Research Article

International Journal of Sports Science & Coaching

International Journal of Sports Science & Coaching
0(0) 1

DOI: 10.1177/17479541211031817
journals.sagepub.com/home/spo
SAGE

Does periodization work? Athletes perform better in major events than in minor competitions

Pekka Matomäki and Aapo Räntilä

Abstract

Previous studies on periodization have led to the view that most athletes fail to peak at major events. These conclusions might be based on definitions of "peak performance" that are too narrow. In this study, a track and field athlete was defined as succeeding in a competition if their outcome was within the acceptable range from the best result of the season. The data are from seven championship finals and semifinals together with 42 Diamond League competitions from the 2010s, altogether 7,087 individual results. All field events and running events up to 400 m were included. The majority of athletes succeeded in major events (67.0% in sprint and 57.5% in field events). Overall, championship finals elicit success rates that are more than 70% higher than the basic level achieved in Diamond League competitions ($p < 0.001$). Success rates were systematically higher (over 60% higher) for the top three versus other competitors in every race ($p < 0.001$). When an acceptable range is adopted for the definition of what a successful result is, the majority of athletes manage to peak at the most important competitions. In addition, finishing in the top three in championship finals typically requires a peak performance.

Keywords

Peak performance, track and field, sport analytics, success rate

Introduction

The many definitions for periodization typically share three characteristics: Periodization aims to minimize overtraining, minimize injury rate and maximize performance for a given date.[1–3]

Recent studies have examined the basis for the theory of periodization.[1,4–10] Most critically, the timelines of periodization studies are typically too short lasting less than 13 weeks.[1,11] Second, in the intervention studies periodization is not always distinguished from training variation[1,5,10] or programming.[11,12] Lastly, the focus of research on periodization is often too narrow: It ignores the practice's multilateral nature and the importance of minor, yet important aspects that affect performance, including nutrition, sleep, stress, medication, skill, mental training and supplementation.[1,6,13,14]

Even though researchers dispute the theoretical basis of periodization,[5,6,11] it is used to guide the training processes of most athletes.[15,16] Performance timing, one of the three main characteristics of the purpose of periodization, has been studied by comparing the outcome in the championship finals to the athlete's season-best result (SB) in swimming[17,18] and track and field.[7,19] These results suggest that athletes and coaches mostly fail to produce a peak performance at major events. Less than 40% of swimmers[17,18] and 16% to 28% of track and field athletes[7] were able to achieve their SB in championship competitions. This translates to about 60% to 80% of athletes failing to periodize their training appropriately.

However, to say that an athlete succeeds only when producing a SB is too strict. Intuitively, it is fair to

Reviewers: Beat Knechtle (University of Zurich, Switzerland)
Ray Stefani (California State University, Long Beach, USA)

Biology of Physical Activity, Faculty of Sport and Health Sciences, University of Jyväskylä, Jyväskylä, Finland

Corresponding author:
Pekka Matomäki, Biology of Physical Activity, Faculty of Sport and Health Sciences, University of Jyväskylä, Jyväskylä, PO Box 35, Finland.
Email: pekka.j.matomaki@jyu.fi

consider a javelin result that is 1 cm less than a SB a successful peak performance of the season. For borderline cases such as this, an "acceptable range" should be introduced to indicate how much a result can differ from SB and still be classified as a successful effort. A logical choice would be to utilize race-to-race variability, which indicates how much random volatility there is in reproducing an athletic result. This variability has previously been calculated in, for example, cycling (0.4%–2.4%),[20] rowing (0.7%–1.4%),[21] track and field (1.0%–3.2%),[22] long-distance running (1.2%–4.2%),[23] swimming (0.6%–1.0%),[24] and the biathlon (2.5%–3.2%).[25]

The study measured athletes' success in peaking at major track and field events with a novel binary variable of success or not, which includes acceptable day-to-day variability. This success rate is then compared to a "basic" competition level from less important competitions. If periodization works in the way it is presumed to work, most athletes will succeed in championship competitions and the success rate will be better than in normal competitions.

Methods

This study was a retrospective analysis of data available in the public domain. There was no recruitment of participants, no experimental treatment or intervention, and no names of individuals were used or can be identified in the manuscript. Consequently, the Ethics Committee of the University of Jyväskylä decided that informed consent or ethics approval were not required.

