimming pool) competition of regional or national level. The time gap between data collection and swimming performance was made in less than two weeks.
Anthropometric measurements were based on arm span (AS) and chest perimeter (CP) assessments measured with a flexible anthropometric tape (RossCraft, Canada). Swimmers were in in the upright orthostatic for the AS simulated the hydrodynamic position for the CP measurements.
For kinematic assessment each swimmer performed three freestyle swim trials of 25 m with underwater start. For further analysis the average value of the three trials was computed. A speedo-meter cable (Swim speedo-meter, Swimsportec, Hildesheim, Germany) was used to collect swimming velocity (v), stroke frequency (SF) and speed fluctuation (dv) during the middle 15 m. Stroke length (SL) was measured as SL=v/SF.
In the hydrodynamic domain, coefficient of active drag (C_{Da}) was assessed with the velocity perturbation method (Kolmogorov & Duplishcheva, 1992). Each swimmer performed two

maximal 25 m freestyle trials (with and without the perturbation object). Stroke index (SI) was computed as SI= SL·v and propelling efficiency (η_p) as reported by Zamparo et al. (2005). Two clustering approaches were used: (i) a hierarchical cluster analysis using Ward's linkage method with the squared Euclidian distance measure; (ii) a k-Means (non-hierarchical) cluster analysis. It was used standardized z-scores of the selected variables in the clustering analysis. To identify the variables with highest influence in each cluster, cluster's ANOVA (including total eta square) and discriminant analysis (stepwise method) tests were computed ($P < 0.05$). MANOVA using cluster group as the independent variable and swimmers' characteristics (i.e. gender and swimming performance) and Bonferroni post-hoc test to verify differences between each cluster in the different TP's were also tested ($P < 0.05$).

RESULTS: Table 1 presents the swimmers' classification computed with k-Means cluster method (k = 3) for TP1, TP2 and TP3, respectively. ANOVA statistics revealed significant variations in TP1 for almost all variables ($P \leq 0.001$), except for dv (P = 0.25). In TP2 there were significant variations for AS, CP, SL, v, η_p, SI ($P < 0.001$) and C_{Da} (P = 0.03), but not for dv (P = 0.16). In TP3 there were significant variations for AS, SL, v, η_p, SI ($P < 0.001$), CP (P = 0.001) and C_{Da} (P = 0.002), but not for dv (P = 0.95) once again.

Table 1. Descriptive and ANOVA statistics by clustering in TP1, TP2 and TP3.

