Prediction of the Brazilian Paralympic Athletes' Participation and Performance in the London 2012 and the Rio 2016 Paralympic Games[*]

ALAN DE CARVALHO DIAS FERREIRA [a*], ARIEL BEHR [b], VERA PEDRAGOSA [c], ALBERTO REINALDO REPPOLD FILHO [b]

[a] CIDEFES – Centro de Investigação em Desporto, Educação Física, Exercício e Saúde – Universidade Lusófona & CIEQV – Centro de Qualidade de Vida – Escola Superior de Desporto de Rio Maior – Instituto Politécnico de Santarém;
[*] Corresponding author: alan.ferreira@esdrm.ipsantarem.pt
[b] UFRGS – Olympic and Paralympic Studies Center – Federal University of Rio Grande do Sul –Porto Alegre, Brazil.
[c] CICEE – Centro de Investigação em Ciências Económicas & Empresariais & CIP – Centro de Investigação em Psicologia – Universidade Autónoma de Lisboa.

ABSTRACT

This study uses data from the 2009-2012 and 2013-2016 Brazilian Paralympic athletes' participation in athletics, and swimming international and national competitions, to predict their participation and performance in the Paralympic Summer Games of 2012 and 2016. Logistic regressions were conducted to examine the impact of the number of competitions, domestic and international, in which Brazilian Paralympic athletes participated in preparation for the Paralympics and their effective participation and performance in the London 2012 and the Rio 2016 Paralympic Games, and which year(s) of their participation in the cycles of the Paralympic competitions determine with more significance their participation in the Paralympic Games. Results document that for the sports in focus, there is a statistically significant relationship between participation and performance in the Paralympic Games and participation in the events of other sport competitions during the years leading up to each Paralympic cycle. Athletes' participation in international competitions exhibits a higher impact on their participation in the Paralympic Games. Participation in international competitions also shows a positive and statistically significant impact on obtaining a medal by an athlete in the London 2012 and the Rio 2016 Paralympic Games. Additionally, participation in international competitions, in year 1 and year 3 preceding the Paralympic Games, has a greater impact on participation in the Games. These findings can contribute to managing event schedules, training sessions, and even sport funding.

I. Introduction

IN ADDITION TO PROMOTING HEALTH, self-esteem, and social integration for athletes with physical disabilities, Paralympic sports at large, are also a competitive sporting activity (Blauwet & Willick, 2012). Swimming and athletics represent most of the medals awarded to para-athletes in international competitions, including the Paralympic Games. The medal-awarding achievements at the London 2012 Paralympic Games, the Para-Panamerican Games of 2007, 2011, and 2015, and the 72 medals won at the Paralympic Rio 2016 Games are evidence of the victorious trajectory of Brazil in international Paralympic sport (Haiachi et al., 2016). Hence, the Paralympic Games are driven by athletes' sporting performance, which raises the demands of coaching planning and decision-making regarding athletes' training and their preparation for competition (Miller, 2015; Blauwet & Willick, 2012; Legg et at., 2011).

In this framework, paralympic athletes' performance, arguably, benefits from technological innovation, especially, in relation to materials, equipment, and training methodologies. This reality creates a kind of virtuous circle, which increases the opportunities for obtaining sport-linked performance funding, from both public and private sources, therefore promoting excellency in Paralympic sporting activity (Haiachi, 2015; Burkett, 2010).

At the same time as Paralympic sporting business is increasing, it is also worth highlighting sports analytics as it encompasses "the segment of data collection and management, predictive modeling and computational methods in order to find valuable information for sport related decision making" (Sarlis & Tjortjis, 2020; Sands et al., 2017; Microsoft, 2017; Silva, 2016; Miller, 2015; SAP, 2015; Gaudino et al., 2013; Wright et al., 2013). Notably, the search for patterns in data analysis aims at optimizing processes and decision-making related to sporting activities (Intel, 2017; IBM, 2016).

Its use makes it possible to obtain competitive advantages that are linked to sports performance and participation in competitions, the decision-making of managers and allows economic advantages to be reaped that are generated through merchandizing, and the financing of elite sporting activity (Link, 2018; Alamar, 2013). Therefore, sports analytics can promote improvements in athletes' performance, effectiveness, and ultimately efficiency (Baker & Kwartler, 2015; Hakes & Sauer, 2006). Additionally, sports analytics can also contribute to improving the strategic planning of sporting organizations and is useful for the financial planning of public authorities when deciding on the distribution of resources for athletes and sporting organizations (Andreu & Pappous, 2016).

Although participation in international competitions is considered one of the success pillars of the sports (Patatas et. al, 2019; De Bosscher et. al, 2015), few initiatives have been developed that attempt to analyze and predict the participation of elite athletes in

the Olympic and Paralympic Games, particularly that of para-athletes (Mello et. al, 2022). Therefore, this paper aims at using a sample of data from the participation of Brazilian Paralympic athletes in domestic and international athletics, and swimming qualifying competitions, to estimate the odds of their participation and performance in the London 2012 and the Rio 2016 Paralympic Summer Games (PSG).

For empirical testing purposes, we conducted logistic regression estimations, using two samples — the Paralympic Games of 2012 and the Paralympic Games of 2016 — drawn from the National Elite Sports Office of Brazil's Ministry of Sports, covering the 2009-2012 and 2013-2016 sampling periods, respectively, with 140 athlete-events for the Paralympic Games of 2012, and 202 athlete-events for the Paralympic Games of 2016.

This research is organized based on the following research objectives. First, to examine the impact of the number of events in which Brazilian Paralympic athletes participated, in domestic and international athletics, and swimming qualifying competitions, on their participation and performance in the London 2012 and the Rio 2016 Paralympic Games, and how this relationship is influenced by the type of sports. Second, identification of the year(s), in which the events occur, which had a higher impact on the prediction of Brazilian Paralympic athletes' participation in Paralympic Games.

