

# The Explanatory Capacity of Talent Identification Tests for Performance in Triathlon Competitions: A Longitudinal Analysis

by

Alba Cuba-Dorado<sup>1</sup>, Oscar Garcia-Garcia<sup>1</sup>, Verónica Morales-Sánchez<sup>2</sup>,  
Antonio Hernández-Mendo<sup>2</sup>

*The aim of this study was to determine the explanatory capacity of the Spanish Triathlon Federation's talent identification tests in relation to performance in competition in subsequent years. We used an exploratory longitudinal study design to establish the relationship between talent identification tests completed by 247 triathletes (97 women and 150 men) aged from 14 to 19 years and the results they obtained over the years in competition. The battery of tests included freestyle swimming (100 and 1000 m) and running (400 and 1000 m). The results indicate that the explanatory capacity of these tests for split places in competition in the corresponding discipline was highest in the 1000-m swimming test, with a value of 0.34 for the adjusted coefficient of determination ( $R^2a$ ) ( $p \leq 0.001$ ), followed by the 1000-m running and 100-m swimming tests, where the highest  $R^2a$  values were 0.26 and 0.19, respectively. No significant model was found for the 400-m running test. It was concluded that the explanatory capacity of the tests analysed for predicting performance in the discipline in competition was low. However, it was higher for the swimming and running tests of longer distance.*

**Key words:** talent identification; sports performance; triathlon performance; endurance sports; youth sports; exercise tests.

## Introduction

Triathlon is a sport that combines the disciplines of swimming, cycling and running, performed consecutively in that order and timed continuously from start to finish without stopping the clock. Between disciplines there is a period known as transition, in which triathletes change their clothing and equipment for the next segment. There are various competition formats; those used in the main world circuit (World Triathlon Series) and in the Olympic Games are sprint (750 m of swimming, 20 km of cycling, 5 km of running) and standard (1500 m, 40 km, 10 km).

As with other endurance sports such as long-distance running and cycling, success in triathlon is related to physiological characteristics

of athletes, especially considering high levels in maximum values of oxygen uptake ( $VO_{2max}$ ) and peak power output in cycling ( $W_{peak}$ ). Similar values to other endurance athletes are also typically observed in submaximal physiological variables, in measurements of the ventilatory threshold (VT) and lactate threshold (LT), although it is difficult to make comparisons due to the multiple definitions that exist (Suriano and Bishop, 2010). Other physiological factors, less studied yet seemingly related to performance, are lactic acid buffering capacity (Baldari et al., 2007) and economy of movement (Millet et al., 2011).

Moreover, some anthropometric variables have been shown to be related to success in competition, such as a low body fat percentage (related to overall performance in competition)

<sup>1</sup> - Sports Performance, physical condition and wellness laboratory. University of Vigo, Pontevedra (Spain).

<sup>2</sup> - Department of Social Psychology, Faculty of Psychology, University of Málaga (Spain).

and long limbs (arms are related to performance in swimming and legs to performance in running and cycling) (Ackland et al., 1998; Landers et al., 2013). Mechanical factors, such as technical capacity in each discipline (Bentley et al., 2002), and tactical factors, such as competition strategy (Bentley et al., 2002; Hausswirth et al., 2010; Vleck et al., 2008), will also have influence on triathlon performance. In order to analyse these factors, it is important to consider fatigue from the previous disciplines, since behaviour will be different if analysed in isolation (Ofoghi et al., 2016).

The popularisation of triathlon in recent years has led several sports institutions to start investing significant resources in the talent identification process, in order to find young males and females with the potential to become elite athletes in adulthood (Swann et al., 2015). Some institutions, such as the Spanish Triathlon Federation (FETRI), use these programmes to provide the most talented athletes with appropriate training (including greater monitoring through national team training camps and competitions and even scholarships at high performance centres) to enable them to join the Spanish elite team, ensuring high competitiveness in international events (Federación Española de Triatlón, 2008). To achieve this goal it is critically important to evaluate the validity of the models applied and the success rates of the programmes carried out (Vaeyens et al., 2008).

