

A COMPARISON BETWEEN DIFFERENT TESTS FOR FUNCTIONAL THRESHOLD POWER DETERMINATION IN RUNNING

COMPARACIÓN ENTRE DIFERENTES TESTS PARA LA DETERMINACIÓN DEL UMBRAL DE POTENCIA FUNCIONAL EN CARRERA

Javier Olaya Cuartero¹; Sergio Sellés Pérez¹; Alberto Ferriz Valero¹; Roberto Cejuela Anta¹

¹Área Educación Física y Deportiva. Departamento de Didáctica General y Didácticas Específicas. Universidad de Alicante, Alicante, España.

Detalles del artículo:

Número de palabras: 4472; Tablas: 5; Figuras: 5; Referencias: 20

Recibido: febrero 2019; Aceptado: abril 2019; Publicado: julio 2019

Conflicto de interés: El autor declara que no existen conflictos de interés.

Correspondencia del autor: Javier Olaya Cuartero, javierolayacafd@gmail.com

Abstract

Introduction: power is an important variable in performance assessment. With the increasing availability of power measurement devices, it is simple to associate watts to the running Functional Threshold Power measured in watts (rFTPw) as indicated by Vance (2016). The main goal of this study is to find the most appropriate methodology for rFTPw determination. **Material and methods:** five different methodologies were carried out in 9 recreational triathletes (22.9 ± 4.8 years) to calculate rFTPw: 3-minute test, 3-minute – 9-minute test, 3-minute – 9-minute Stryd test, 3-laps – 6-lap test and 30-minute test. All tests were performed on an athletics track with a Stryd footpod. **Results:** the 3-minute – 9-minute test presented a lower average error in comparison to the mean rFTPw value (rFTPw M) of all power and pace measurement tests. **Conclusions:** the 3-minute – 9-minute test could be the best choice regardless of the distance or duration of the competition because rFTPw changes depending on the duration of each test. The watts associated with Critical Speed (CS) and obtained in the 3-minute test are not a valid measure, given their greater average error in both power and pace. The 30-minute test could be an alternative to determine rFTPw through the data obtained in a training session or competition of similar length. The tests with the lowest average error in power are the ones with the lowest error in pace. Therefore, pace remains the main variable to monitor external load in running.

Key words: Triathletes, accelerometry, estimation, pace.

Vance (2016) posits that power in running allows to obtain watts associated with “running Functional Threshold Power measured in watts” (rFTPw), through the Stryd footpod (Stryd Power Meter, Boulder, CO, USA), used in the most recent studies (Austin, Hokanson, McGinnis & Patrick, 2018; García-Pinillos, Roche-Seruendo, Marcén-Cinca, Marco-Contreras & Latorre-Román, 2018; Aubry, Power & Burr, 2018). The same company Stryd® also proposes a series of field tests to determine the rFTPw (Vance, 2016), assuming that power is not measured but estimated with a complex set of calculations and assumptions through different formulas and algorithms (Van Dijk & Van Megen, 2017).

The main goal of this study is to determine which test provides a lower error in rFTPw estimation by conducting a comparative analysis between the five different methodologies. Moreover, the pace in seconds per kilometer (s/km) associated with rFTPw will also be calculated and compared.

MATERIAL AND METHODS

Participants

Table 1 shows the characteristics of the sample of 9 recreational triathletes who participated in the study. All of them were part of the same group, trained under the same triathlon coach and had two years of training experience. All results are expressed as Mean (**M**) ± Standard Deviation (SD). The training consisted of 5 hours per week on 5 different days (2 swimming, 1 running, 1 strength, 1 cycling). They all signed the PAR-Q (Revision of the Physical Activity Readiness Questionnaire) (Thomas, Reading & Shepard, 1992) and an informed consent to participate in the study.