We find that a one-unit increase in the number of international competitions in which a Brazilian Paralympic athlete participates in preparation for the Paralympic Games is associated with an increase in the odds of being in an event at the Paralympic Games by 1.893 (1.619 for swimming and 2.199 for athletics). For the Paralympic Games of London 2012 and Rio 2016, a one-unit increase in the number of international events performed per athlete is associated with a 2.415 and 1.816 increase in the odds of being in the Paralympics (1.992 for swimming and 3.197 for athletics in London 2012; and, 1.447 for swimming and 2.292 for athletics in Rio 2016), respectively. Additionally, results also document that participation in domestic and international competitions shows both a positive and statistically significant impact on obtaining a medal by an athlete. We also found that a one-unit increase in the number of international events in year 1 (2009) and year 3 (2011) is associated with a 2.003 and 1.523 increase in the odds of being in the Paralympic Games, respectively. For the Paralympic Games of Rio 2016, a unit increase in the number of domestic events in year 3 (2015), is associated with a -2.007 decrease in the log-odds of the participation in the Paralympic Games 2016. Finally, we also document that the participation of athletes in more overall events decreases the odds by 0.305 of being in the Paralympic Games.

Overall, our findings suggest that there is a positive impact on the number of international events in which Brazilian Paralympic athletes participated, in preparation for the Paralympics, and their participation and performance (number of medals) in the Paralympic Games, e.g., London 2012 and Rio 2016. Additionally, we document a trend for years 1 and 3 (2009 and 2011, respectively), in which the international competitions occur, with higher impact on the prediction of Brazilian Paralympic athletes' participation in Paralympic Games. However, concerning Rio 2016, participation in

domestic competitions in year 3 (2015) documented a negative impact on predicting Paralympic athletes' participation.

II. Literature Review and Research Questions

The overall performance of the sports industry could benefit from better analytics, and a clearer understanding of how to mitigate the cognitive biases, which may undermine the quality of management decision-making processes (Ward et al. 2019; Ofoghi et al., 2013). For example, assigning athletes or teams to participate in competitions aiming at optimizing funding allocation, as well as marketing strategies (Link, 2018; De Bosscher et. al, 2015).

Studies conducted by Houlihan and Green (2008) and the results presented by the consortium *Sports Policy Leading to International Sporting Success* (SPLISS) (De Bosscher et. al, 2015) have demonstrated that participation in international competitions is a factor that drives the strategy forward and one that can contribute to the international sporting success of countries. However, relatively few studies approach examining the activity data of sporting events quantitatively (Sun & Lin 2022; Bohlmann & Van Heerden, 2008). As Ofoghi et al., (2013) postulate, data analytical methods, where the focus is upon exploration and developing new insights, are becoming increasingly useful tools in analyzing the performance data of elite sports and supporting decision making that is crucial to gaining success.

There is also growing interest in developing intelligent models and prediction systems for elite sports (Link, 2018). However, if we focus on studies that are principally or exclusively based on quantitative data related to sports management, we identify studies that relate to sports with betting and lotteries as being the most common. For example, to predict the behavior of sales in sports betting in China (Sun & Li, 2022), the monthly sales series in sports betting are analyzed and the volume of sales in sports betting in the same region is predicted.

Similarly, Yang et al. (2012) establish a prediction of sports betting consumption in the province of Sichuan, China to improve future promotional strategies. Studies by Spann and Skiera (2009) include a comparative analysis of diverse prediction methods in the betting market, or the application of predictive methods to evaluate time series data in social networks and sports debate forums (Mao et al., 2015; Ofoghi et al., 2013). Quantitative sports data analysis methods have also been used to predict results in the English Premier League (EPL) (Yiannakis et al., 2006) or in order to measure consumer expenditure on sport using multivariate models (Camps & Pappous, 2016).

According to Horvat and Job's (2020), recent review of the mainstream machine learning (ML) models for sports outcome prediction, logistic regression ranks number four in terms of the number of reviewed papers, using that specific ML algorithm (Ishwarya & Nithya, 2021; Kasera & Johari, 2021). In other words, knowing the number of events that can or should be competed in by each team or athlete in a particular sport, in a phase of the Paralympic cycle, can even assist with the planning and management of Paralympic teams, since it may indicate their participation capability in the next Paralympic Games.

According to the data analyzed in this study, for each competition of athletics in one of the Paralympic Games, an average of 86 events were competed in by Brazilian athletes in the four years before the Games, which means more than 28 events per year. The analysis carried out also demonstrated that this reality is different for each sport studied. For swimming, on average, they held 5 preparation events for each event at the 2012 and 2016 Paralympics.

As highlighted by Gold and Gold (2007), since 1988, there has been a convergence process that brought the Paralympic Games to the forefront of the Olympics arena. In the process, new sports were included, and a wider range of disabilities represented, which increased the number of events for each Paralympic team and athlete.

The Paralympic Games also highlight the change in sports competition as a therapy for para-athletes to elite events that carry intrinsic prestige, with growing rivalry for medals. According to the Nielsen Sports Report (2016), "The Rise of Parasports, The Growth of Paralympic Games and Opportunities for Fans and Brands", since London 2012, more than 4,000 para-athletes, from 164 different countries, are competing in the Summer Paralympic Games. This represents a 10 percent increase in the number of athletes and countries competing in the games, plus a 3.8 billion cumulative tv audience and 2.7 million tickets sold to spectators, only in London 2012 (Nielsen Sports Report, 2016).

Considering the increase in the number of paralympic competitions and events, predicting the performance and participation of para-athletes in international competitions has become very important for sports organizations and government, in addition to the companies and brands that finance paralympic sports. In the meantime, the importance of these decisions is often matched by their complexity. For example, as noted by Ofoghi et al. (2013) in a review of sports competition data analysis, the question, "what are your crucial problems?", was interpreted by coaches at several different levels: a) which events in the game have the greatest influence on the result or b) which events in the game influence the "percentage of shots" and "which central passes lead to shots on goal". In contrast, but with the same complexity, the type of sports data analysis in our study can contribute to the management of participation in competitions on multiple levels, including the calendar of events defined for each athlete or team, the management of training and even the profitable management of sport organisations and public entities (Budovich, 2021), since there are considerable investments and logistics related to the participation of para-athletes in international and national competitions.