Data collection

The sports used were all field events and all running events up to 400 m in track and field. In these events, the tactical elements are minimal,[26] meaning the athletes give their best efforts in each competition. It follows that in these events their competition results reveal their present level of performance. Moreover, the championship competitions are the one where all elite track and field athletes primarily aim to peak.

The data (from the website of World Athletics www.worldathletics.org during spring 2020) consist of three full seasons of Diamond League competitions covering the 2010s: 2013 ($n=14$), 2016 ($n=14$) and 2019 ($n=14$). Furthermore, all seven finals of the championship competitions—world championships (WC) and Olympic games (OG)—from the 2010s were included: 2011 (WC), 2012 (OG), 2013 (WC), 2015 (WC), 2016 (OG), 2017 (WC), 2019 (WC), as well as the semifinals of running events from the championships. Altogether, 7,087 individual results (4,039 sprint and 3,048 field results) from 49 competitions were included.

Success rate

In this study, an athlete was defined as succeeding in a competition if the athlete's result was within an acceptable range from SB. A logical choice for the acceptable range is race-to-race variability that covers the natural volatility in race performances. These coefficients of variations have been calculated in track and field events,[22] and they range from 1% to 3%, as shown in Table 1, and they form the acceptable range in this study. In running events, for example, an athlete is said to have a successful result if the finish time is no more than 1.0% slower than their SB. This results in a binary variable for success: an athlete either succeeds in a competition or not.

The *success rate* of a competition was defined to be a percentage which indicates the portion of athletes that succeeded in each competition. Similarly, *season best rate* (SB rate) was defined to be a percentage which indicates the portion of athletes that achieved their SB in a particular competition. If the acceptable range is chosen to be 0%, the success rate coincides naturally with the SB rate.

Random success rates in championship competitions

The Diamond League events were intended to provide a base level to which the championships were compared. To study the chance that the observed success rates in championship competitions would be attained merely randomly the following was assumed: Success rates in Diamond League competitions are independent and identically distributed (i.i.d.). Applying this i.i.d. assumption, central limit theory is usable, and one can utilize a normal distribution to calculate the probability that an observed success rate in a major event occurs randomly.

To explore whether the performance level was increased during the season, each Diamond League season (14 competitions per season) was divided into two parts: the first seven competitions and the last seven, which were then compared.

Statistics

The descriptive statistics, means, and standard deviations were weighted by the number of participants in each competition. Because the variance between groups was different, Welch's t test was used for comparison. The magnitude of differences was calculated with a corrected effect size Hedge's g,[27,28] and a 95% confidence interval (CI) was also calculated for g. Following the recommendation,[29,30] p values were reported as

Table 1. Definition of an acceptable range: How much competition result can differ from season best (SB) while still being described as a successful result. Acceptable difference is a coefficient of variation from different events.[22]

Event	Acceptable difference from SB
Sprint run (up to 400 m, with and without hurdles)	1.0%
High jump, triple jump	1.5%
Long jump, pole vault	2.5%
Javelin, hammer throw, discus, shot put	3.0%