Table 1. Participants' characteristics (M ± SD)

	Sample (n = 9)
Age (years)	22.9 ± 4.8
Weight (kg)	74.1 ± 8.2
Height (cm)	175.6 ± 9.8
BMI (kg)	23.2 ± 1.8
Fat mass (kg)	7 ± 1.5
Muscle mass (kg)	36.7 ± 3.4
Σ8 SS (mm)	61.4 ± 12.8

*BMI (Body Mass Index)

*SS (Skinfold Sum)

Observational design

All the tests were performed during the first two weeks of the second month of training. A previous counterbalance of the sample (Table 2) was carried out to determine the order of the field tests and avoid a possible learning effect with respect to the previous tests. The test were performed on an approved athletics track (lane 1), with a rest period of at least 24 hours between one test and the next under regular weather conditions and with clear sky: average temperature ($\pm 18.9 \pm 0.3$ °C), minimum temperature (15 ± 1.1 °C), maximum temperature (24.4 ± 1 °C) and wind speed (6.3 ± 1.7 km/h).

Table 2. Order for field tests

Tests Order Participant	1	2	3	4
1	3 MT	TS 3min-9min	T 3L-6L	T 30min
2	TS 3min-9min	T 3L-6L	T 30min	3 MT
3	T 3L-6L	T 30min	3 MT	TS 3min-9min
4	T 30min	3 MT	TS 3min-9min	T 3L-6L
5	3 MT	TS 3min-9min	T 3L-6L	T 30min
6	3 MT	TS 3min-9min	T 3L-6L	T 30min
7	TS 3min-9min	T 3L-6L	T 30min	3 MT
8	T 3L-6L	T 30min	3 MT	TS 3min-9min
9	T 30min	3 MT	TS 3min-9min	T 3L-6L

*3-minute test (3 MT)

*3-minute – 9-minute Stryd test (TS 3 min-9 min)

*3-lap – 6-lap test (T 3L-6L)

*30-minute test (T 30 min)

Procedures

Estimation of body composition

At 8m on the day of the first test, a fasting estimation of each triathlete’s body composition was carried out through anthropometric ISAK (International Society for the Advancement of Kinanthropometry) method. All measurements were taken in the same tent, at ambient temperature ($22^{\circ} \pm 1^{\circ}\text{C}$) and by the same researcher (level 3). Measurements followed Ross and Marfell-Jones’s methods (Ross & Marfell-Jones, 1991) and were taken three times for each subject. The equipment used included a Holtain skinfold caliper (Holtain Ltd. U.K), a Holtain bone breadth calliper (Holtain Ltd., U.K), scales, a stadiometer and anthropometric tape (SECA LTD., Germany). Physical characteristics were measured in the following order: age, weight and stature. The following measurements were also taken: bi-epicondylar humerus, bi-styloid and bi-epicondylar femur breadths; relaxed arm, flexed and tense arm; mid-thigh and calf girths; sub-scapular, biceps, triceps, suprailiac, supraspinale, front thigh, medial calf and abdominal skinfolds. The Lee equation (Lee et al., 2000) was used to calculate muscle mass and the Withers equation (Withers, Craig, Bourdon & Norton, 1987) to calculate fat mass.

3-minute all-out exercise test (3 MT)

This 3-minute all-out exercise test (3 MT) is the gold standard method to determine critical speed (CS) (Pettit, Jammick & Clarck, 2012). These authors state that CS is determined by the pace of the last 30 seconds of the test. The power and pace of the last 30 seconds were associated with rFTPw. All subjects were informed that they had to achieve maximum speed as soon as possible (5 seconds) and maintain it throughout the test. A whistle indicated the beginning and end of the test. All participants were informed that the test would last 3 minutes, but no information was provided on the remaining time or the pace. A standard 15-minute warm-up was previously performed.

3-minute – 9-minute test (T 3 min – 9 min) and 3-minute – 9-minute Stryd test (TS 3 min – 9 min)

Only one test was performed per participant. The proposed protocol consisting of a maximum effort of 3 minutes, a recovery of 30 minutes, and another maximum effort of 9 minutes (Vance, 2016) was followed. A standard 15-minute warm-up was previously performed. However, two different methodologies were used to calculate rFTPw with this test. In the first one, known as the 3-minute – 9-minute test and proposed by the original author of the test (Vance, 2016), rFTPw was calculated by adding together the average power of the two efforts, dividing this value by two, and selecting 90% of this value. According to the second methodology, called 3-minute – 9-minute Stryd test and put forward by the company Stryd®, rFTPw can be calculated by entering the average power and pace of the first and second effort on Stryd's Power Center platform®. In this test pace can only be calculated using this methodology on Stryd's platform.