However, there is a lack of scientific research on talent identification processes in sports in general (Johnston et al., 2018), and especially in triathlon. Bottoni et al. (2011) tried to identify factors to be considered in talent identification programmes, suggesting that they should be based not only on measuring physical parameters, but should also take account of other variables. Cuba-Dorado et al. (2015) analysed the talent identification tests developed by the Spanish Triathlon Federation (FETRI), which has one of the highest ITU country rankings worldwide. They found that the tests had very limited explanatory capacity for overall performance in competition in the year the tests were conducted.

This suggests that the tests carried out to choose the most promising Spanish triathletes do not predict their performance. Nevertheless, it seems necessary to check these findings in a

longitudinal study, which can help avoid biases in talent identification, reducing the risk of prematurely attributing causality to the results (Johnston et al., 2018). The objective of this study, therefore, was to determine the explanatory capacity of talent identification tests carried out by the FETRI in relation to performance in competition in subsequent years in the discipline corresponding to each test.

## **Methods**

### *Participants*

We obtained the data for all triathletes aged from 14 to 19 years between 2009 and 2015 who underwent the talent identification tests developed by the FETRI. The study included triathletes who obtained a minimum of 8 points on the FETRI's scale (explained in the investigation design) in at least 3 of the 4 tests, showing a good level of performance. 247 triathletes met this criterion: 150 men (60.7%) and 97 women (39.3%). Many of these triathletes subsequently participated in the Spanish Championships in their categories in the year when the tests were carried out and in the following years. In the year of the tests (Year 1), 200 of triathletes tested took part in the Spanish Championships; the next year (Year 2), the figure was 167, followed by 119 in Year 3, 71 in Year 4, 40 in Year 5, 29 in Year 6, 14 in Year 7, and 11 in Year 8, seven years after the tests were completed. The research protocol followed the principles of the Helsinki Declaration on biomedical research involving human subjects (64<sup>th</sup> World Medical Assembly, 2013). The study was approved by the Ethics Committee of the University of Vigo.

### *Measures, Design and Procedures*

An exploratory longitudinal study design with an associative strategy was used to establish the relationship between the variables obtained in the Spanish Triathlon Federation's tests and the split results obtained over the years in the various segments of the Spanish Triathlon Championships. We performed an independent analysis for each of the FETRI tests: S100 (100 m of freestyle swimming), S1000 (1000 m of freestyle swimming), R400 (400 m of running) and R1000 (1000 m of running). Swimming tests were conducted in a 25-m pool and running tests on a 400-m athletics track.

Predictive variables were used in each of

these analyses:

- **Points:** points obtained in each test according to the FETRI's scale. This scale establishes a score according to the time reached in each test, taking into account age and sex of the triathlete. The minimum score is 0 points and the maximum is 10 for triathletes aged 18 or 19, 11 for those aged 16 or 17 and 12 for those aged 15.
- **Time:** time of the test, measured in seconds.
- **Best Proportion of All (All):** time percentage for the triathlete who achieved the best time in the test relative to all participants of the same age and sex between 2007 and 2016.
- **Best Proportion of 4 Results (4R):** time percentage for the triathlete who achieved the best time in the test relative to all participants of the same age and sex between 2007 and 2016, taking into account only athletes who scored 8 or more points in at least three out of the four talent identification tests.

We used the place percentage obtained by the triathlete in the discipline corresponding to the test analysed, relative to the number of participants in the Spanish Triathlon Championship, as a dependent or response variable. In this percentage a low number indicates a result in the highest places, whereas 100 signifies that the athlete finished in the last place. We analysed this response variable over the following eight years, taking Year 1 (SCYear1) as the year in which the tests were carried out, Year 2 (SCYear2) as the next year, and so on. The championships in the youth, junior, under-23 and elite categories were also considered (since the last two categories include triathletes competing in the same competition, the result that was more positive for the triathlete was selected).

We collected the data for FETRI's tests and the Spanish Championship results from 2009, 2010, 2011, 2012, 2013, 2014, and 2015, verifying that triathletes fulfilled the criterion of a minimum of 8 points in at least three tests. Once the sample had been selected, we collected the triathletes' results in the Spanish Championships in their categories from the year of the test and the following years. The competition split place

percentage (relative to the number of participants) was calculated for the discipline corresponding to each test.