III. Research Design, Data Description and Empirical Implementation

Since the explained variable assumes a dichotomic nature (dummy variable), the logistic regression estimation method (logit) was used.[1] Logistic regression is a widely used statistical method for analyzing and modeling relationships between one or more

---

[1] The linear probabilistic alternative model, despite being simpler to implement, may be prone to some shortcomings, notably, the non-normality of the errors' distribution, and heteroscedastic variance distributions.

predictor variables and a categorical dependent variable. Reasons why logistic regression may be suitable for use, apart from the binary nature of the dependent variable, are, e.g.: coefficients interpretability, as the effect of a predictor variable on the log-odds of the dependent variable; robustness and flexibility on handling a wide range of predictor variables, including continuous, categorical, and ordinal variables; easy to implement; linearity assumption. Overall, logistic regression is a powerful and flexible method, and it is widely used in many fields, including sports, finance, and social sciences (Hosmer et al. 2013; Menard, 2002).

The specification of the logistic regression model for this study is:

$$\frac{(POG_i=1)}{1-(POG_i=1)} = \exp(\beta_0 + \beta_1 NEvents\_Domest_i + \beta_2 NEvents\_Inter_i + \beta_3 Dum\_NEvents\_Athl_i) \qquad \text{(i)}$$

$$\log\left(\frac{(POG_i=1)}{1-(POG_i=1)}\right) = \beta_0 + \beta_1 NEvents\_Domest_i + \beta_2 NEvents\_Inter_i + \beta_3 Dum\_NEvents\_Athl_i \qquad \text{(ii)}$$

where POG_i denotes the participation in Paralympic Games per athlete and event and is specified as 1 for athletes that participated in an event at the Paralympic Games and 0 for athletes that did not participate in the said event (an alternative specification denotes this variable as the performance in Paralympic Games per athlete and event and is specified as 1 for athletes that were classified in the three first places and received a medal in an event at the Paralympic Games and 0 for athletes that did not receive a medal in the said event; $NEvents_i$ denotes the total number of competitions, domestic (*Domest*) and international (*Inter*), in which each athlete participated in preparation for an event at the Paralympic Games; $Dum\_NEvents\_Athl_i$ denotes a dummy for the number of events per athlete (an athlete who participated in the Paralympic Games in more than one event was coded with the value 1, and athletes who only participated in one event were coded with 0).

Equation (i) describes the odds of an athlete participating in an event of the Paralympic Games (and alternatively their performance. The odds for the event in Paralympic Games is $(POG_i=1)\big/1-(POG_i=1)$ such that *POG* is the probability of the event.

Equation (ii) states the natural logarithm of the odds as a function of the *NEvents* domestic and international in which an athlete participated (the log odds).[2]

For this empirical analysis, we collected data on the participation of the athletics, and swimming Brazilian Paralympic athletes in sports competitions between the 2009 and 2016 Paralympic cycles, available from the National Elite Sports Office of Brazil's Ministry of Sports, described in Table 1.

---

[2] The log odds is the natural logarithm of the odds ratio. The odds ratio is the ratio of the probability of an event occurring to the probability of the event not occurring.

Table 1
**Data on the Brazilian athletics, and swimming paralympic athletes,
available from the National Elite Sports Office of Brazil's Ministry of Sports**

Collected data	Description
Competition name	Name of the sporting competition
Calendar year	Year of the competition
Entity	Organization responsible for the organization of the competition
Country	Country of the sporting competition
Level	Geographic coverage of the competition (World, International, Pan American, South American, National)
Athlete name	Name of the athlete that participated in the competition
Sport	Athlete sport
Ranking	Performance/Classification of the athlete

Data collection was carried out, in three stages, using printed and electronic documents in addition to a standardized electronic form, in order to collect data from the competitions and athletes (Figure 1).

Figure 1
**Data collection stages to predict the participation and performance of
Brazilian athletes in the Paralympic Games of London 2012 and Rio 2016.**
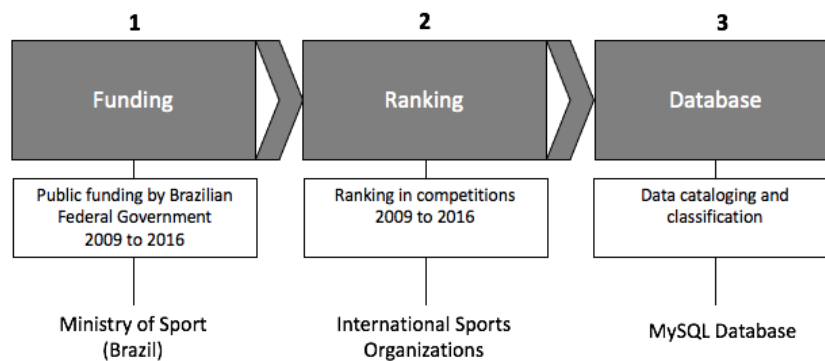
In the first stage of the data collection, information on the public funding of the athletes studied was drawn from the National Elite Sports Office in Brazil's Ministry of Sports. Secondly, data on the participation of paralympic athletes and their results in international and national competitions were collected from the official websites of the Brazilian Paralympic Committee (CPB), the International Paralympic Committee (IPC), and International Sports Organization, for each of the athletics, and swimming sports analyzed. Athletics and swimming represent almost 80 percent of the competitions participated in by Brazilian Paralympic athletes, which were financed by the national federal government.

In the third stage, the data were coded and classified in a MySQL database (DB) by sport, subdivided by competition, event, and athletes' functional class, in a computerized model developed by Ferreira (2018).