3-lap – 6-lap test (T 3L – 6L)

This test was carried out according to the original protocol after a standard 15-minute warm-up, similar to the previous test but informing to the athlete of the laps completed (distance), rather than the minutes spent running (time). Involving a maximum effort of 1,200 meters, a recovery of 30 minutes, and another maximum effort of 2,400 meters (Van Dijk & Van Megen, 2017). Power and pace values were also entered on Stryd's Power Center®.

30-minute test (T 30 min)

This test consisted of a maximum effort of 30 minutes, the rFTPw value being the average power associated with the last 20 minutes (Vance, 2016). In the same way as in the other tests, a standard 15-minute warm-up was previously carried out and power and pace values were entered on Stryd's Power Center®.

Stryd Summit footpod (Stryd Summit Model, Boulder, CO, USA)

All data were recorded second by second with the same wearable device (Phoenix 3hr, Garmin, Taiwan) connected to the Stryd Power Meter (Stryd Summit Model, Boulder, CO, USA), which is a carbon fiber-reinforced footpod (attached to the athlete's shoe) that weighs 9.1 grams and is based on a 6-axis inertial motion sensor (3-axis gyroscope and 3-axis accelerometer).

Statistical Analyses

The data obtained were analyzed statistically with the software Statistical Package for The Social Sciences (v.24.0 SPSS Inc., Chicago, IL, USA). The normality of the sample was checked by the Shaphiro-Wilk test. The parametric results were interpreted with the Mann-Whitney U test, taking into account the significant differences of $p < 0.05$. Bland and Altman's 95% limits of agreement (LOA) and Pearson correlations (r) were applied to determine the concordance between different tests and methodologies. To interpret the magnitude of correlations between measurement variables the following criteria were adopted: < 0.1 (trivial), 0.1-0.3 (small), 0.3-0.5 (moderate), 0.5-0.7 (large), 0.7-0.9 (very large), and 0.9-1.0 (almost perfect) (Hopkins, Marshall, Batterham & Hanin, 2009). Stryd's Power Center platform® was used for estimating rFTPw (Stryd Power Meter, Boulder, CO, USA).

RESULTS

Table 3 shows the power values obtained for 9 participants who performed the field tests. All results are expressed as Mean (M) ± Standard Deviation (SD). The rFTPw values obtained are represented in absolute watts (W) and in watts relative to weight (W/kg). Higher values (315 W ± 40 W) were obtained in the most intense and shortest test 3 MT. Consequently, the lowest values (263 W ± 33 W) were obtained in the longest test (T 30 minutes).

Table 3. Power (W) rFTPw values of different tests (M ± SD)

Test Participant	3 MT		T 3min-9min		TS 3min-9min		T 3L-6L		T 30 min	
	W	W/kg	W	W/kg	W	W/kg	W	W/kg	W	W/kg
1	314	4.4	287	4.1	263	3.7	282	4	269	3.8
2	344	3.9	301	3.5	296	3.4	341	3.9	289	3.3
3	361	5.2	317	4.5	331	4.7	327	4.7	287	4.1
4	320	4.7	272	4	253	3.7	270	3.9	253	3.7
5	241	3.3	246	3.4	231	3.2	225	3.1	226	3.1
6	340	5.3	288	4.5	272	4.3	301	4.7	265	4.1
7	259	3.8	227	3.3	199	2.9	240	3.5	198	2.9
8	320	3.8	298	3.5	298	3.5	304	3.6	289	3.4
9	339	4.1	299	3.6	299	3.6	313	3.8	293	3.6
M	315	4.3	282	3.8	271	3.7	289	3.9	263	3.6
±	±	±	±	±	±	±	±	±	±	±
SD	40	0.7	29	0.5	40	0.5	39	0.5	33	0.4

*Absolute watts (W)

*Watts per kilogram (W/kg)

*3-minute test (3 MT)

*3-minute – 9-minute test (T 3 min – 9 min)

*3-minute – 9-minute Stryd test (TS 3 min-9 min)

*3-lap – 6-lap test (T 3L-6L)

*30-minute test (T 30 min)

Table 4 shows the mean (M) P (W) values of all tests performed by each subject to determine a value of all power and pace measurement tests (rFTPw M). Test with negative values have higher rFTPw values than the rFTPw M. It can be seen that 3 MT all values are higher than the average obtained from all tests.

