# How Does Power During Running Change when Measured at Different Time Intervals?

Authors

**Felipe García-Pinillos[1], Víctor M. Soto-Hermoso[2], Pedro Á. Latorre-Román[3], Juan A. Párraga-Montilla[4], Luis E. Roche-Seruendo[5]**

**Affiliations**

1 Department of Physical Education, Sports and Recreation, Universidad de La Frontera, Temuco, Chile
2 Sport and Physical Education, Universidad de Granada, Granada, Spain
3 Department of Didactics of Music, Plastic and Corporal Expression, University of Jaén, Jaén, Spain
4 Physical Activity and Sports Science, University of Jaen, Jaen, Spain
5 Universidad San Jorge Facultad de Ciencias de la Salud, Health Sciences, Villanueva de Gallego, Spain

**Key words**

endurance, foot pod, kinetics, runners

**accepted** 29.05.2019

**Bibliography**

**DOI** https://doi.org/10.1055/a-0946-2159
Published online: 2019
Int J Sports Med
© Georg Thieme Verlag KG Stuttgart · New York
ISSN 0172-4622

**Correspondence**

Felipe García-Pinillos, PhD
Department of Physical Education,
Sports and Recreation,
Universidad de La Frontera (Temuco, Chile),
Calle Uruguay, 1980 Temuco, Chile
fegarpi@gmail.com

## ABSTRACT

This study aimed to examine how the power output changes while running at a continuous comfortable velocity on a motorized treadmill by comparing running power averaged during different time intervals. Forty-nine endurance runners performed a running protocol on a treadmill at self-selected comfortable velocity. Power output (W) was estimated with the Stryd™ power meter, and it was examined over six recording intervals within the 3-min recording period: 0–10 s, 0–20 s, 0–30 s, 0–60 s, 0–120 s and 0–180 s. The ANOVAs showed no significant differences in the magnitude of the power output between the recording intervals ($p = 0.276$, $F = 1.614$, partial $Eta^2 = 0.155$). An almost perfect association was also observed in the magnitude of the power output between the recording intervals ($ICC \geq 0.999$). Bland-Altman plots revealed no heteroscedasticity of error for the power output in any of the between-intervals comparisons ($r^2 < 0.1$), although longer recording intervals yield smaller systematic bias, random errors, and narrower limits of agreement for power output. The results show that power data during running, as measured through the Stryd™ system, is a stable metric with negligible differences, in practical terms, between shorter (i. e., 10, 20, 30, 60 or 120 s) and longer recording intervals (i. e., 180 s).

## Introduction

The inclusion of power meters in cycling revolutionized both training and competition [1], and the same may soon be accomplished for runners. Traditionally, the use of this metric during running was restricted to laboratory settings, because it was difficult to estimate. However, in recent years with the advent of inertial motion sensors, power can be estimated during running, and some projects are focused on its use and application with runners. A recent study by García-Pinillos et al. [2] reported a strong linearity for the power-velocity relationship during running in a wide range of submaximal intensities. Another study, Aubry et al. [3] examined the relationship between power and metabolic demand during running, whereas Austin et al. [4] examined the correlation between running power and running economy at comfortable velocity.

Today, some devices provide feedback on power during running (e. g., RunScribe™, Stryd™ or Myotest™), even real-time feedback. Therefore, the limitation is not how to collect the data but how to interpret the data, and some methodological considerations might have an influence on that point (e. g., accuracy, reliability or duration of data collection). So far, studies analyzing running power through the aforementioned wearables [2–4] averaged power output data collected during intervals with different durations. For example, power data has been averaged in previous studies for 3-min [2], 2-min [3] or 1-min stages [4]. In this context, a question arises: at a given running velocity and in no-fatigue conditions, how does the power during running change when measured at different time intervals? The importance of those two conditions (i. e., a given running velocity and in no-fatigue conditions) is noteworthy since previous studies have shown the influence of both running velocity [2] and fatigue on power output [5].

Therefore, the aim of this study was to examine how the power output changes during running at a continuous comfortable velocity on a motorized treadmill, by comparing running power averaged during different time intervals. It is hypothesized that, under those conditions, power output would keep constant over time.