For this empirical investigation, three samples were developed, one for the Paralympic Games of 2012 with all the athlete-events for the sampling period 2009 to 2012, another for the Paralympic Games of 2016 with all the athlete-events for the sampling period 2013 to 2016, and another with all data on the athlete-events for the sampling period 2009 to 2016 (including the Paralympic Games of 2012 and 2016). Using the above-described stages, we end up with a sample of 140 athlete-events for the Paralympic Games of 2012, a sample of 202 athlete-events for the Paralympic Games of 2016, and 342 athlete-events for the joint sample of the Paralympic Games of 2012 and 2016.

This empirical investigation is organized around three-research objectives. The first aims at examining the impact of the number of events in which Brazilian Paralympic athletes participated, in domestic and international athletics, and swimming qualifying competitions, on their participation and performance in the Paralympic Games (London 2012 and Rio 2016). The second is how is the aforementioned relationship influenced by the type of sport. The third is the identification of the year(s), in which the events occur, which had a higher impact on the prediction of Brazilian Paralympic athletes' participation in Paralympic Games.

Addressing these research objectives, we performed a descriptive analysis of the data, frequencies, mean, median, standard deviation and maximum, followed by a logistic regression estimation of the number of competitions in which each athlete participated, domestically and internationally, in preparation for the Paralympic Games (also by year) on their participation and performance in the Paralympic Games 2012 and 2016 per athlete and event.

To conduct these tests, we specified our regression models as:

$$\Pr(POG_i = 1) = \beta_0 + \beta_1 NEvents\_Domest_i + \beta_2 NEvents\_Inter_i + \beta_3 Dum\_NEvents\_Athl_i + \varepsilon_i \qquad (1)$$

$$\Pr(POG\_2012_i = 1) = \beta_0 + \beta_1 NEvents\_Domest_i + \beta_2 NEvents\_Inter_i + \beta_3 Dum\_NEvents\_Athl_i + \varepsilon_i \qquad (2)$$

$$\Pr(POG\_2016_i = 1) = \beta_0 + \beta_1 NEvents\_Domest_i + \beta_2 NEvents\_Inter_i + \beta_3 Dum\_NEvents\_Athl_i + \varepsilon_i \qquad (3)$$

$$\Pr(POG_i = 1) = \beta_0 + \beta_{1\_to\_4} NEvents\_Domest\_Y1\_to\_4_i + \beta_{5\_to\_8} NEvents\_Inter\_Y1\_to\_4_i + \varepsilon_i \qquad (4)$$

$$\Pr(POG\_2012_i = 1) = \beta_0 + \beta_{1\_to\_4} NEvents\_Domest\_Y1\_to\_4_i + \beta_{5\_to\_8} NEvents\_Inter\_Y1\_to\_4_i + \varepsilon_i \qquad (5)$$

$$\Pr(POG\_2016_i = 1) = \beta_0 + \beta_{1\_to\_4} NEvents\_Domest\_Y1\_to\_4_i + \beta_{5\_to\_8} NEvents\_Inter\_Y1\_to\_4_i + \varepsilon_i \qquad (6)$$

Equations (1) to (3) address the first and second research objectives, and Equations (4) to (6) examine the third research objective, for the Paralympic Games (London 2012 and Rio 2016). The variables specification is presented in table 2.

Table 2
Variable Specification

Variable	Variable Specification
POG - Participation (and Performance) in the Paralympic Games per athlete and event	1 for athletes that participated in an event at the Paralympic Games and 0 for athletes that did not participate (alternative specification as 1 for athletes that were classified in the three first places and received a medal in an event at the Paralympic Games and 0 for athletes that did not receive a medal in the said event)
*POG*_2012 - Participation (and Performance) in the Paralympic Games 2012 per athlete and event	1 for athletes that participated in an event at the Paralympic Games 2012 and 0 for athletes that did not participate (alternative specification as 1 for athletes that were classified in the three first places and received a medal in an event at the Paralympic Games 2012 and 0 for athletes that did not receive a medal in the said event)
*POG*_2016 - Participation (and Performance) in the Paralympic Games 2016 per athlete and event	1 for athletes that participated in an event at the Paralympic Games 2016 and 0 for athletes that did not participate (alternative specification as 1 for athletes that were classified in the three first places and received a medal in an event at the Paralympic Games 2016 and 0 for athletes that did not receive a medal in the said event)
NEvents_Domest – Number of Events	Total number of domestic events in which each athlete participated in preparation for the Paralympic Games
NEvents_Inter - Number of Events	Total number of international events in which each athlete participated in preparation for the Paralympic Games
Dum_NEvents_Athl	Dummy for the number of events at the Paralympic Games per athlete (an athlete who participated in more than one event was coded with the value 1, and athletes who only participated in one event were coded with 0)
NEvents_Domest_Y1_to_4	Total number of domestic events in which each athlete participated in preparation for the Paralympic Games in Years 1 to 4; Year 1 (2009 for Paralympic Games 2012 and 2013 for Paralympic Games 2016); Year 2 (2010 for Paralympic Games 2012 and 2014 for Paralympic Games 2016); Year 3 (2011 for Paralympic Games 2012 and 2015 for Paralympic Games 2016); Year 4 (2012 for Paralympic Games 2012 and 2016 for Paralympic Games 2016)
NEvents_Inter_Y1_to_4	Total number of international events in which each athlete participated in preparation for the Paralympic Games in Years 1 to 4

IV. Results and Discussion

Table 3 displays the summary statistics on the variables used in the empirical testing for the Paralympic Games of 2012 and Paralympic Games of 2016 samples. We found that the *NEvents* (*Domest* and *Inter*) mean is higher in the preparation for the Paralympic Games of 2012 than for 2016, 2.007, 1.743 and 1.678, 1.253 competitions, respectively.