	Cluster 1 (n = 15)		Cluster 2 (n = 7)		Cluster 3 (n =11)				
TP1	**Mean ± 1SD**	**z**	**Mean ± 1SD**	**z**	**Mean ± 1SD**	**z**	**F**	**P**	**η^2**
AS [cm]	165.26 ± 9.91	0.66	152.00 ± 7.70	- 0.54	151.81 ± 7.98	- 0.56	9.27	0.001	0.40
CP [cm]	82.55 ± 4.32	0.74	73.82 ± 3.94	- 0.75	75.07 ± 4.57	- 0.54	13.98	< 0.001	0.50
SL [m]	1.68 ± 0.17	0.76	1.13 ± 0.21	- 1.27	1.41 ± 0.12	- 0.23	27.15	< 0.001	0.66
v [m·s^{-1}]	1.37 ± 0.12	0.73	0.88 ± 0.14	- 1.50	1.20 ± 0.08	- 0.04	43.31	< 0.001	0.76
dv [dimensionless]	0.10 ± 0.03	0.02	0.11 ± 0.04	0.48	0.08 ± 0.02	- 0.33	1.47	0.25	0.10
C_{Da} [dimensionless]	0.30 ± 0.08	- 0.27	0.23 ± 0.09	- 0.72	0.48 ± 0.19	0.82	9.02	0.001	0.38
η_p [%]	30.55 ± 2.77	0.65	21.95 ± 4.48	- 1.25	27.24 ± 2.55	- 0.08	18.11	< 0.001	0.57
SI [m^2·c^{-1}·s^{-1}]	2.31 ± 0.36	0.81	1.02 ± 0.33	- 1.37	1.70 ± 0.22	- 0.22	41.87	< 0.001	0.75
TP2	**Cluster 1 (n = 7)**		**Cluster 2 (n = 8)**		**Cluster 3 (n = 18)**				
AS [cm]	173.00 ± 9.73	1.18	159.75 ± 9.05	- 0.08	156.16 ± 7.40	- 0.42	10.39	< 0.001	0.44
CP [cm]	87.78 ± 3.66	1.22	79.50 ± 5.42	- 0.15	77.93 ± 4.73	- 0.41	11.19	< 0.001	0.29
SL [m]	1.16 ± 0.09	- 0.03	1.45 ± 0.11	1.44	1.04 ± 0.10	- 0.63	44.09	< 0.001	0.76
v [m·s^{-1}]	0.94 ± 0.09	- 0.21	1.30 ± 0.13	1.54	0.87 ± 0.07	- 0.60	61.25	< 0.001	0.81
dv [dimensionless]	0.13 ± 0.06	0.63	0.10 ± 0.02	- 0.09	0.09 ± 0.03	- 0.21	1.92	0.16	0.12
C_{Da} [dimensionless]	0.33 ± 0.05	0.27	0.38 ± 0.15	0.67	0.25 ± 0.09	- 0.40	4.20	0.03	0.24
η_p [%]	19.92 ± 1.30	- 0.48	27.84 ± 2.95	1.48	19.99 ± 2.05	- 0.47	39.26	< 0.001	0.73
SI [m^2·c^{-1}·s^{-1}]	1.10 ± 0.15	- 0.22	1.96 ± 0.34	1.55	0.92 ± 0.16	- 0.60	66.76	< 0.001	0.83
TP3	**Cluster 1 (n = 11)**		**Cluster 2 (n = 4)**		**Cluster 3 (n = 18)**				
AS [cm]	162.00 ± 7.40	- 0.03	180.75 ± 5.12	1.81	158.43 ± 7.80	- 0.38	14.74	< 0.001	0.59
CP [cm]	82.78 ± 5.59	0.11	91.00 ± 2.16	1.46	79.67 ± 5.05	- 0.39	8.44	0.001	0.45
SL [m]	1.20 ± 0.12	- 0.98	1.75 ± 0.14	1.40	1.49 ± 0.16	0.29	25.19	< 0.001	0.64
v [m·s^{-1}]	0.99 ± 0.10	- 1.12	1.54 ± 0.05	1.54	1.29 ± 0.10	0.34	58.31	< 0.001	0.81
dv [dimensionless]	0.10 ± 0.02	0.06	0.09 ± 0.01	0.05	0.09 ± 0.03	- 0.05	0.05	0.95	0.03
C_{Da} [dimensionless]	0.32 ± 0.12	- 0.16	0.64 ± 0.36	1.55	0.30 ± 0.11	- 0.24	7.86	0.002	0.42
η_p [%]	22.49 ± 1.97	- 1.15	29.07 ± 1.18	0.56	29.15 ± 2.46	0.58	33.14	< 0.001	0.69
SI [m^2·c^{-1}·s^{-1}]	1.20 ± 0.22	- 1.08	2.70 ± 0.18	1.69	1.94 ± 0.26	0.28	64.09	< 0.001	0.83

In TP1, cluster 1 was related to a high CP, SI, SL, cluster 2 was related to a high dv and cluster 3 to a high C_{Da}. The variables that better discriminate the clusters in TP1 were the v (F = 43.31; $P < 0.001$), the SI (F = 41.87; $P < 0.001$), and the SL (F = 27.15; $P < 0.001$). MANOVA showed non-significant multivariate gender effect (F = 2.082; P = 0.142), but a significant one for the swimming performance (F = 7.018; P = 0.003) in cluster groups ($\Lambda_{Wilk's}$ = 0.625, P = 0.008; $\Lambda_{Pillai's}$ = 0.378, P = 0.012). Bonferroni pairwise comparisons showed significant differences ($P = 0.05$) between cluster 1 and 2. In TP2, cluster 1 was related to a high CP and AS, cluster 2 was related to a high SI, v and η_p and cluster 3 was related to a high dv and C_{Da}. As for TP1, the variables that better discriminate the clusters in TP2 were the SI ($F = 66.76$; $P < 0.001$), the v ($F = 61.25$; $P < 0.001$) and the SL ($F = 44.09$; $P < 0.001$). MANOVA showed non-significant multivariate gender effect ($F = 1.171$; $P = 0.324$), but a significant one for the swimming performance ($F = 7.344$; $P = 0.003$) in cluster groups ($\Lambda_{Wilk's} = 0.659$, $P = 0.015$; $\Lambda_{Pillai's} = 0.346$, $P = 0.021$). Bonferroni pairwise comparisons showed significant differences ($P = 0.05$) between cluster 1 and 2. In TP3, cluster 1 was related to a high CP and dv. Cluster 2 was related to a high AS and SI. And cluster 3 was related to a high η_p and v. The variables that better discriminate the clusters in TP3 were the SI ($F = 64.09$; $P < 0.001$), the v ($F = 58.31$; $P < 0.001$) as happened for TP1 and TP2, plus the η_p ($F = 33.14$; $P < 0.001$). MANOVA showed significant multivariate gender ($F = 3.521$; $P = 0.042$) and swimming performance ($F = 9.449$; $P = 0.001$) effects, in cluster groups ($\Lambda_{Wilk's} = 0.582$, $P = 0.003$; $\Lambda_{Pillai's} = 0.431$, $P = 0.005$). For gender, Bonferroni pairwise comparisons showed significant differences ($P = 0.05$) between cluster 2 and 3. As for performance, there were significant differences between cluster 1 and 2; and cluster 2 and 3. So, cluster 1 can be tagged as “anthropometrics”, cluster 2 “efficiency” and cluster 3 “hydrodynamics” with a performance but not a gender effect on it.