# Materials and Methods

## Subjects

A group of 49 amateur endurance runners, 44 men and 5 women (n = 49; age: 26 ± 8 years; height: 1.74 ± 0.07 m; body mass: 71 ± 10 kg) voluntarily participated in this study. All participants met the inclusion criteria: (i) older than 18 years old, (ii) able to run 10 km in less than 50 min, (iii) not suffering from any injury in the last 6 months before the data collection. This study meets the ethical standards of the IJSM [6]. After receiving detailed information on the objectives and procedures of the study, each subject signed an informed consent form in order to participate, which complied with the ethical standards of the World Medical Association's Declaration of Helsinki (2013). It was made clear that the participants were free to leave the study, if they saw fit. The study was approved by the Institutional Review Board.

## Procedures

Participants were individually tested on one specific day. Prior to all testing, subjects refrained from severe physical activity for at least 48 h, and all tests were conducted at least 3 h after eating. Tests were performed with the subjects' usual training shoes to measure their typical performance.

Participants performed a running protocol on a motorized treadmill (WOODWAY Pro XL, Woodway, Inc., Waukesha, WI, USA). The initial speed was set at 8 km·h$^{-1}$, and speed increased by 1 km·h$^{-1}$ every minute until participants felt comfortable. Then, running velocity was fixed (i. e., self-selected comfortable running velocity: 11.7 ± 1.3 km·h$^{-1}$). Since previous studies [7, 8] on human locomotion have shown that accommodation to running on a treadmill occurs in ~6–8 min, an 8-min accommodation program was performed at that self-selected velocity. Immediately after the accommodation interval, the recording period started. It lasted 3 min, and it was performed at the same running velocity. The slope was maintained at 0 % over the entire protocol.

## Materials and testing

For descriptive purposes, body height (cm) and body mass (kg) were determined using a precision stadiometer and weighing scale (SECA 222 and 634, respectively, SECA Corp., Hamburg, Germany). All measurements were taken with the participants wearing underwear.

Power output (in W) was estimated with the Stryd™ power meter (Stryd Power meter, Stryd Inc. Boulder CO, USA). Stryd™ is a relatively new carbon fiber-reinforced foot pod (attached to the shoe) that weights 9.1 grams. Based on a 6-axis inertial motion sensor (3-axis gyroscope, 3-axis accelerometer), this device provides twelve metrics to quantify performance: pace, distance, elevation, running power, form power, cadence, ground contact time, vertical oscillation, leg stiffness. To the best of the authors' knowledge, just one study has examined its validity and reliability (in this case, to measure spatiotemporal gait characteristics [9]), with no data to demonstrate the validity and reliability of this device for measuring power and related variables. For this study, only two out of twelve metrics were used (running velocity and power output). Those variables were obtained from Stryd's website (https://www.stryd.com/powercenter/analysis) in the .fit file. Then, data were analyzed using the publically available software (Golden Cheetah, version 3.4) and exported as .csl files. Those files were imported from Excel® (2016, Microsoft, Inc., Redmond WA) and laps were done every 3 min. Power output was examined over six recording intervals within the 3-min recording period: 0–10 s, 0–20 s, 0–30 s, 0–60 s, 0–120 s and 0–180 s. Mean and standard deviation (SD) were calculated for those intervals.

## Statistical analysis

Descriptive statistics are represented as mean (SD). The normal distribution of data and homogeneity of variances were confirmed through the Kolmogorov-Smirnov and Levene´s tests, respectively ($p > 0.05$). One-way repeated measures ANOVA, with Bonferroni post-hoc corrections, were conducted on the magnitude of power output to examine possible differences between the recording intervals (0–10 s, 0–20 s, 0–30 s, 0–60 s, 0–120 s, 0–180 s). The association of the magnitude of the power output between the recording intervals was quantified through the intraclass correlation coefficient (ICC). Based on the characteristics of this experimental design and following the guidelines reported by Koo and Li [10], the authors decided to conduct a "two-way random-effects" model (ICC [2,k]), "mean of measurements" type, and "absolute" definition for the ICC measurement. The interpretation of the ICC was based on the benchmarks reported by a previous study [11]: ICC < 0 reflects 'poor' reliability, 0–0.20 'slight', 0.21–0.40 'fair', 0.41–0.60 'moderate', 0.61–0.80 'substantial', and > 0.81 'almost perfect' reliability. The Bland-Altman [12] limits of agreement method (mean difference ± 1.96 SD) was used to examine differences in power output between each recording interval (0–10 s, 0–20 s, 0–30 s, 0–60 s and 0–120 s) and the longest interval (0–180 s). Heteroscedasticity of error was defined as an $r^2 > 0.1$ [13]. The level of significance used was $p < 0.05$. Data analysis was performed using the SPSS software (version 21, SPSS Inc., Chicago, Ill).