Table 3
Summary Statistics – Paralympic Games (2012 and 2016)

Variable	Paralympic Games					London 2012					Rio 2016				
	N	Mean	P50	Std	Max	N	Mean	P50	Std	Max	N	Mean	P50	Std	Max
POG	342	0.655	1	0.476	1	140	0.629	1	0.485	1	202	0.673	1	0.470	1
Performance (POG)	342	0.743	1	0.438	1	140	0.836	1	0.372	1	202	0.678	1	0.468	1
NEvents_Domest	342	1.813	1	2.150	6	140	2.007	1	2.287	6	202	1.678	1	2.044	6
NEvents_Inter	342	1.453	1	1.500	5	140	1.743	1	1.533	5	202	1.253	1	1.446	5
Dum_NEvents_Athl	342	0.866	1	0.342	1	140	0.907	1	0.291	1	202	0.837	1	0.371	1
NEvents_Year1	342	0.725	0.5	0.807	2	140	0.886	1	0.814	2	202	0.614	0	0.785	2
NEvents_Year2	342	0.725	0	0.825	2	140	0.75	1	0.806	2	202	0.708	0	0.840	2
NEvents_Year3	342	1.251	1	1.167	4	140	1.486	1	1.096	4	202	1.089	1	1.190	4
NEvents_Year4	342	0.564	0	0.663	2	140	0.629	0.5	0.703	2	202	0.520	0	0.633	2

Logistic regression results on the impact of the number of competitions in which Brazilian Paralympic athletes participated, in domestic and international athletics, and swimming qualifying competitions, on their participation and performance in Paralympic Games, are reported in table 4, columns (1) and (2), respectively. Columns (3) and (4) show the relationship between the number of competitions in which the athlete participated and their participation in Paralympic Games by the sports most participated in, swimming, and athletics, respectively.

Table 4
Logistic Regression Estimation Results (*logit*) – Paralympic Games, Domestic and International Competitions, and Sports

Variables	(1) All Dum_POG	Odds Ratio	(2) All Dum_POG (Performance)	Odds Ratio	(3) By Sport Swimming	Odds Ratio	(4) By Sport Athletics	Odds Ratio
NEvents_Domest	0.0426	1.044	7.5260***	1855.7	-0.136	0.873	0.183	1.201
	(0.0611)		(1.5073)		(0.101)		(0.417)	
NEvents_Inter	0.638***	1.893	5.5555***	258.6	0.482***	1.619	0.788***	2.199
	(0.0997)		(1.5307)		(0.168)		(0.128)	
Dum_NEvents_Athl	-1.187***	0.305	-1.3514**	0.259	-1.040*	0.353	-0.688	0.503
	(0.409)		(0.6059)		(0.568)		(0.852)	
Constant	0.848**	2.336	-3.3638***	0.035	1.764***	5.833	-0.0282	0.972
	(0.411)		(0.7829)		(0.598)		(0.832)	
Observations	342		342		170		164	
Pseudo R-squared	0.1347		0.8787		0.068		0.175	
Wald chi2	52.26		25.12		11.98		39.68	
Prob chi2	0.00000		0.00000		0.00740		0.00000	

Results document that for a unit increase in the number of international competitions in which each athlete participated a 0.638 increase is expected in the log-odds of the participation in an event in the Paralympic Games and is statistically significant at the 1 percent level. The participation in domestic and international competitions is both positive and statistically significant, at the 1 percent level, impact on obtaining a medal by the athlete (column (2)). The expected value of the log-odds of the participation in an

event in the Paralympic Games for athletes who competed in more than one competition (vis-à-vis those who competed in just one event) is -1.187 and is statistically significant at the 1 percent level. Similar results, coefficients, signs, and statistical significance are documented for the participation in the Paralympic Games by sports, columns (3) and (4) of table 4. The number of domestic competitions in which an athlete participated is not statistically significant in explaining their participation in an event in the Paralympic Games.

Table 5, columns (1) and (2), and by sports in columns (3) and (4), presents the results on the impact of the number of competitions in which Brazilian Paralympic athletes participated, domestic and international, on their participation and performance in an event in the London 2012 Paralympic Games. Findings document that for a unit increase in the number of international competitions in which each athlete participated a 0.882 increase is expected in the log-odds of the participation in an event in the London 2012 Paralympic Games, and is statistically significant at the 1 percent level. In the London 2012 Paralympic Games, the participation in domestic and international competitions is both positive and statistically significant, at the 1 percent level, impact on athletes' performance (column (2)). Similar results, coefficients, signs, and statistical significance are documented for the participation in the Paralympic Games by sports, columns (3) and (4) of table 5. The number of domestic competitions in which an athlete participated is also statistically significant at the 5 percent level, in explaining their participation in an event in London 2012, however, its impact is lower when compared to the participation in international competitions.

Table 5
Logistic Regression Estimation Results (logit) – London 2012, Domestic and International Competitions, and Sports

Variables	(1) London 2012	Odds Ratio	(2) London 2012 (Performance)	Odds Ratio	(3) By Sport Swimming	Odds Ratio	(4) By Sport Athletics	Odds Ratio
NEvents_Domest	0.216**	1.242	33.264***	2.80e+14	0.150	1.162	0.638	1.893
	(0.0962)		(1.728)		(0.147)		(0.724)	
NEvents_Inter	0.882***	2.415	34.276***	7.69e+14	0.689***	1.992	1.162***	3.197
	(0.162)		(0.909)		(0.236)		(0.274)	
Dum_NEvents_Athl	-1.174	0.309	-15.414***	2.02e-07	-1.094	0.335	1.989**	7.309
	(0.720)		(1.062)		(0.895)		(0.948)	
Constant	-0.0849	0.919	-17.157***	3.54e-08	0.357	1.429	-3.805***	0.022
	(0.721)		(0.688)		(0.851)		(1.204)	
Observations	140		140		72		65	
Pseudo R-squared	0.2407		0.8802		0.1498		0.3328	
Wald chi2	35.14				12.63		18.17	
Prob chi2	0.00000				0.0055		0.0004	

Results on the impact of the number of competitions in which Brazilian Paralympic athletes participated, domestic and international, on their participation in the Rio 2016

Paralympic Games, are reported in table 6, columns (1) and (2), and by sports in Columns (3) and (4).