Stepwise discriminant analysis, for TP1, extracted 2 functions including CP, v and C_{Da} (fig. 1A). Function 1 is mainly defined by v and CP, explaining 69.3 % of variance ($\Lambda = 0.105$; $X^2(6) = 65.22$; $P < 0.001$). Function 2 is mainly defined by C_{Da}, explaining 30.7 % of variance ($\Lambda = 0.426$; $X^2(2) = 24.75$; $P < 0.001$). Classification functions (90.9 % of original grouped correctly classified) were: Cluster 1 = $4.438 \cdot CP + 59.136 \cdot v - 41.329 \cdot C_{Da} - 218.227$; Cluster 2 = $4.241 \cdot CP + 24.009 \cdot v - 41.61 \cdot C_{Da} - 163.88$; Cluster 3 = $3.894 \cdot CP + 51.29 \cdot v - 24.039 \cdot C_{Da} - 172.327$. Stepwise discriminant analysis, for TP2, extracted 2 functions including AS and SI (fig. 1B). Function 1 is mainly defined by SI, explaining 89.2 % of variance ($\Lambda = 0.089$; $X^2(4) = 71.37$; $P < 0.001$). Function 2 is mainly defined by AS explaining 10.8 % of variance ($\Lambda = 0.593$; $X^2(1) = 15.41$; $P < 0.001$). Classification functions (87.9 % of original grouped correctly classified) were: Cluster 1 = $2.82 \cdot AS - 26.602 \cdot SI - 230.789$; Cluster 2 = $2.571 \cdot AS - 26.152 \cdot SI - 189.336$; Cluster 3 = $2.292 \cdot AS + 1.608 \cdot SI - 186.065$. Stepwise discriminant analysis, for TP3, extracted 2 functions including AS, SL, v, C_{Da} and SI (fig. 1C). Function 1 is mainly defined by SI, v and SL, explaining 65.4 % of variance ($\Lambda = 0.03$; $X^2(10) = 95.81$; $P < 0.001$). Function 2 is mainly defined by AS and C_{Da} explaining 34.6 % of variance ($\Lambda = 0.23$; $X^2(4) = 40.63$; $P < 0.001$). Classification functions (100 % of original grouped correctly classified) were: Cluster 1 = $3.81 \cdot AS + 2285.71 \cdot SL + 2387.87 \cdot v - 80.711 \cdot CDa - 1842.30 \cdot SI - 1758.55$; Cluster 2 = $4.08 \cdot AS + 2108.83 \cdot SL + 2202.50 \cdot v - 59.25 \cdot CDa - 1694.27 \cdot SI - 1609.38$; Cluster 3 = $4.31 \cdot AS + 2264.28 \cdot SL + 2328.76 \cdot v - 69.32 \cdot CDa - 1836.37 \cdot SI - 1756.15$

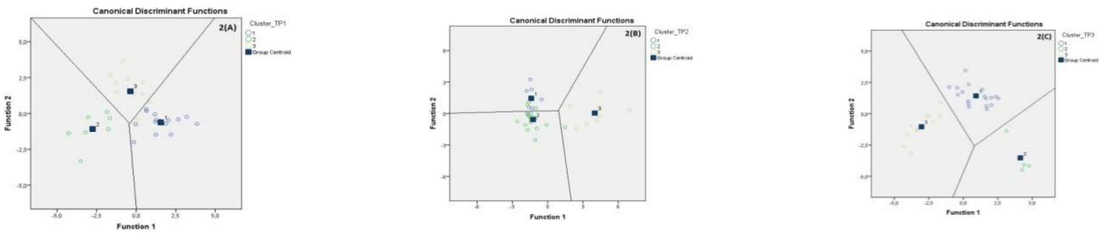

Figure 1: Territorial map of the two canonical discriminant functions in TP1 (1A), TP2 (1B) and TP3 (1C), respectively.