# Results

The ANOVAs showed no significant differences in the magnitude of the power output between the recording intervals ($p = 0.276$, $F = 1.614$, partial $Eta^2 = 0.155$). An almost perfect association was also observed in the magnitude of the power output between the recording intervals (ICC ≥ 0.999) (▶ **Table 1**).

Bland-Altman plots revealed no heteroscedasticity of error for the power output in any of the between-interval comparisons ($r^2 < 0.1$). Longer recording intervals yield smaller systematic bias and narrower limits of agreement for power output during running on a treadmill at comfortable velocity (▶ **Fig. 1**).

# Discussion

This study aimed at comparing power output during running at a constant velocity obtained from averaging different time intervals (i. e., 10, 20, 30, 60, 120 and 180 s), to examine the influence of the volume of data (i. e., longer intervals equal more data) on the stability over time of the power estimation. The results show that power data during running, as measured through the Stryd™ system, is a stable metric with negligible differences, in practical terms, between shorter (i. e., 10, 20, 30, 60 or 120 s) and longer recording intervals (i. e., 180 s).

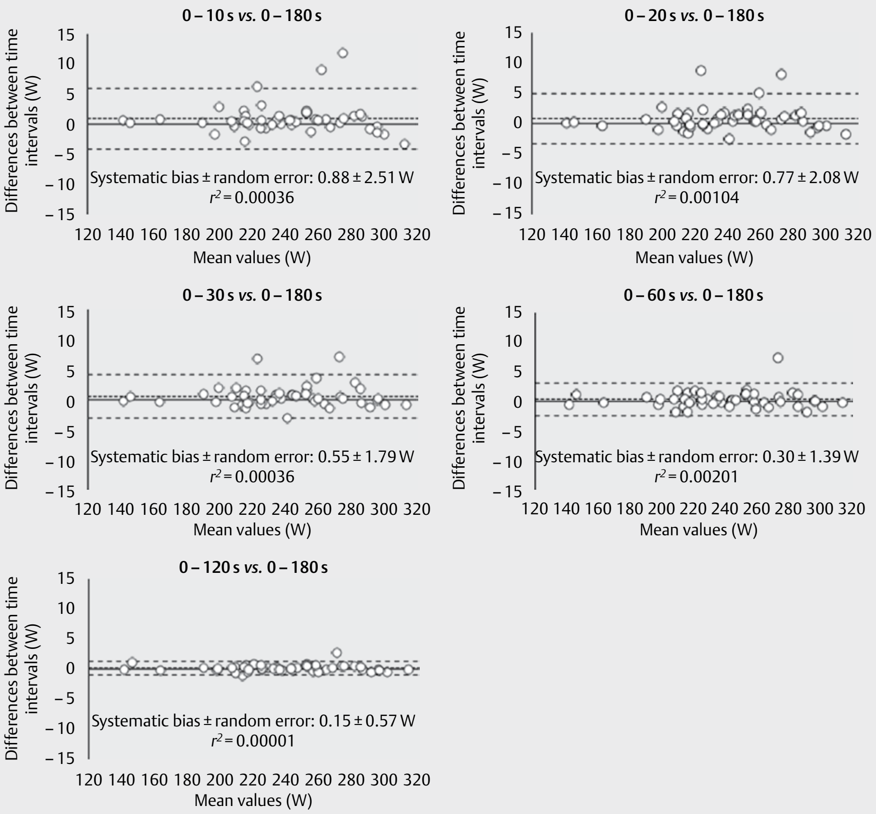
First of all, the context and conditions in which these data were collected must be considered to properly interpret these results. This protocol was performed on a treadmill, at 0 % gradient and at a constant velocity (i. e., self-selected comfortable velocity). All these points create a 'controlled environment' [14] that may influence the stability of the running power over time and make the authors doubt the repeatability of these findings during overground running. There is no consensus in the literature regarding the ability to extrapolate results for adult treadmill running to overground. Specifically, no previous studies addressed an overground vs. treadmill comparison regarding power output during running, but some works have focused on kinetic data (i. e., influencing variables on power data calculation). Whereas a previous work concluded that joint kinematics and ground reaction forces for level running were generally similar between overground and treadmill conditions [15], other researchers found kinetic differences between conditions [16]. Therefore, there are reasons to be cautious about the repeatability of these findings during overground conditions. Also the duration of the protocol, with the recording period lasting 3 min in order to minimize the potential effect of fatigue on the power output. It is well known that fatigue induced might influence the average power output when compared with data collected during longer intervals, since the duration that exercise can be maintained decreases as the power requirements increase, and vice versa [5]. Therefore, this finding cannot be extrapolated to runs in the presence of fatigue.