#### **Statistical analysis**

In order to verify that the data were reliable (Gr) and generalisable (Ga) a variance component analysis was implemented, using the least squares and maximum likelihood estimation procedures ( $p < 0.01$ ) and generalisability analysis (Gr > 0.800; Ga > 0.800). Generalisability theory (G theory) is a theory of the multifaceted errors of measurement; it assumes that any measurement situation contains infinite sources of variation. Generalisability analysis combines the concepts of reliability, validity and precision (Blanco-Villaseñor et al., 2014). We then developed a multiple linear regression analysis for each of the 8 years of Spanish Championships. We used the statistical packages SAS System for Windows v. 9.1. (SAS Institute Inc., Cary, NC, USA), SAGT v.1. (University of Málaga, Spain) and SPSS v19.0 for Windows (SPSS Inc., Chicago, IL, USA), respectively.

#### **Results**

In all generalisability analysis models the residual error coincided in the least-squares (VARCOMP Type I) and maximum likelihood (GLM General Linear Model) estimation procedures, as it was confirmed that the sample distribution was normal, linear and homoscedastic (Searle et al., 1992). The precision models of the S100, S1000 and R1000 tests showed high levels of Gr (between 0.99 and 1) and Ga (between 0.93 and 0.99), S1000 being the highest value that the two coefficients could take. However, R400 delivered no result (no comma) with any of the estimation procedures used. The degree of independence was calculated using the Durbin-Watson (D-W) test, in which it could be assumed that a value located between 1.5 and 2.5 showed independence. The values of all the models lay within this range, as can be observed in Tables 1, 2 and 3.

The results of the significant model with the independent variables that showed the greatest explanatory capacity in each of the years analysed are presented below. Those for the R1000 test are shown in Table 1, those for the S1000 test in Table 2 and those for the S100 test in Table 3. No significant model was found for the

R400 test.

In Table 1 we can see that the greatest explanatory capacity of the 1000-m running test for split places in running was obtained in the year of the tests. The value of the adjusted coefficient of determination ( $R^2_a$ ) ( $p \leq 0.001$ ) allowed us to state that 26.2% of the total variance of the run split place in the Spanish Championship was explained by the variables All and Points in the linear regression model, obtaining a moderate model adjustment (Figure 1). It can be expressed as follows:

$$\text{SCYear1} = 184.060 - 4.207 \cdot \text{Points} - 1.347 \cdot \text{All}$$

In subsequent years this explanatory capacity decreased significantly.

In Table 2 we can see that the greatest explanatory capacity of the 1000-m swimming test for split places in swimming was obtained one year after the tests. The value of  $R^2_a$  ( $p \leq 0.001$ ) allowed us to state that 34.0% of the total variance of the swim split place in the Spanish

Championship was explained by the variables 4R and Points in the linear regression model, obtaining a moderate model adjustment (Figure 2). It can be expressed as follows:

$$\text{SCYear2} = 135.779 - 0.804 \cdot \text{Points} - 1.226 \cdot 4R$$

In subsequent years this explanatory capacity decreased significantly.

In Table 3 we can see that the greatest explanatory capacity of the 100-m swimming test for split places in swimming was obtained four years after the tests. The value of  $R^2_a$  ( $p = 0.007$ ) allowed us to state that 19.5% of the total variance of the swim split place in the Spanish Championship was explained by the variables 4R and Time in the linear regression model, obtaining a moderate model adjustment (Figure 3). It can be expressed as follows:

$$\text{SCYear5} = 299.743 - 2.372 \cdot 4R - 0.962 \cdot \text{Time}$$

Unlike the previous tests, no significant model was found in the third year after this test.