Table 6
Logistic Regression Estimation Results (logit) – Rio 2016, Domestic and International Competitions, and Sports

Variables	(1) Rio 2016	Odds Ratio	(2) Rio 2016 (Performance)	Odds Ratio	(3) By Sport Swimming	Odds Ratio	(4) By Sport Athletics	Odds Ratio
NEvents_Domest	-0.0728	0.930			-0.429***	0.651	-0.123	0.884
	(0.0814)				(0.149)		(0.548)	
NEvents_Inter	0.596***	1.816	4.381***	79.895	0.369	1.447	0.830***	2.292
	(0.141)		(1.645)		(0.232)		(0.175)	
Dum_NEvents_Athl	-1.209**	0.299	-1.091	0.336	-1.126	0.324	-0.884	0.413
	(0.509)		(0.7577)		(0.793)		(0.894)	
Constant	1.273**	3.573	-3.321***	0.361	3.209***	24.760	0.516	1.676
	(0.516)		(0.9433)		(0.903)		(0.861)	
Observations	202		92		98		99	
Pseudo R-squared	0.103		0.78.55		0.1127		0.1352	
Wald chi2	25.82		8.05		11.89		24.18	
Prob chi2	0.00000		0.0178		0.0078		0.00000	

Results show that for a unit increase in the number of international competitions in which each athlete participated, a 0.596 increase is expected in the log-odds of the participation in Rio 2016 paralympic games and is statistically significant at the 1 percent level. Additionally, the participation in international competitions is also positive and statistically significant, at the 1 percent level, impact on obtaining a medal by an athlete in the Rio 2016 Paralympic Games (column (2)). The number of domestic competitions in which an athlete participated is not statistically significant in explaining their participation in Rio 2016 Paralympic Games. Per type of sports, participating in domestic swimming competitions decreased the log-odds by 0.429 of their participation in the Rio 2016 Paralympic Games, statistically significant at the 1 percent level, and participating in international athletics competitions increased the log-odds by 0.830 of their participation in Rio 2016 Paralympic Games.

Table 7 shows the regression results on the identification of the year(s), in which the competitions occur (domestic and international), with a higher impact on the prediction of Brazilian Paralympic athletes' participation in Paralympic Games (London 2012 and Rio 2016).

Table 7
Logistic Regression Estimation Results (*logit*) – considering the *Domestic*
and *International NEvents* per year before the Paralympic Games

Variables	(1) All		(4) 2012		(3) 2016	
	Dum_POG	Odds Ratio	Dum_POG	Odds Ratio	Dum_POG	Odds Ratio
NEvents_Domest_Year1	0.422	1.525	-0.569	0.566	1.009	2.744
	(0.563)		(0.982)		(0.680)	
NEvents_Inter_Year1	0.695*	2.003	1.597***	4.937	0.191	1.210
	(0.359)		(0.530)		(0.568)	
NEvents_Domest_Year2	0.392	1.479	0.0854	1.089	0.298	1.348
	(0.275)		(0.531)		(0.364)	
NEvents_Inter_Year2	-	1	-	1	-	1
NEvents_Domest_Year3	-0.820*	0.441	0.986	2.682	-2.007***	0.134
	(0.467)		(0.724)		(0.764)	
NEvents_Inter_Year3	0.420**	1.523	0.488*	1.629	0.710***	2.035
	(0.175)		(0.269)		(0.253)	
NEvents_Domest_Year4	0.0672	1.070	0.414	1.513	-0.0808	0.922
	(0.272)		(0.450)		(0.382)	
NEvents_Inter_Year4	0.641	1.898	-2.656**	0.070		1
	(1.584)		(1.167)			
Dum_NEvents_Athl	-1.235***	0.291	-1.235	0.291	-1.187**	0.305
	(0.425)		(0.759)		(0.490)	
Constant	0.902**	2.464	-0.0424	0.959	1.132**	3.102
	(0.425)		(0.779)		(0.491)	
Observations	308		122		184	
Pseudo R-squared	0.1043		0.2187		0.110	
Wald chi2	37.59		28.64		28.40	
Prob chi2	0.00000		0.00040		0.000186	

Results in table 7 document that for a unit increase in the number of international competitions in year 1 and year 3 in which each athlete participated, a 0.695 and a 0.420 increase are expected, respectively, in the log-odds of the participation in the Paralympic Games, a relationship that is statistically significant at the 10 and 5 percent levels, respectively.

Overall, regression results indicate that there is a positive impact on the number of competitions in which Brazilian Paralympic athletes participated in preparation for the Paralympics and their participation and performance in an event in the Paralympic Games (London 2012 and Rio 2016). Indeed, the number of international competitions in which each athlete participated increased in the log-odds of the participation in an event in the Paralympic Games, while in most cases the number of domestic competitions in which an athlete participated is not statistically significant in explaining the participation in the Paralympic Games.

Research on public and private organizations funding for countries and athletes to obtain results in international competitions, especially in the Olympic and Paralympic Games, has gained increasing attention as it provides valuable insights (Gratton et al.,

2018; Bernard & Busse, 2004). These studies highlight the importance of adequate financial investment for sporting success in elite competitions. For instance, results presented by De Bosscher et al. (2019) indicate that countries that allocate substantial resources to sports development are more likely to achieve better results and win a greater number of medals at the Olympics.

In this regard, results documented in this paper can contribute to the management of participation in competitions, such as the competition schedule for each athlete or team, and even financial management of sports organizations, once the relevance of funding allocation optimization grows in relation to the current economic context, in which it is valued due to the allocation of public financial resources in an efficient way, using a transparent criterion for the distribution of these resources.