Despite the repeated measures ANOVA and the ICCs showed an almost perfect agreement between the different time intervals, the information provided by the Bland-Altman plots builds up the interpretation of the results. These plots show that the systematic bias and random error assumed in power output during shorter time intervals (i. e., 10, 20, 30 or 60 s) is higher than during longer intervals (i. e., 120 s), and this should be taken into consideration. As mentioned earlier, the magnitude of the differences between longer and shorter recording intervals was rather small. For example, the systematic bias ± random error in the 0–10 s interval compared to the 0–180 s interval was 0.88 ± 2.51 W, whereas in the

▶ **Table 1** Descriptive values and association between the magnitude of the power output obtained from six time intervals.

| | **Mean (standard deviation)** | | | | | | **ICC (0–10 s vs. 0–180 s)** | **ICC (0–20 s vs.0–180 s)** | **ICC (0–30 s vs. 0–180 s)** | **ICC (0–60 s vs. 0–180 s)** | **ICC (0–120 s vs 0–180 s)** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | **0–10 s** | **0–20 s** | **0–30 s** | **0–60 s** | **0–120 s** | **0–180 s** | **(95 % CI)** | **(95 % CI)** | **(95 % CI)** | **(95 % CI)** | **(95 % CI)** |
| Power | 239.85 | 239.73 | 239.52 | 239.27 | 239.12 | 239.97 | 0.999 | 0.999 | 0.999 | 1.000 | 1.000 |
| (W) | (37.80) | (37.82) | (37.82) | (37.81) | (37.75) | (37.75) | (0.998–0.999) | (0.998–1.000) | (0.999–1.000) | (0.999–1.000) | (1.000–1.000) |

ICC, intraclass correlation coefficient; 95 % CI, 95 % confidence interval.



▶ **Fig. 1** Bland-Altman plot with the power output (W) during running obtained at different time intervals (0–10 s, 0–20 s, 0–30 s, 0–60 s and 0–120 s, compared to 0–180 s). The plot includes the mean difference (dotted line) and 95 % limits of agreement (dashed lined), along with the regression line (solid line). Systematic bias and Pearson's multivariate coefficient of determination ($r^2$) are also presented.

0–120 s vs. 0–180 s comparison was 0.15 ± 0.57 W. Such small-magnitude difference between recording intervals must be gauged in practical contexts, such as when assessing big groups of athletes, or where logistical issues difficult long-lasting assessment protocols, but it must be considered when maximum accuracy is required (e. g., scientific approach).

Finally, some limitations must be addressed. First, these results could be restricted to amateur endurance runners [17] and to a running protocol performed on a treadmill at a comfortable velocity [14]. Second, the footwear was not standardized, but all runners wore their own footwear to increase the ecological validity of the study. There is no evidence about the potential influence of the footwear on the running power but that variable was not controlled in the current study and it must be recognized. Third, the foot strike pattern was not controlled and, even though no previous paper has directly addressed the influence of foot strike pattern on power output, its influence on spatiotemporal gait characteristics is known [18], which might influence power output during running. Fourth, the validity and reliability of the power output data from the Stryd™ system is still unknown. To the best of the authors' knowledge, no previous study has analyzed the concurrent validity of power output during running according to a reference system (e. g., force platform) nor the reliability of this data from different time intervals. Notwithstanding those limitations, the current study provides more information about the use of power output during running at submaximal velocity by comparing the power output during running obtained from averaging different time intervals in 49 amateur endurance runners.