Table 2 presents the cluster membership at TP2 and TP3, tabulated against cluster membership at TP1. In the 3 TPs, cluster 3 (i.e. hydrodynamics) presented the highest stability ranging between 50 % (TP2 vs TP3) and 81.8 % (TP1 vs TP3), followed by cluster 1 (i.e. anthropometrics) ranging from 28.6 % (TP2 vs TP3) to 47 % (TP1 vs TP2). Cluster 2 (i.e. efficiency) showed the lowest stability ranging from 0 % (TP1 vs TP3) to 28.6 % (TP1 vs TP2). Overall, it seems to exist moderate-high stability in the clustering membership.

DISCUSSION: In all 3 TP's the variables that better discriminate the clustering solutions were mainly the v, SI and SL. These variables are reported in the literature as highly correlated and/or with direct effect with swimming performance in children (Barbosa et al., 2010a). There was a non-significant gender effect in TP1 and TP2. However, a gender effect was verified between cluster 2 and cluster 3 in TP3. This difference was related to anthropometric variables, meaning that biological maturation starts to play a role. Performance had a significant effect in all 3 TP's. Swimmers with high kinematic skills (in the 3 TP's) are the fastest ones, as these variables were those that better discriminated the clustering solutions. Stepwise discriminant analysis extracted 2 functions in all 3 TP's. The variance reported by these functions, was mainly explained by the "anthropometrics" and "kinematics" once again, as it happened for the cluster analysis. In TP 1, the 2 functions included v, CP and C_{Da} explaining 69.3 % and 30.7 % of variance, respectively. In TP2, included SI and AS, explaining 89.2 % and 10.8 % of variance, respectively. And in TP3, included SI, v, SL and AS and C_{Da} explaining 65.4 % and 34.6 % of variance, respectively. Regarding cluster and discriminant analysis, young swimmers can be classified according to their "anthropometric", "kinematic" and "hydrodynamic" characteristics. As for the cluster membership, the highest changes in TP1 vs TP2 and TP2 vs TP3 were due to an improvement in the hydrodynamic position. In TP1 vs TP3, changes were due to anthropometric variables. Young swimmers change their kinematic pattern during a competitive season, and this might be due to their anthropometric growth and development processes (Lätt et al., 2009). Data from this research suggests that there is a wide intra-individual stability (i.e., between moderate to high) since the cluster membership throughout the season changes for a large part of the swimmers assessed.

Table 2. Cross-tabulations between cluster membership at different TP's.

	Cluster 1		Cluster 2		Cluster 3	
TP1 vs TP2	**n**	**%**	**n**	**%**	**n**	**%**
Cluster 1	7	47	0	0	0	0
Cluster 2	3	20	2	28.6	3	27.3
Cluster 3	5	33	5	71.4	8	72.7
TP2 vs TP3						
Cluster 1	2	28.6	0	0	9	50
Cluster 2	3	42.8	1	12.5	0	0
Cluster 3	2	28.6	7	87.5	9	50
TP1 vs TP3						
Cluster 1	5	33.3	4	57.2	2	18.2
Cluster 2	4	26.7	0	0	0	0
Cluster 3	6	40	3	42.8	9	81.8

CONCLUSION: Young swimmers' classification is mainly determined by anthropometric, kinematic/efficiency and hydrodynamic features. Through the season their changes in the clustering solution suggests moderate-high stability.

REFERENCES:

Barbosa, T.M., Costa, M., Marinho, D.A., Coelho, J., Moreira, M., Silva, A.J. (2010a). Modeling the links between young swimmer´s performance: energetic and biomechanical profiles. *Pediatric Exercise Science*, 22, 379-391.

Kolmogorov, S. & Duplishcheva, O. (1992). Active drag, useful mechanical power output and hydrodynamic force in different swimming strokes at maximal velocity. *Journal of Biomechanics*, 25, 311-318.

Lätt, E., Jürimäe, J., Haljaste, K., Cicchela, A., Purge, P., & Jürimäe, T. (2009). Physical development and swimming performance during biological maturation in young female swimmers. *Collegium Antropologicum*, 33, 117-122.

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, 134–144.