**Table 1**

*Summary of linear regression models for the 1000 m running test*

	Predictors	R <sup>2</sup>	R <sup>2</sup> <sub>a</sub>	R	Error	Sig	D-W
SCYear1	All and Points	0.270	0.262	0.520	18.58	.000	1.94
SCYear2	4R, Time and All	0.128	0.111	0.358	21.18	.000	2.18
SCYear3	4R, Time and Points	0.228	0.206	0.477	19.63	.000	1.99
SCYear4	Points	0.109	0.094	0.330	20.39	.009	1.85
SCYear5							
SCYear6							
SCYear7							
SCYear8							

No significant model was found

**Table 2**

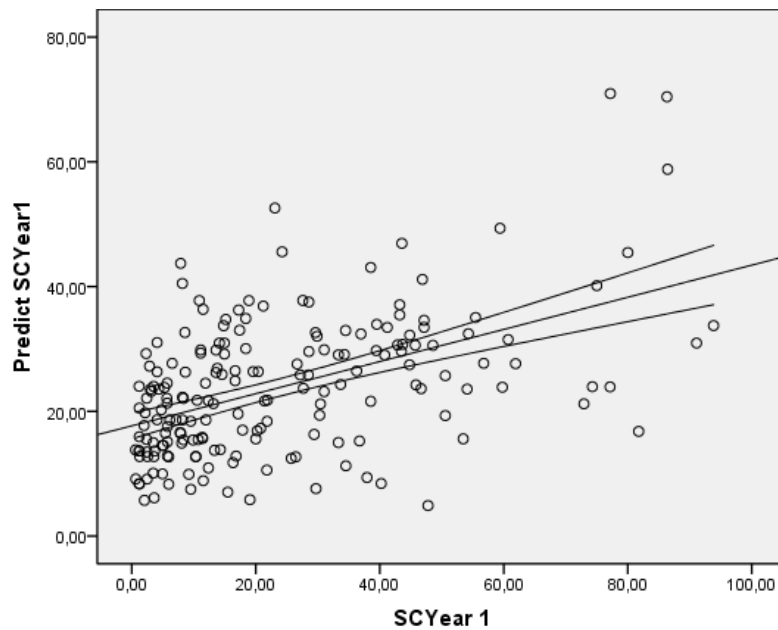
*Summary of linear regression models for the 1000 m swimming test*

	Predictors	R <sup>2</sup>	R <sup>2</sup> <sub>a</sub>	R	Error	Sig	D-W
SCYear1	4R and Points	0.345	0.339	0.588	10.91	.000	1.91
SCYear2	4R and Points	0.348	0.340	0.590	12.44	.000	2.18
SCYear3	All, 4R and Points	0.199	0.179	0.447	14.52	.000	1.87
SCYear4	Time	0.094	0.081	0.307	14.62	.009	1.86
SCYear5							
SCYear6							
SCYear7							
SCYear8							