In addition, considering current research shows that funding of high-performance sports is directly linked to winning medals at Olympic and Paralympic Games (De Bosscher et al., 2019; Gratton et al., 2018), the results on the prediction of the performance and participation of athletes in the Paralympic Games and their relationship vis-à-vis their participation in competitions during the Paralympic cycle, presented in this paper, can be used to collaborate with the funding effectiveness verification of the high-performance Paralympic athletes and teams. These results are consistent with the previous literature that discussed how participation in international competitions is considered to be one of the success pillars of sports (De Bosscher et. al, 2015; Houlihan & Green, 2008). In brief, studies have demonstrated that participation in international competitions is a factor that drives the strategy forward and that can contribute to the international sporting success of many countries.

V. Summary and Concluding Remarks

Since the use of sports data analysis techniques can assist sports management and sports funding, this paper explores data from one of the success pillars of sports, international and national competitions, over the 2009-2016 period, in order to study the impact of the number of competitions in which Brazilian Paralympic athletics and swimming athletes participated, to predict their participation in the London 2012 and the Rio 2016 Paralympic Games.

Logistic regression estimations results, using samples of the Paralympic Games of 2012 and the Paralympic Games of 2016, document that a one-unit increase in the number of competitions in which a Brazilian Paralympic athlete participates in international competition is associated with a 1.619 and 2.199 increase in the odds of being in an event at the Paralympic Summer Games, respectively. Similar results, coefficients, signs, and statistical significance are documented for the participation in the Paralympic Games by the sports studied (athletics and swimming).

On the contrary, the results also confirm that the number of domestic competitions in which an athlete participated is not statistically significant (or lower than the participation in international competitions) in explaining their participation in an event in the Paralympic Games.

We also found that participation in international competitions in different years of the Paralympic cycle may have different impacts on the presence of these athletes in the Paralympic Games. However, we were able to document a trend for years 1 and 3, in which the events occur, with higher impact on the prediction of Brazilian Paralympic athletes' participation in Paralympic Games.

Given that we found that, for the two sports in point, there is a robust relationship between participation in international competitions and athletes' participation in the Paralympic Games, future studies can broaden the analysis presented in this paper, including data from athletes from other countries, which may allow the similarities and differences between sports and between different nations to be confirmed as well as conduct experiments to identify the most effective form of financial support for athletes.

Given that there is considerable investment and logistics related to the participation of para-athletes in international and national competitions, it should be noted that this type of sports data analysis can contribute to managing their attendance in competitions, such as the calendar of each athlete or team, and also the financial management of sport organizations and public entities. The relevance of funding allocation optimization is growing in relation to the current economic context, in which it is valued due to the allocation of public financial resources in an efficient way, using a transparent criterion for the distribution of these resources. In addition, in the context of major sports events and attempts to predict the participation of elite athletes in the Olympic and Paralympic Games, this work is a contribution to the advancement of knowledge in this area. We suggest that further studies address the prediction of athletes' participation in competitions using machine learning models for detailed study of the influence of each year of the competition cycle on the probability of participation in Olympic and Paralympic Games, which was a limitation of the study. Likewise, since this study is limited in that it only uses data from Brazilian athletes, future studies may expand the sample to athletes from other countries and from different sports, other than athletics and swimming.

REFERENCES

Alamar, B. C. (2013). *Sports Analytics: A Guide for Coaches, Managers, and Other Decision Makers*. Columbia University Press.

Baker, R. E., & Kwartler, T. (2015). Sport analytics: Using open source logistic regression software to classify upcoming play type in the NFL. *Journal of Applied Sport Management, 7*(2), 11.

Bernard, A. B., & Busse, M. R. (2004). Who wins the Olympic Games: Economic resources and medal totals. *Review of Economics And Statistics, 86*(1), 413-417.

Blauwet, C., & Willick, S. E. (2012). The paralympic movement: Using sports to promote health, disability rights, and social integration for athletes with disabilities. *Pm&r, 4*(11), 851-856.

Bohlmann, H. R., & Van Heerden, J. H. (2008). Predicting the economic impact of the 2010 FIFA World Cup in South Africa. *International Journal of Sport Management And Marketing, 3*(4), 383-396.

Budovich, L. S. (2021). Business management of the sport industry by considering the digitalization (2021). *Journal of Human Sport and Exercise*, 16, 4, 1619-1624.

Burkett, B. (2010). Technology in Paralympic sport: performance enhancement or essential for performance? *British Journal of Sports Medicine, 44*(3), 215-220.

Camps. A., & Pappous, A. (2016). Predicting the Evolution of Sports Federation Membership:{An Important} Tool to Asses National Governing Bodies' Strategic Planning. *Journal of Sports Science, 4*(2), 57-69.

De Bosscher, V.; Shibli, S.; Weber.; Andreas, C. Is prioritization of funding in elite sport effective? An analysis of the investment strategies in 16 countries. European Sport Management Quarterly 2019, 19, 221-243. https://doi.org/10.1080/16184742.2018.1505926

De Bosscher, V., Shibli, S., Westerbeek, H., & Van Bottenburg, M. (2015). *Successful elite sport policies: an international comparison of the sports policy factors leading to international sporting success (SPLISS 2.0) in 15 nations*. Meyer & Meyer Sport.

Ferreira, A. (2018). *A model for the management of elite sport information in Brazil (*PhD Thesis). Federal University of Rio Grande do Sul, Rio Grande do Sul - UFRGS, RS, Brazil. https://www.lume.ufrgs.br/handle/10183/180547.

Gaudino, P., Iaia, F. M., Alberti, G., Strudwick, A. J., Atkinson, G., & Gregson, W. (2013). Monitoring training in elite soccer players: systematic bias between running speed and metabolic power data. *International Journal of Sports Medicine, 34*(11), 963-968.

Gold, J. R., & Gold, M. M. (2007). Access for all: The rise of the paralympic games. *Journal of the Royal Society for the Promotion of Health, 127*(3), 133-141.

Gratton, C., Shibli, S., & Coleman, R. (2018). Sport economics. Routledge.