Table 4. Difference between Power (W) associated to rFTPw M of all tests and rFTPw of each test

Test Participant	rFTPw M	3 MT	T 3min-9min	TS 3min-9 min	T 3L-6L	T 30min
1	283	-30.9	-4.2	20	1	14
2	314	-29.5	13.4	18	-27	25
3	324.6	-36.6	7.9	-6.4	-2.4	37.6
4	273.4	-46	1.8	20.4	3.4	20.5
5	233.8	-7.1	-12.2	2.8	8.8	7.8
6	301	-77.7	12.8	29	0	36
7	224.5	-34.1	-2.5	25.5	-15.5	26.5
8	304.9	-31.1	7.4	6.9	0.9	15.9
9	308.6	-30	9.4	9.6	-4.4	15.6
M	285.3	-35.9	3.7	14	-3.9	22.1
±	±	±	±	±	±	±
SD	33.4	18.7	8.6	11.5	10.9	10.1

*3-minute test (3 MT)

*3-minute – 9-minute test (T 3 min – 9 min)

*3-minute – 9-minute Stryd test (TS 3 min-9 min)

*3-lap – 6-lap test (T 3L-6L)

*30-minute test (T 30 min)

Table 5 shows the differences between the rFTPw M value of all test and each of these tests for pace measured in seconds per kilometer (s/km). Negative numerical values mean a slower pace was obtained in that test, i.e. more seconds per kilometer than the rFTPw M.

Table 5. Difference in pace (s/km) associated to rFTPw M of all tests and rFTPw of each test

Test Participant	rFTPw M	3 MT	TS 3min-9min	T 3L-6L	T 30 min
1	266	33	-9	-21	6
2	231	35	-14	9	-16
3	219	30	9	-14	-34
4	269	56	-13	-13	-15
5	320	19	-4	-15	5
6	246	76	-18	-18	-23
7	332	41	-10	15	-34
8	229	24	7	-28	-8
9	227	23	0	-6	-17
M	260	37	-6	-10	-15
±	±	±	±	±	±
SD	41	18	9	14	14

*3-minute test (3 MT)

*3-minute – 9-minute test (T 3 min – 9 min)

*3-minute – 9-minute Stryd test (TS 3 min-9 min)

*3-lap – 6-lap test (T 3L-6L)

*30-minute test (T 30 min)

The Bland-Altman plots below show the difference between rFTPw M and 3 MT with a bias \pm random error of -35.9 ± 18.7 W (Figure 1). The bias \pm random error for the rFTPw M and the T 3 min – 9 min (Figure 2) and TS 3 min – 9 min (Figure 3) was 3.7 ± 8.6 W and 14 ± 11.5 W respectively. The bias \pm random error for the rFTPw M and the T 3L – 6L obtained through the previous test was -3.9 ± 10.9 W (Figure 4). Finally, the bias \pm random error of 22.1 ± 10.1 W (Figure 5). The concordance of test was evaluated with Bland-Altman plots.