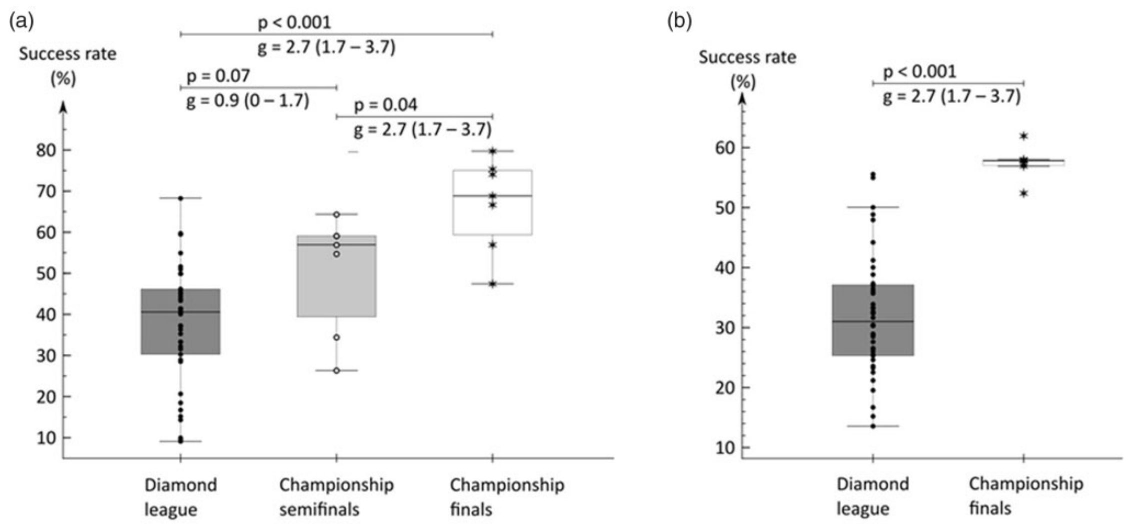

Figure 1. (a) Success rates of sprint runs (all events up to 400 meters) in Diamond League competitions, in championship semifinals, and in championship finals. (b) Success rates of all field events in diamond league competitions and in championship finals. P: p value; g: Hedge's g value for effect size (95% CI for g).

equalities unless they were very small (<0.001). The Shapiro-Wilk test was used to examine the normality, and no significant differences from normal distribution were detected.

Results

Championship finals elicit higher success rates than Diamond League competitions (Figure 1).

In Table 2, success rates and SB rates are reported along with the probabilities that the observed success rates from championship finals occurred by chance, assuming that success rate is merely a random process.

The correlation between SB rate and success rate among all observed competitions was $r = 0.82$ for sprint events and $r = 0.90$ for field events.

Success rates for top three finishers vs. other competitors are listed in Table 3.

Comparing the beginning (the first seven competitions) of a Diamond League season to the ending (the last seven competitions), the success rate of the first seven competitions in sprint events was lower ($33.8 \pm 10.4\%$ vs. $42.5 \pm 15.8\%$, $p = 0.043$, $g = 0.64$, 95% CI: 0–1.3), while there was no difference in that of field events ($33.3 \pm 10.4\%$ vs. $31.0 \pm 9.4\%$, $p = 0.45$, $g = 0.23$, 95% CI: –0. –1.5).

Discussion

An elaborated success rate, which contains the acceptable day-to-day variability of the performance, was introduced. The main finding was that applying it, the majority of athletes succeed in major events (Table 2): 67.0% in sprint and 57.5% in field events. The introduced day-to-day variability explains why previous research,[7,17–19] based on a strict SB-or-not rule, failed to observe the successful timing of the peak performance to be the norm. Moreover, achieving a top three finish in championship finals, or even qualifying for the finals, requires a peak performance (Table 3), which has also been reported in literature.[17–19]

As was hypothesized, the championship competitions seemed to elicit higher success rates compared to a basic performance level (Figure 1). By assuming that peak performance is a random process, the

Table 2. Weighted season best (SB) and success rates and their standard deviations (*SD*) in different competitions, and probability that the success rate in finals occurred by chance.

	Sprint run			Field events		
Competition	SB rate (*SD*) %	Success rate (*SD*) %	Probability of seeing such success rate by chance	SB rate (*SD*) %	Success rate (*SD*) %	Probability of seeing such success rate by chance
Diamond league	11.3 (6.6)	38.5 (14.1)		7.2 (5.4)	32.1 (9.8)	
Championship semifinals	14.4 (5.5)	50.8 (14.3)				
Championship finals	30.0 (12.2)	67.0 (11.3)		20.8 (2.9)	57.5 (2.8)	
2011 WC final		56.9	9.5%		62.0	0.1%
2012 OG final		79.7	0.2%		57.9	0.4%
2013 WC final		68.8	1.6%		57.3	0.5%
2015 WC final		74.1	0.6%		58.0	0.4%
2016 OG final		75.3	0.4%		56.9	0.6%
2017 WC final		47.4	26.3%		52.4	2.0%
2019 WC final		66.7	2.3%		57.8	0.4%

WC: world championship; OG: Olympic games.

Table 3. Weighted mean success rates (SD) (%) for places 1–3 vs. 4 or lower, in championship semifinals those who qualified for the finals vs. those who were disqualified, and those who qualified for the finals from the semifinals vs. the finals. The *p* values and effect sizes (Hedge's *g*) for their differences are calculated, as well as 95% CI for the effect size.