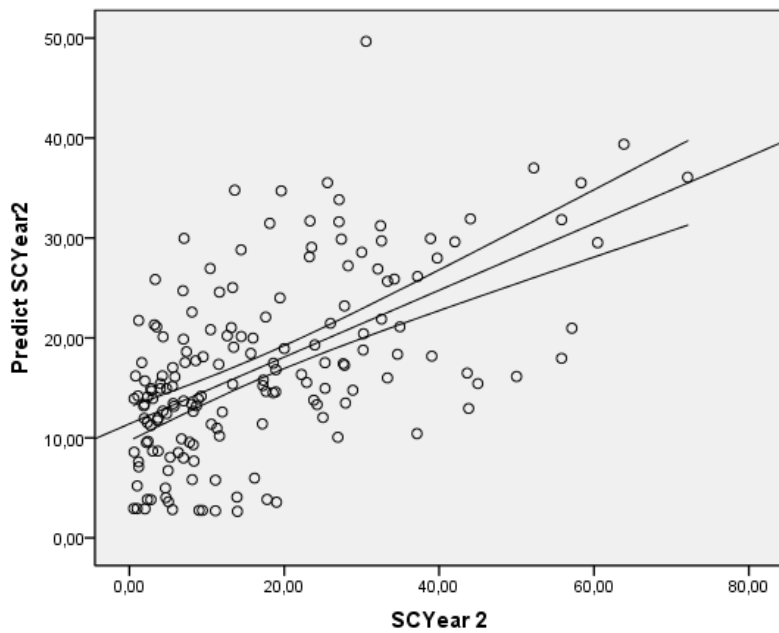
No significant model was found

**Table 3**  
*Summary of linear regression models for the 100 m swimming test*

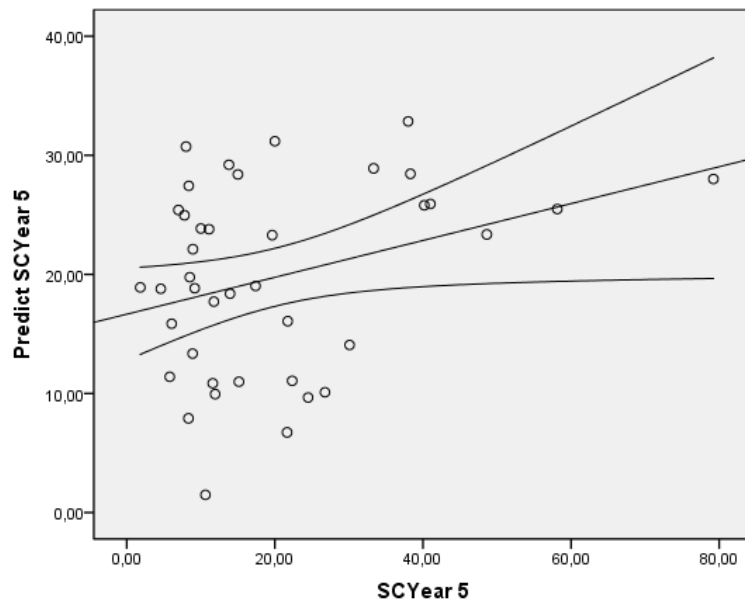
	Predictors	R <sup>2</sup>	R <sup>2</sup> <sub>a</sub>	R	Error	Sig	D-W
SCYear1	4R and Time	0.108	0.099	0.328	12.74	.000	1.79
SCYear2	All and Time	0.128	0.117	0.358	14.39	.000	2.01
SCYear3	4R	0.051	0.043	0.225	15.68	.014	1.99
SCYear4	No significant model was found						
SCYear5	4R and Time	0.237	0.195	0.486	14.68	.007	2.18
SCYear6	No significant model was found						
SCYear7	No significant model was found						
SCYear8	No significant model was found						



**Figure 1**  
 Scatter plot (SCYear 1 vs. Predicted SCYear1) for the 1000 m running test with a linear trend, the 95% confidential interval for the linear trend

**Figure 2**

*Scatter plot (SCYear 2 vs. Predicted SCYear 2) for the 1000 m swimming test with a linear trend, the 95% confidential interval for the linear trend*

**Figure 3**

*Scatter plot (SCYear 5 vs. Predicted SCYear 5) for the 100 m swimming test with a linear trend, the 95% confidential interval for the linear trend*

## Discussion

The main finding of this study is that the explanatory capacity of the tests used by the FETRI in their talent identification process for split performance in triathlon competitions is low (1000 m running test: 26.2%, 1000 m swimming test: 34%, 100 m swimming test: 19.5%).

The results also indicate that the explanatory capacity is related to the length of the test (the most explanatory test was S1000, followed by R1000 and S100, whereas the explanatory capacity of R400 was zero). These results are in line with those obtained by Cuba-Dorado et al. (2015), who analysed the explanatory capacity of the same tests for the overall results in triathlon competitions in the year of the tests.

Explanatory capacity seems to be related to energy demands: the 1000-m swimming test, which has the greatest explanatory capacity, is considered a long-term effort, in which the aerobic pathway is the most used. The 1000-m running test, which has the second highest explanatory capacity, constitutes a medium-term effort, with a balanced use of the aerobic and anaerobic pathways. The other tests, i.e., 100 m swimming and 400 m running, are regarded as short-term efforts with predominance of the anaerobic over the aerobic pathway (Capelli et al., 1998; Gastin, 2001; Spencer and Gastin, 2001).

A reasonable explanation of these findings could be a relationship between predictive capacity and the predominant pathway in the test, which is related, in turn, to physiological performance factors associated with endurance sports. Thus, the S1000 test would be associated with  $VO_{2max}$ ,  $VT_2$ , lactic acid buffering capacity and economy of movement, and the R1000 test would be associated with  $VO_{2max}$  and also with the capacity to withstand concentrations of hydrogen ions in muscle and blood. On the other hand, to our knowledge the test with greater predominance of the anaerobic pathway does not seem to be related to performance factors in triathlon that have been studied (Baldari et al., 2007; Millet et al., 2011; Suriano and Bishop, 2010). However, these hypotheses still need to be tested, and this would allow us to suggest that a closer relationship with the performance factors of triathlon will help increase the explanatory capacity of talent

identification tests.