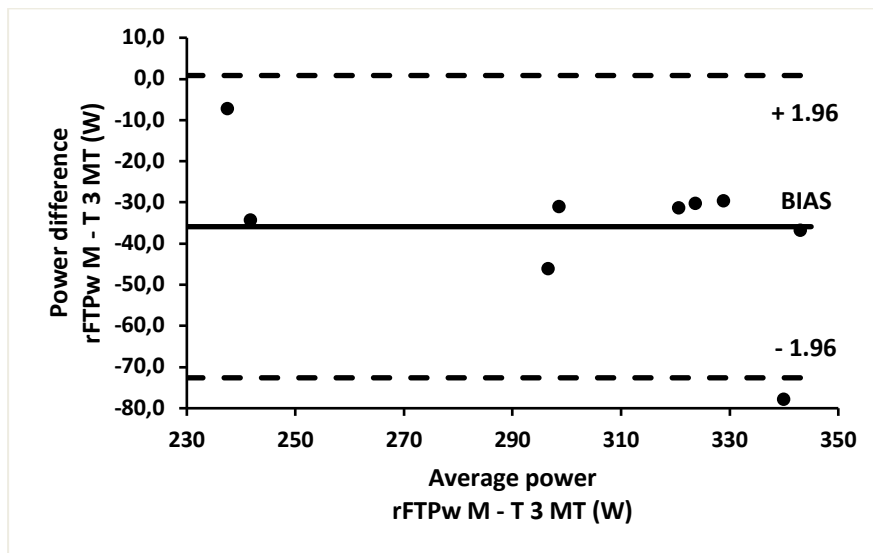


Figure 1. Bland-Altman plot between M rFTPw M and 3 MT

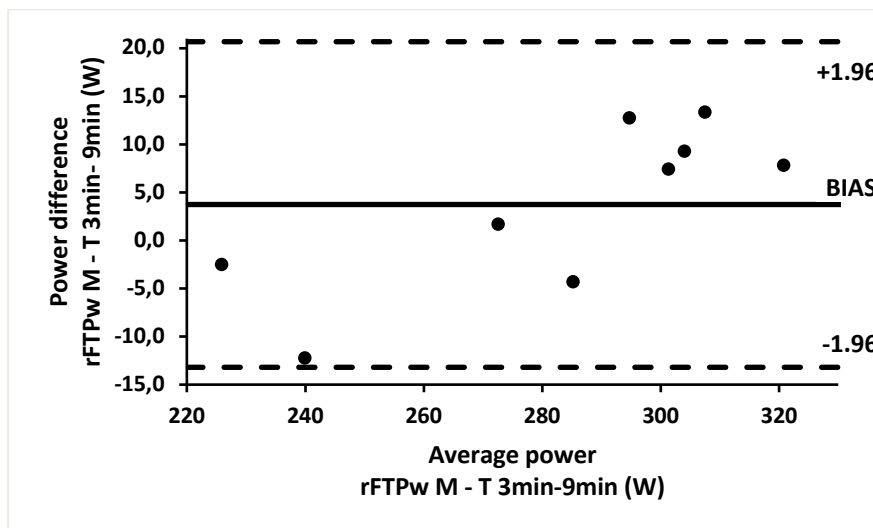


Figure 2. Bland-Altman plot between M rFTPw M and T 3min-9min

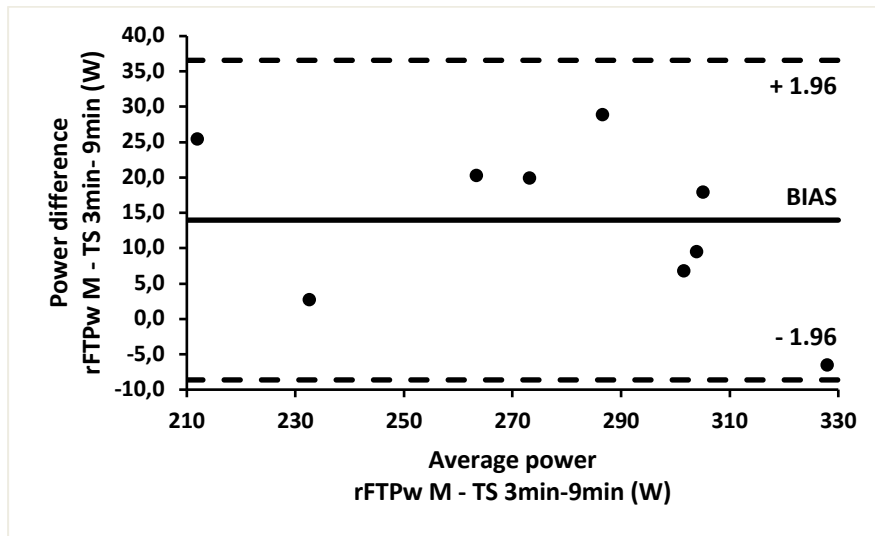


Figure 3. Bland-Altman plot between M rFTPw M and TS 3min-9min

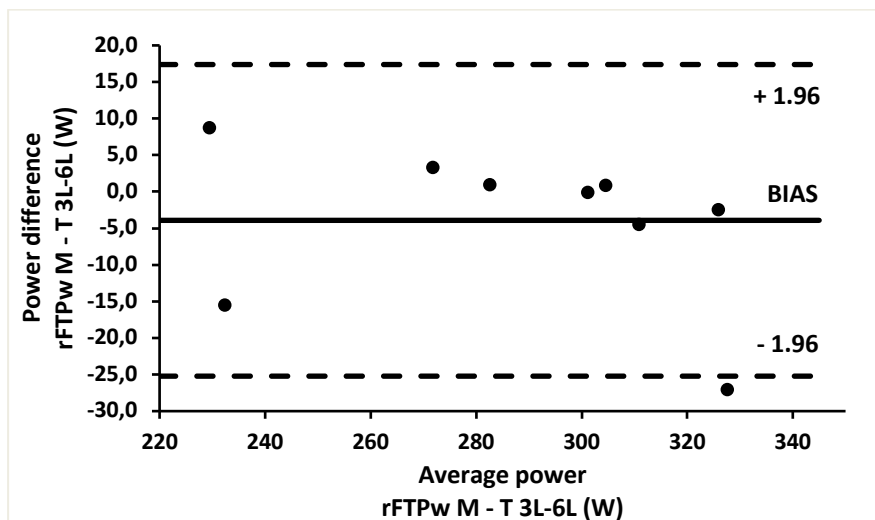


Figure 4. Bland-Altman plot between M rFTPw M and T 3L – 6L

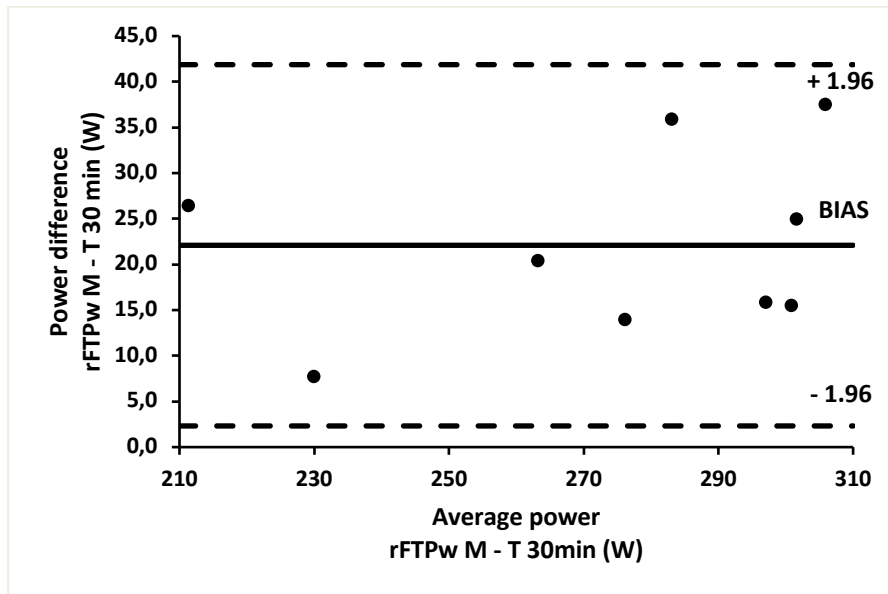


Figure 5. Bland-Altman plot between M rFTPw M and T 30min

DISCUSSION

The 3-minute – 9-minute test shows the lowest mean error relative to rFTPw values calculated through the methodology proposed by Vance (2016). However, the 3-min – 9-minute Stryd test, which is the same test but obtaining the rFTPw value through Stryd’s Power Center®, has a greater error. Regarding pace, T 3 min – 9 min calculated through Vance’s methodology is once again the one with the lowest average error; therefore, this test is proposed as the most suitable methodology to calculate rFTPw in P (W) and pace (s/km) associated with this intensity through the Power Center. The error of T 3L-6L in power (3.7 W; -3.9 W) and pace (-6 s/km; -10 s/km) is similar to that of the previous test, which could be explained by the fact that in recreational athletes performing 3- and 9-minute efforts is very similar to completing 3 laps and 6 laps changing the way in which information is given to the athlete. Probably, a very different result could be obtained if athletes of elite level run faster or recreational athletes run slower, because the time test would have been different. It should be emphasized that almost perfect and significant correlations ($r = 0.919, p < 0.001$) were found between power values in both tests. Compared to previous research, this minimum difference in watts and pace could indicate that in recreational population the power could be a useful tool (Aubry, Power & Burr, 2018).