		1–3	4 or lower	*p* value	Effect size (95% CI)
Diamond league	Sprints	51.9 (21.4)	30.4 (12.3)	<0.001	1.2 (0.8–1.7)
	Field	51.2 (15.8)	21.8 (10.2)	<0.001	2.2 (1.7–2.7)
Championships finals	Sprints	87.7 (12.4)	53.5 (13.1)	<0.001	2.5 (1.1–3.9)
	Field	84.3 (3.7)	47.9 (3.1)	<0.001	9.9 (6.1–13.7)
Championships semifinals	Sprints	Qualified for finals 61.1 (13.6)	disqualified from finals 45.1 (15.1)	0.06	1.0 (–0.1 to 2.2)
Semifinals vs. finals	Sprints	Qualified for finals from semifinals 61.1 (13.6)	Championship Finals 67.0 (11.3)	0.4	0.4 (–0.6 to 1.5)

probabilities that the observed success rates in championship finals occurred by chance (Table 2) is below 2.5% for 12 out of the 14 analyzed finals. These strongly imply that the underlying distribution for the success rates in the championship finals differs drastically from that of Diamond League competitions. Similar implications were made in Konings and Hettinga,[31] where it was stated that the championship competition environment seems to increase the athletes' performance level.

There are two possible explanations. Either the athletes manage to achieve a peak performance level for a given date, or only those who by chance have a peak performance level at championships will advance to the finals. To minimize the possibility for a false positive deduction, sprint semifinals from championship competitions were included in the analysis. They reveal that semifinals also differed from the basic performance level in Diamond League competitions (Figure 1), albeit not so radically as in the finals. This suggests that the athletes generally do succeed better in major events.

By its definition, if periodization has been successful, then the performance at a major event will be successful. An equivalent proposition is that if the performance at a major event has been unsuccessful, then periodization has been unsuccessful (Figure 2). Hence, one could deduce a failure of periodization from negative outcomes from major events. However, the inverse proposition is not necessarily true: if performance at a major event has been successful, one cannot say anything certain about periodization (Figure 2).

Naturally, a successful performance may stem from successful use of periodization. However, other explanations might emerge. These include recognized short term augmentations such as successful tapering[32] and the intrinsic motivation to perform well at major events.[33] Also, more hypothetical events might affect: more continuous and larger financial and social

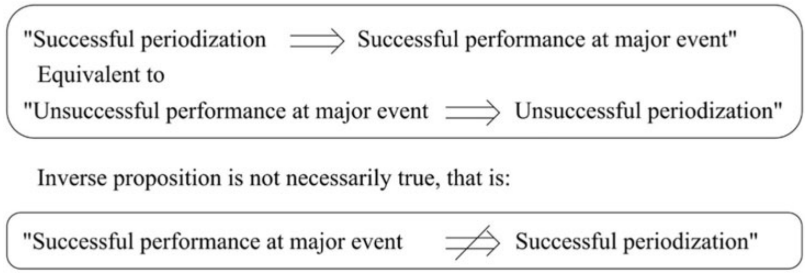

Figure 2. How success at major events and successful periodization are linked.

Table 4. Success rates in major events in swimming. Normal distribution was utilized, as the original articles stated that no deviations from normality were detected. The mean and standard deviation (SD) in the case of Mujika et al.[17] was calculated by combining the means and SD of men and women.

n	Major events	Event	Mean (*SD*) finishing time difference compared to preSB	Estimated success rate (%)	Study
424	Athens Olympic games	All swimming finals	0.58 (1.13) %	58	18
7832	All WC and OG between 2011 and 2017	All swimming finals	0.36 (1.20) %	64	17

preSB: season best of the athletes at the time entering the observed competition; WC: world championship; OG: olympic game.

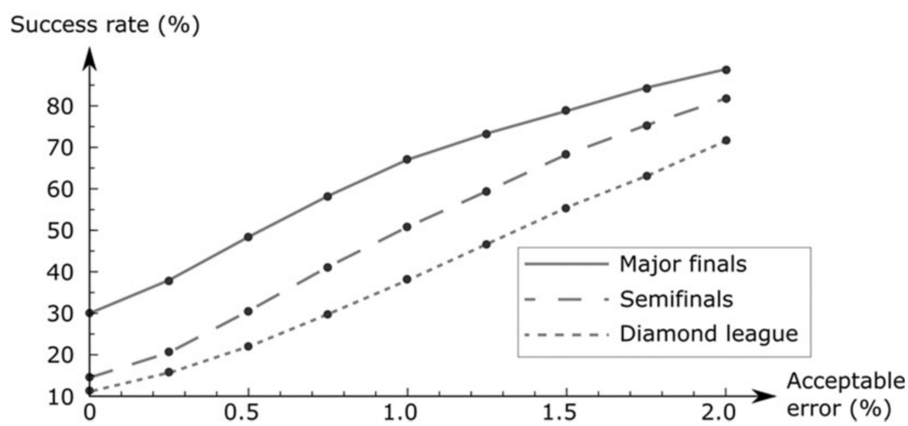

Figure 3. Mean success rate of sprint events in major finals, semifinals, and Diamond League when acceptable error is adjusted to 0–2.0%.

support toward the important competitions, prioritizing the key end of season competition, good weather and wind circumstance, luck, and so on. The main difference to periodization is that when periodization is a structured construction that tries to maximize an athlete's capacity to succeed, the above features "may just happen in the right way", that is, without planning.