In relation to the tests with predominance of the anaerobic pathway, greater explanatory capacity was observed in the S100 test compared to R400. Despite the fact that these tests are of very similar duration ( $62.0 \pm 4.1$  s for S100 and  $62.5 \pm 5.7$  s for R400), the greater explanatory capacity of the swimming test may be related to the influence of speed over the first metres on triathlon performance (Vleck et al., 2008). Another possible explanation for the low predictive capacity of the talent identification tests in relation to performance factors in triathlon is that these tests are exclusively focused on the physiological component, contrary to the proposals of Abbot and Collins (2004), and do not take account of other factors with a higher influence on triathlon performance, such as anthropometric (Ackland et al., 1998; Landers et al., 2013), psycho-social (Ruiz-Tendero and Martin, 2012), mechanical (Bentley et al., 2002; Ofoghi et al., 2016) and tactical factors (Hauswirth et al., 2010; Ofoghi et al., 2016). Nor do they consider one of the distinctive features of triathlon compared to other long-distance cyclical sports: the influence of the preceding discipline on performance in the following segment (Bentley et al., 2007; Delextrat, 2005; Millet and Vleck, 2000).

## Conclusions

The explanatory capacity of FETRI's tests for predicting performance in the discipline in competition is low. It varies substantially according to the type of the test and the number of years between the test and the competition. However, it seems to be related to the performance factors involved in triathlon.

## Practical implications

Coaches and sports managers should reconsider the suitability of the 100 m and 1000 m swimming and 400 m and 1000 m running talent identification tests for elite triathlon performance. It seems necessary to investigate a new model for the talent identification programme, with tests more appropriate to the performance factors of the sport, so as to identify and select triathletes in a more precise way. This would contribute to the better use of the resources spent on these programmes. Moreover, such programmes can substantially determine the development of female and male athletes, by providing facilities for the most outstanding ones, and conversely

they can discourage those who do not stand out and lead them to abandon the sport, and it is

therefore even more necessary to reconsider the relevance of the identification tests used.