As far as the duration of the efforts is concerned, the test with the greatest average error in power and pace is the shortest one, 3 MT qualified as gold standard method to calculate CS (Pettit, Jamnick & Clarck, 2012). It is worth noting that 3 MT is the only test in which rFTPw is higher than rFTPw M in all subjects in both measures P (W) and pace (s/km), the pace associated with rFTPw being 37 ± 18 s faster than the pace value for rFTPw M. In fact, the weakest – but still very large and significant – correlations were obtained between 3 MT and T 3 min -9 min ($r = 0.868, p = 0.002$); TS 3 min - 9 min ($r = 0.806, p = 0.009$), T 3L-6L ($r = 0.878, p = 0.002$) and T 30 min ($r = 0.809, p = 0.008$). For these reasons, it can be concluded that the watts associated with CS are not valid as rFTPw. These results are consistent with those obtained in previous studies which conclude that running power estimated through

accelerometry is more sensitive to changes in cadence and other biomechanical factors than running economy or caloric cost (Austin, Hokanson, McGinnis & Patrick, 2018). An increase in speed and cadence may lead to a greater increase in the watts associated with this rFTPw. The main influencing factor could be, however, that recreational athletes are not prepared to maintain a high intensity for a long time, whether it is 9 minutes, 6 laps, or 30 minutes.

Finally, almost perfect and very large correlations were found in T 30 minutes, whose mean error relative to the rFTPw M was greater than that for the T 3 min – 9 min ($r = 0.971$, $p < 0.001$), for TS 3 min – 9 min ($r = 0.944$, $p = 0.001$), and for T 3L-6L ($r = 0.898$, $p = 0.001$), but lower than that for the 3 MT ($r = 0.809$, $p = 0.008$) for both variables – P (W) and pace (s/km) – associated with rFTPw. Therefore, T 30 minutes could be an alternative to calculate rFTPw using paces of a training session or competition of similar duration.