A strong correlation ($r > 0.8$) was found between success rates and the SB rates. Thus, the usage of the SB-or-not rule to estimate success rates seems reliable. In swimming[17,18] and track and field,[7,19] SB rates in major events were reported to be between 16% and 38%. In the present study, the SB rates in championship finals (13%–38%, data not shown) were comparable to those. Hence, it is probable that when adding the acceptable range to the above-mentioned research results, success rates similar to the present paper should arise.

To test this hypothesis, success rates from swimming championship competitions from 2004 as well as from between 2011 and 2017 were estimated by utilizing Gaussian distribution, applying the mean and *SD* of the result difference to the SB prior to each major event (preSB) given in literature.[17,18] The coefficient of variation of 0.8% for swimming performance[24,34] was used for the acceptable range. As preSB were used in these studies, instead of SB as in the present study, slightly overestimated values are possible. The results are presented in Table 4. The estimated success rates (58%–66%) are compatible to the success rates seen in the present study (47%–80%, Table 2). The basic level from less important competitions cannot be verified from these studies.

The right choice for an acceptable range is a matter of considerable debate, because skill, race dynamics, environment, and the level of the athlete all contribute to the coefficient of variation.[34] Previous studies have used SB rate,[17–19] or in other words, a 0% acceptable range. In the present paper, 1.0% was used for sprinters, which is a justifiable group average.[22] Altering the acceptable range from 0 to 2.0%, the success rates vary considerably (Figure 3). However, the main outcome, that athletes succeed better in major events than in minor races, remains valid irrespective of acceptable range. This emphasizes how athletes succeed in Championships races. Any choice greater than 0.55% (which converts to approximately 55–60 ms in the 100 m dash) guarantees that a majority of finalists succeed in major events.

A limitation of the study is that the generic form of the data does not give any specific reason for the success seen in major events. Moreover, individualized acceptable range were not calculated for athletes. Instead, a rougher group average was used.

Conclusion

The present study proposed a concept, success rate, which included day-to-day variance of performance in its definition. This construct helped show that the majority of athletes succeed in major events. Furthermore, comparing the performance level from major events to the basic level from less important races revealed a clear difference in performance. Moreover, without the peak form, success in major events is unlikely.

Athletes' performance level is elevated in major events. Periodization, in the form used by athletes today, is one possible explanation for this, but a clear conclusion cannot be made based on generic data from this study.

Acknowledgements

We thank everyone who has given feedback on the earlier drafts of the paper.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Pekka Matomäki https://orcid.org/0000-0002-0241-8406

References

1. Afonso J, Nikolaidis PT, Sousa P, et al. Is empirical research on periodization trustworthy? A comprehensive review of conceptual and methodological issues. *J Sport Sci Med* 2017; 16: 27–34.
2. Harries S, Lubans D and Callister R. Systematic review and meta-analysis of linear and undulating periodized resistance training programs on muscular strength. *J Strength Cond Res* 2015; 29: 1113–1125.
3. Pihkala L. Specialization in track sports, what it is, and what it is not. *Am Phys Educ Rev* 1913; 18: 154–159.
4. Kiely J. New horizons for the methodology and physiology of training periodization – block periodization: new horizon or a false dawn? *Sports Med* 2010; 40: 803–807.
5. Kiely J. Periodization paradigms in the 21st century: periodization paradigms in the 21st century. *Int J Sports Physiol Perform* 2012; 7: 242–250.
6. Mujika I, Halson S, Burke LM, et al. An integrated, multifactorial approach to periodization for optimal performance in individual and team sports. *Int J Sports Physiol Perform* 2018; 13: 538–561.
7. Loturco I, Nakamura FY. Training periodization: an obsolete methodology? *Aspetar Sports Med J* 2016; 5: 110–115.
8. Kiely J, Pickering C and Halperin I. Comment on "biological background of block periodized endurance training: a review". *Sports Med* 2019; 49: 1475–1477.
9. Cissik J, Hedrick A and Barnes M. Challenges applying the research on periodization. *Strength Cond J* 2008; 30: 45–51.
10. Afonso J, Rocha T, Nikolaidis PT, et al. A systematic review of meta-analyses comparing periodized and non-periodized exercise programs: why we should go back to original research. *Front Physiol* 2019; 10: 1–7.
11. Hornsby WG, Fry AC, Haff GG, et al. Addressing the confusion within periodization research. *JFMK* 2020; 5: 68.
12. Cunanan AJ, DeWeese BH, Wagle JP, et al. The general adaptation syndrome: a foundation for the concept of periodization. *Sports Med* 2018; 48: 787–797.
13. Blumenstein B and Orbach I. Periodization of psychological preparation within the training process. *Int J Sport Exerc Psychol* 2020; 18: 13–23.
14. Jeukendrup AE. Periodized nutrition for athletes. *Sports Med* 2017; 47: 51–63.