In summary, these results show that power output during running, measured through the Stryd™ system, is stable over time when velocity is constant and under controlled conditions, with no differences between different time intervals recorded during a 3-min run. Nevertheless, it is worth noting that the analysis conducted shows that longer recording intervals yield smaller systematic bias and narrower limits of agreement. Therefore, if maximum accuracy is required (e. g., scientific approach), longer recording periods must be used (i. e., 2–3 min). However, from a practical standpoint, shorter recording intervals (e. g., 10–20 s) might be a more time-efficient option for clinicians or coaches working with big groups of athletes, or where logistical issues make it difficult to conduct long-lasting assessment protocols (e. g., athletes with pain during running).

## Acknowledgements

The authors would like to thank to all the participants. This study was funded by the University of La Frontera (Universidad de La Frontera, Temuco, Chile, Project DFP18-0023).

## Conflict of Interest

The authors declare no conflict of interest.

## References

[1] Passfield L, Hopker J, Jobson S, Friel D, Zabala M. Knowledge is power: Issues of measuring training and performance in cycling. J Sports Sci 2017; 35: 1426–1434

[2] García-Pinillos F, Latorre-Román PÁ, Roche-Seruendo LE, García-Ramos A. Prediction of power output at different running velocities through the two-point method with the Stryd™ power meter. Gait Posture 2019; 68: 238–243

[3] Aubry RL, Power GA, Burr JF. An assessment of running power as a training metric for elite and recreational runners. J Strength Cond Res 2018; 32: 2258–2264

[4] Austin C, Hokanson J, McGinnis P, Patrick S. The relationship between running power and running economy in well-trained distance runners. Sports (Basel) 2018; 6: E142

[5] Burnley M, Jones AM. Power-duration relationship: Physiology, fatigue, and the limits of human performance. Eur J Sport Sci 2018; 18: 1–12

[6] Harriss D, Macsween A, Atkinson G. Standards for ethics in sport and exercise science research: 2018 update. Int J Sports Med 2017; 38: 1126–1131

[7] Schieb DA. Kinematic accommodation of novice treadmill runners. Res Q Exerc Sport 1986; 57: 1–7

[8] Lavcanska V, Taylor NF, Schache AG. Familiarization to treadmill running in young unimpaired adults. Hum Mov Sci 2005; 24: 544–557

[9] García-Pinillos F, Roche-Seruendo LE, Marcen-Cinca N, Marco-Contreras LA, Latorre-Román PA. Absolute reliability and concurrent validity of the Stryd system for the assessment of running stride kinematics at different velocities. J Strength Cond Res 2018; doi: 10.1519/JSC.0000000000002595

[10] Koo TK, Li MY. A Guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016; 15: 155–163

[11] Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977; 33: 159–174

[12] Bland JM, Altman DG. Comparing methods of measurement: Why plotting difference against standard method is misleading. Lancet 1995; 346: 1085–1087

[13] Atkinson G, Nevill AM. Statistical methods for assessing measurement error (Reliability) in variables relevant to sports medicine. Sport Med 1998; 26: 217–238

[14] Riley PO, Dicharry J, Franz J, Della Croce U, Wilder RP, Kerrigan DC. A kinematics and kinetic comparison of overground and treadmill running. Med Sci Sports Exerc 2008; 40: 1093–1100

[15] Firminger CR, Vernillo G, Savoldelli A, Stefanyshyn DJ, Millet GY, Edwards WB. Joint kinematics and ground reaction forces in overground vs. treadmill graded running. Gait Posture 2018; 63: 109–113

[16] Rozumalski A, Novacheck TF, Griffith CJ, Walt K, Schwartz MH. Treadmill vs. overground running gait during childhood: A qualitative and quantitative analysis. Gait Posture 2015; 41: 613–618

[17] Nakayama Y, Kudo K, Ohtsuki T. Variability and fluctuation in running gait cycle of trained runners and non-runners. Gait Posture 2010; 31: 331–335

[18] Di Michele R, Merni F. The concurrent effects of strike pattern and ground-contact time on running economy. J Sci Med Sport 2014; 17: 414–418