## References

- Abbott A, Collins D. Eliminating the dichotomy between theory and practice in talent identification and development: considering the role of psychology. *J Sports Sci*, 2004; 22(5): 395–408. doi: 10.1080/02640410410001675324
- Ackland TR, Blanksby BA, Landers G, Smith D. Anthropometric profiles of elite triathletes. *J Sci Med Sport*, 1998; 1(1): 52–56
- Baldari C, Di Luigi L, Da Silva SG, Gallota MC, Emerenziani GP, Pesce C, Guidetti L. Relationship between optimal lactate removal power output and Olympic triathlon performance. *J Strength Cond Res*, 2007; 21(4): 1160–65. doi: 10.1519/R-21336.1
- Bentley DJ, Libicz S, Jouglu A, Coste O, Manetta J, Chamari K, Millet GP. The effects of exercise intensity or drafting during swimming on subsequent cycling performance in triathletes. *J Sci Med Sport*, 2007; 10(4), 234–43. doi: 10.1016/j.jsams.2006.05.004
- Bentley DJ, Millet GP, Vleck VE, McNaughton LR. Specific aspects of contemporary triathlon: implications for physiological analysis and performance. *Sports Med*, 2002; 32(6): 345–59. doi: 10.2165/00007256-200232060-00001
- Blanco-Villaseñor A, Castellano J, Hernández-Mendo A, Sánchez-López CR, Usabiaga O. Application of the generalizability theory in sport to study the validity, reliability and estimation of samples. *Revista de Psicología del Deporte*, 2014; 23(1): 131–37
- Bottoni A, Gianfelici A, Tamburri R, Faina M. Talent selection criteria for olympic distance triathlon. *J Hum Sport Exerc*, 2011; 6(2): 293–304. doi: 10.4100/jhse.2011.62.09
- Capelli C, Pendergast DR, Termin B. Energetics of swimming at maximal speeds in humans. *Eur J Appl Physiol Occup Physiol*, 1998; 78(5): 385–93
- Cuba-Dorado A, García-García O, Hernández-Mendo A. Explanatory capacity triathletes performance through talent detection of Spanish federation. *Cuad psicol deporte*, 2015; 15(2): 105–11
- Delextrat A, Brisswalter J, Hausswirth C, Bernard T, Vallier JM. Does prior 1500-m swimming affect cycling energy expenditure in well-trained triathletes?. *Can J Appl Physiol*, 2005; 30(4): 392–403
- Federación Española de Triatlón (2008). *Programa Nacional de Tecnificación Deportiva [National Sports Technification Program]*. Retrieved 22 March 2013 from <http://triatlon.org>
- Gastin PB. Energy system interaction and relative contribution during maximal exercise. *Sports Med*, 2001; 31(10): 725–41
- Hausswirth C, Le Meur Y, Bieuzen F, Brisswalter J, Bernard T. Pacing strategy during the initial phase of the run in triathlon: influence on overall performance. *Eur J Appl Physiol*, 2010; 108(6): 1115–23. doi: 10.1007/s00421-009-1322-0
- Johnston K, Wattie N, Schorer J, Baker J. Talent Identification in Sport: A Systematic Review. *Sports Med*, 2018; 48(1): 97–109. doi: 10.1007/s40279-017-0803-2
- Landers GJ, Ong KB, Ackland TR, Blanksby BA, Main LC, Smith D. Kinanthropometric differences between 1997 World championship junior elite and 2011 national junior elite triathletes. *J Sci Med Sport*, 2013; 16(5): 444–49. doi: 10.1016/j.jsams.2012.09.006
- Millet GP, Vleck VE. Physiological and biomechanical adaptations to the cycle to run transition in Olympic triathlon: review and practical recommendations for training. *Br J Sports Med*, 2000; 34(5): 384–90
- Millet GP, Vleck V, Bentley DJ. Physiological requirements in triathlon. *J Hum Sport Exerc*, 2011; 6(2): 184–204. doi: 10.4100/jhse.2011.62.01
- Ofoghi B, Zeleznikow J, Macmahon C, Rehula J, Dwyer DB. Performance analysis and prediction in triathlon. *J Sports Sci*, 2016; 34(7): 607–12. doi: 10.1080/02640414.2015.1065341



- Ruiz-Tendero G, Martin JJS. Psycho-social factors determining success in high-performance triathlon: compared perception in the coach-athlete pair. *Percept Mot Skills*, 2012; 115(3): 865–80. doi: 10.2466/08.25.PMS.115.6.865-880
- Searle SR, Casella G and McCulloch CE. *Variance components*. New York: John Wiley & Sons, 1992
- Spencer MR, Gastin PB. Energy system contribution during 200-to 1500-m running in highly trained athletes. *Med Sci Sports Exerc*, 2001; 33(1): 157–62
- Suriano R, Bishop D. Physiological attributes of triathletes. *J Sci Med Sport*, 2010; 13(3): 340–47. doi: 10.1016/j.jsams.2009.03.008
- Swann C, Moran A, Piggott D. Defining elite athletes: Issues in the study of expert performance in sport psychology. *Psychol Sport Exerc*, 2015; 16: 3–14. doi: 10.1016/j.psychsport.2014.07.004
- Vaeyens R, Lenoir M, Williams AM, Philippaerts RM. Talent identification and development programmes in sport: current models and future directions. *Sports Med*, 2008; 38: 703–14. doi: 10.2165/00007256-200838090-00001
- Vleck VE, Bentley DJ, Millet GP, Bürgi A. Pacing during an elite Olympic distance triathlon: comparison between male and female competitors. *J Sci Med Sport*, 2008; 11(4): 424–32. doi: 10.1016/j.jsams.2007.01.006

**Corresponding author:****Alba Cuba-Dorado**

Faculty of Education and Sport Sciences. Campus A Xunqueira s/n. Pontevedra. 36005. Spain

Telephone: 0034609155805

Fax: 0034986801701

E-mail : acuba@uvigo.es