In summary, when calculating rFTPw, the same tests that show a greater or lower error in power, also show them in pace. Therefore, these tests are used to calculate rFTPw for competitions or training sessions but could be possible to train by pace too. This is an extremely important conclusion considering that recently runners have been able to utilize GPS technology embedded in wrist-watches or smart-phone applications to monitor running pace in real time (Puleo & Abraham, 2018). As this is easy, chip, and useful, pace is still the main external measure to control load in running, although power could be useful in popular triathletes to monitor the training load in running. It would be interesting in future research to compare the power measurements in cycling and running for well-trained triathletes and study the concordance and relationship between these two measures at different intensities.

CONCLUSIONS

For power, the most adequate test to calculate the rFTPw is the T 3 min – 9 min using Vance's methodology.

For pace, the TS 3 min – 9 min and T 3L – 6L methodologies could be valid for recreational athletes.

Therefore, T 3 min – 9 min could be the best choice without taking into account the distance or duration of the competition, rFTPw changes depending on the duration of each test.

Power associated with rFTPw obtained in the 3 MT is not a valid measure to determine CS because its average error relative to the rFTPw M is the greatest for both power and pace.

T 30 min could be an alternative to calculate rFTPw with the data obtained in a training session or competition of similar duration.

The tests with the lowest average error in power are the ones with the lowest error in pace; therefore, once the rFTPw is calculated, it is possible to train through power and also by pace.

Although the power could be a useful tool, the pace would remain the main measure to control the external load in running in real time.