15. Durell DL, Pujol TJ and Barnes JT. A survey of the scientific data and training methods utilized by collegiate strength and conditioning coaches. *J Strength Cond Res* 2003; 17: 368–373.
16. Ebben WP and Carroll R. Strength and conditioning practices of national hockey league strength and conditioning coaches. *J Strength Cond Res* 2004; 18: 889–897.
17. Mujika I, Villanueva L, Welvaert M, et al. Swimming fast when it counts: a 7-year analysis of Olympic and world championships performance. *Int J Sports Physiol Perform* 2019; 14: 1132–1139.
18. Issurin V, Kaufman L, Lustig G, et al. Factors affecting peak performance in the swimming competition of the Athens Olympic games. *J Sports Med Phys Fitness* 2008; 48: 1–8.
19. Casado A, Renfree A, Maroto-Sánchez B, et al. Individual performances relative to season bests in major track running championship races are distance-, position- and sex-dependent. *Eur J Hum Mov* 2020; 44: 146–161.
20. Paton CD and Hopkins WG. Variation in performance of elite cyclists from race to race variation in performance of elite cyclists from race to race. *Eur J Sprot Sci* 2006; 6: 25–31.
21. Smith TB and Hopkins WG. Variability and predictability of finals times of elite rowers. *Med Sci Sports Exerc* 2011; 43: 2155–2160.
22. Hopkins WG. Competitive performance of elite track and field athletes. Variability and smallest worthwhile enhancements. *Sportscience* 2005; 9: 17–20.
23. Hopkins WG and Hewson DJ. Variability of competitive performance of distance runners. *Med Sci Sports Exerc* 2001; 33: 1588–1592.
24. Pyne DB, Trewin CB and Hopkins WG. Progression and variability of competitive performance of Olympic swimmers. *J Sports Sci* 2004; 22: 613–620.
25. Skattebo Ø and Losnegard T. Variability, predictability, and race factors affecting performance in elite biathlon. *Int J Sports Physiol Perform* 2018; 13: 313–319.
26. Hanley B and Hettinga FJ. Champions are racers, not pacers: an analysis of qualification patterns of Olympic and IAAF world championship middle distance runners. *J Sports Sci* 2018; 36: 2614–2620.
27. Lakens D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Front Psychol* 2013; 4: 1–12.
28. Hedges LV. *Statistical methodology in meta-analysis*. 1982. ERIC Information Analysis Products (071). Report no: ERIC-TM-83. Institution: ERIC Clearinghouse on Tests, Measurement, and Evaluation, Princton, N.J.
29. Greenland S, Senn SJ, Rothman KJ, et al. Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. *Eur J Epidemiol* 2016; 31: 337–350.
30. Goodman S. A dirty dozen: twelve P-value misconceptions. *Semin Hematol* 2008; 45: 135–140.
31. Konings MJ and Hettinga F. The impact of different competitive environments on pacing and performance. *Int J Sports Physiol Perform* 2018; 13: 701–708.
32. Le Meur Y, Hausswirth C and Mujika I. Tapering for competition: a review. *Sci Sport* 2012; 27: 77–87.
33. Mallett CJ and Hanrahan SJ. Elite athletes: why does the 'fire' burn so brightly? *Psychol Sport Exerc* 2004; 5: 183–200.
34. Malcata RM and Hopkins WG. Variability of competitive performance of elite athletes: a systematic review. *Sports Med* 2014; 44: 1763–1774.