Haiachi, M. D. C., Cardoso, V. D., Reppold Filho, A. R., & Gaya, A. C. A. (2016). Reflexões sobre a carreira do atleta paraolímpico brasileiro. *Ciência & Saúde Coletiva, 21*, 2999-3006.

Haiachi, M. C. (2015). Inovação nas relações entre atleta, treinamento e carreira esportiva em entidades gestora do esporte paralímpico. *Deslandes A, Da Costa LP, Miragaya A, organizadores. O Futuro dos Mega-eventos Esportivos. Rio de Janeiro: Engenho Arte e Cultura*, 479-488.

Hakes, J. K., & Sauer, R. D. (2006). An economic evaluation of the Moneyball hypothesis. *Journal of Economic Perspectives, 20*(3), 173-185.

Horvat, T., & Job, J. (2020). The use of machine learning in sport outcome prediction: A review. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 10*(5), e1380.

Hosmer, D. W., Jr., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied Logistic Regression* (3rd ed.). Wiley.

Houlihan, B., & Green, M. (2008). *Comparative Elite Sport Development*. Butterworth-Heinemann.

IBM (2016). *Watson IoT and sports: change the game*. https://www.ibm.com/blogs/internet-of-things/watson-iot-sports/.

Intel (2017). *The digitization of sports*. https://www.newsroom.intel.com/press-kits/digitization-of-sports.

Ishwarya, K., & Nithya, A. A. (2021, December). Relative analysis and performance of machine learning approaches in sports. *In 2021 5th International Conference on Electronics, Communication and Aerospace Technology (ICECA)* (pp. 1084-1089). IEEE.

Kasera, M., & Johari, R. (2021). Prediction using machine learning in sports: A case study. *In Data Analytics and Management: Proceedings of ICDAM* (pp. 805-813). Springer Singapore.

Legg, D., De Bosscher, V., Shibli, S., & Van Bottenburg, M. (2011). What is the ideal paralympic sport system? Future research directions. *ICSSPE Bulletin, 1*(61), 52-63.

Mao, L. L., Zhang, J. J., & Connaughton, D. P. (2015). Determinants of Demand for Sports Lottery: Insights from a Multilevel Model". Asian Economic and Financial Review 5, n.º 8, 973–87.

Link, D. (2018). Sports Analytics-wie aus (kommerziellen) Sportdaten neue Möglichkeiten für die Sportwissenschaft entstehen. *German Journal of Exercise and Sport Research, 48*(1), 13-26.

Mello, J., Pedretti, A., Caporal, G., Mahique, J., Ferreira, A., Reppold-Filho, A., & Gaya, A. (2020). Secular trend of sports practice of Brazilian children and young people in the decade of mega-sport events. *Journal of Human Sport and Exercise* 17, n.º 1.

Menard, S. (2002). *Applied Logistic Regression Analysis* (2nd ed.). Sage Publications.

Microsoft (2017). *Microsoft sports performance platform.* https://www.news.microsoft.com/de-de/microsoft-sports-performance-platform.

Miller, T. W. (2015). Sports analytics and data science: winning the game with methods and models. *FT press.*

Nielsen Sports Report (2016). *Paralympics and para-sports. the rise of para-sports, the growth of the paralympic games and the opportunities for fans and brands.* https://pdfslide.net/documents/paralympics-and-para-sports-nielsen-the-rise-of-para-sports-the-growth-of.html?page=1

Ofoghi, B., Zeleznikow, J., MacMahon, C., & Raab, M. (2013). Data mining in elite sports: a review and a framework. *Measurement in Physical Education and Exercise Science, 17*(3), 171-186.

Patatas, J. M., De Bosscher, V., Derom, I., & De Rycke, J. (2020). Managing parasport: An investigation of sport policy factors and stakeholders influencing para-athletes' career pathways. *Sport Management Review, 23*(5), 937-951.

Sands, W. A., Kavanaugh, A. A., Murray, S. R., McNeal, J. R., & Jemni, M. (2017). Modern techniques and technologies applied to training and performance monitoring. *International Journal of Sports Physiology and Performance, 12*(s2), S2-63.

SAP (2015). *Unveils SAP sports one solution for soccer.* http://news.sap.com/sap-unveils-sap-sports-one-solution-for-soccer/.

Spann, M. and Skiera, B. (2009), Sports forecasting: a comparison of the forecast accuracy of prediction markets, betting odds and tipsters. J. Forecast., 28: 55-72.

Sarlis, V., & Tjortjis, C. (2020). Sports analytics—Evaluation of basketball players and team performance. *Information Systems, 93*, 101562.

Silva, R. M. (2016). *Sports analytics.* PhD Dissertation, Department of Statistics and Actuarial Science, Faculty of Science, Simon Fraser University, Greater Vancouver, Canada.

Sun, G., & Lin, Y. (2022). Evaluation Model of Sports Culture Industry Competitiveness Based on Fuzzy Analysis Algorithm, Mathematical Problems in Engineering, vol. 2022, Article ID 9708646, 8 pages, 2022.

Ward, P., Windt, J., & Kempton, T. (2019). Business Intelligence: How Sport Scientists Can Support Organization Decision Making in Professional Sport, International Journal of Sports Physiology and Performance, 14(4), 544-546.

Wright, C., Atkins, S., Jones, B., & Todd, J. (2013). The role of performance analysts within the coaching process: Performance Analysts Survey 'The role of performance analysts in elite football club settings.' *International Journal of Performance Analysis in Sport, 13*(1), 240-261.

Yang, Y. L., Cheng, L. L., & Zgao, Y. T. (2012). Measurement Model of Sports Lottery Sales Volume and the Concerning Promotion Strategy: a case study of Sichuan Province. *Journal of Chengdu Sport University,* 9.

Yiannakis, A., Selby, M. J., Douvis, J., & Han, J. Y. (2006). Forecasting in sport: the power of social context—a time series analysis with English premier league soccer. *International Review for the Sociology of Sport, 41*(1), 89-115.