ACKNOWLEDGMENTS

To the recreational triathlon training group from University of Alicante, always willing to help in all studies and activities.

REFERENCES

- Allen, H., & Coggan, A. (2006). *Training and Racing with a PowerMeter*. Boulder, CO: VeloPress.
- Aubry, R. L., Power, G. A., & Burr, J. F. (2018). An assessment of running power as a training metric for elite and recreational runners. *The Journal of Strength & Conditioning Research*, 32(8), 2258-2264.
- Austin, C., Hokanson, J., McGinnis, P., & Patrick, S. (2018). The Relationship between Running Power and Running Economy in Well-Trained Distance Runners. *Sports*, 6(4), 142.
- Bland, J. M., & Altman, D. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The lancet*, 327(8476), 307-310.
- Bourdon, P. C., Cardinale, M., Murray, A., Gatin, P., Kellmann, M., Varley, M. C., ... & Cable, N. T. (2017). Monitoring athlete training loads: consensus statement. *International journal of sports physiology and performance*, 12(Suppl 2), S2-161.
- Elvira, J. L. (1998). *Development of a new software for the calculation of the mechanical power in the jump with force platform*. Study of the reproducibility of the values obtained in various jumping tests (Doctoral dissertation, Master thesis, INEF of Castilla y León).
- Faulkner J. (1968). Physiology of swimming and diving. En H. Falis (ed.), *Exercise physiology*. Baltimore: Academic Press.
- García-Pinillos, F., Roche-Seruendo, L. E., Marcén-Cinca, N., Marco-Contreras, L. A., & Latorre-Román, P. A. (2018). Absolute Reliability and Concurrent Validity of the Stryd System for the Assessment of Running Stride Kinematics at Different Velocities. *Journal of strength and conditioning research*.
- Hopkins, WG, Marshall, SW, Batterham, AM, and Hanin, J. Progressive statistics for studies in sports medicine and exercise science. *Med Sci Sports Exerc* 41: 3–12, 2009.
- Jeukendrup, A., & Diemen, A. V. (1998). Heart rate monitoring during training and competition in cyclists. *Journal of Sports Sciences*, 16(1), 91-99.
- Johnson, M. A., Sharpe, G. R., & Brown, P. I. (2007). Inspiratory muscle training improves cycling time-trial performance and anaerobic work capacity but not critical power. *European Journal of Applied Physiology*, 101(6), 761-770.
- Lee, R. C., Wang, Z., Heo, M., Ross, R., Janssen, I., & Heymsfield, S. B. (2000). Total-body skeletal muscle mass: development and cross-validation of anthropometric prediction models. *The American Journal of Clinical Nutrition*, 72(3), 796-803.
- Marroyo, J. A. R., & López, J. G. (2015). Trabajo, potencia y energía. In *Biomecánica básica aplicada a la actividad física y el deporte* (pp. 149-172). Paidotribo.
- Pettitt, R. W., Jamnick, N., & Clark, I. E. (2012). 3-min all-out exercise test for running. *International journal of sports medicine*, 33(06), 426-431.
- Puleo, N. A., & Abraham, K. A. (2018). External Feedback Does Not Affect Running Pace in Recreational Runners. *International Journal of Exercise Science*, 11(5), 384.
- Ross, W. D., & Marfell-Jones, M. J. (1991). *Kinanthropometry*. Physiological testing of elite athlete. Human Kinetics Books. Champaign Il.
- Thomas, S., Reading, J., & Shephard, R. J. (1992). Revision of the physical activity readiness questionnaire (PAR-Q). *Canadian journal of sport sciences*.
- Van Dijk, H., & Van Megen, R. (2017). *The Secret of Running: Maximum Performance Gains Through Effective Power Metering and Training Analysis*. Meyer & Meyer Sport.
- Vance, J. (2016). *Run with power: The complete guide to power meters for running*.
- Withers, R. T., Craig, N. P., Bourdon, P. C., & Norton, K. I. (1987). Relative body fat and anthropometric prediction of body density of male athletes. *European Journal of Applied Physiology and Occupational Physiology*, 56(2), 191-200.