©Journal of Sports Science and Medicine (2018) **17**, 305-311
http://www.jssm.org

Research article

Validity and Reliability of the PowerTap P1 Pedals Power Meter

Jesús G. Pallarés ✉ and José Ramón Lillo-Bevia
Human Performance and Sports Science Laboratory, University of Murcia, Murcia, Spain

Abstract
To validate the new PowerTap P1® pedals power meter (PP1), thirty-three cyclists performed 12 randomized and counterbalanced graded exercise tests (100–500 W), at 70, 85 and 100 rev·min^{-1} cadences, in seated and standing positions. A scientific SRM system and a pair of PP1 pedals continuously recorded cadence and power output data. Significantly lower power output values were detected for the PP1 compared to the SRM for all workloads, cadences, and pedalling conditions (2–10 W, $p < 0.05$), except for the workloads ranged between 150 W to 350 W at 70 rev·min^{-1} in seated position ($p > 0.05$). Strong Spearman's correlation coefficients were found between the power output values recorded by both power meters in a seated position, independently from the cadence condition (rho ≥ 0.987), although slightly lower concordance was found for the standing position (rho = 0.927). The mean error for power output values were 1.2%, 2.7%, 3.5% for 70, 85 and 100 rev·min^{-1}, respectively. Bland-Altman analysis revealed that PP1 pedals underestimate the power output data obtained by the SRM device in a directly proportional manner to the cyclist's cadence (from -2.4 W to -7.3 W, rho = 0.999). High absolute reliability values were detected in the PP1 pedals (150–500 W; CV = 2.3%; SEM < 1.0 W). This new portable power meter is a valid and reliable device to measure power output in cyclists and triathletes for the assessment, training and competition using their own bicycle, although caution should be exercised in the interpretation of the results due to the slight power output underestimation of the PP1 pedals when compared to the SRM system and its dependence on both pedalling cadence and cyclist's position (standing vs. seated).

Key words: Cycling, mobile power meter, testing, cycle ergometer, power output.

Introduction

Mobile power meters became commercially available in the 1980's, allowing direct measurement of power output in field conditions (Nimmerichter et al., 2017). Since then, scientists, coaches and cyclists have been able to measure bicycle power output during cycling training and competition, as traditionally performed in a laboratory setting. The SRM power meter soon became the *gold standard* of the mobile power meters. It consists of a crankset that allows the measurement of torque via strain gauges, located between the crank and the chain rings, and angular velocity from the cadence. Therefore, power output is calculated as the product of torque and angular velocity.

Several valid and reliable laboratory-specialized ergometers and power meters have been developed so far to monitor exercise performance while cycling: Lode (Earnest et al., 2005; Reiser et al., 2000), Ergoline (Maxwell et al., 1998), Monark (MacIntosh et al., 2001; Maxwell et al., 1998), Velotron (Abbiss et al., 2009; Astorino and Cottrell, 2012), Wattbike (Hopker et al., 2010; Wainwright et al., 2017). It should be noted that it is not possible to use them for field testing, moreover their size, weight and especially their price, could make difficult their use in laboratories with low financial resources (Peiffer and Losco, 2011). In addition, even if the cyclist can customize the position of the ergometer´s handlebars, saddle and pedals (not always possible), there would be considerable variations with their own bicycles in some decisive metrics such as the crank width (Q-factor), crank length, and other differences related to the specific geometry of the bicycle itself, which could affect comfort, pedalling performance and might even increase injury incidence (Disley and Li, 2014).

Currently, there are some mobile power meters whose validity and reliability have been confirmed, such as Garmin Vector (Bouillod et al., 2016; Nimmerichter et al., 2017; Novak and Dascombe, 2016), or PowerTap Hub (Bertucci et al., 2005b; Bouillod et al., 2016; Gardner et al., 2004). Nevertheless, there are others whose results are reproducible, but whose validity remains in question, such as Stages (Bouillod et al., 2016; Granier et al., 2017). Finally, there are others considered unreliable power meters, for example, Look Keo Power pedal (Sparks et al., 2015).

The validity and reliability of power meters is linked to the usefulness of the information obtained, since it is well known that poor reliability in power output measurement does not allow for optimisation of the training program, in comparison with previous or future tests, nor an accurate analysis of the data (Jeukendrup et al., 2000). Changes in performance and training status cannot be determined without a high level of reliability for the measurement of power output (Garcia-Lopez et al., 2016; Hopkins et al., 2001; Jeukendrup et al., 2000; Pallares et al., 2016; Paton and Hopkins, 2001). For the evaluation of the effect of training or detraining with power output measurement, it is important to know the variation due to the technical error of the power meter (Bertucci et al., 2005a). Specifically, Vanpraagh et al. (1992) suggested that the range of the technical error for workload recorded using ergometers should be within 5%. When using a power meter to test high-level athletes, it would be advisable for this technical error to be closer to 2%, due to the fact that elite male cyclists have typical variation of ~1% for time trials lasting ~1 hour (Paton and Hopkins, 2001).

The recent development of the PowerTap P1 pedals (PP1, CycleOps, Madisson, USA) has introduced another mobile power measuring tool to the market with a reduced price (~$999.99). In a similar way to others, this manufacturer claims that the PP1 pedals are accurate to within

Received: 20 November 2017 / Accepted: 16 April 2018 / Published (online): 14 May 2018

1.5%, with a very limited extra weight (~150 g) compared with mid or top range clipless road pedals. They are built with eight strain gauges, which work with a 'Multipole Ring', a sensor made of 20 small magnets around the pedal spindle. It allows cyclists to use their own bicycle in tests or training sessions carried out on laboratory ergometers, indoor trainers, rollers or in the field, by just replacing the pedals. The power measurement comes directly from the point of contact with the bicycle, reducing the loss of power output due to mechanical connections (Jones, 1998).

As far as we know, the PowerTap P1 pedals have not been previously validated. For this reason, the purpose of this study is to examine the validity, reliability and accuracy of a new powermeter placed in the pedals of the bike under laboratory cycling conditions.

Methods

Experimental approach to the problem

A descriptive, cross-sectional, quantitative study was conducted. During a period of three weeks, each participant performed several tests, conducted on separate days, in the same exercise laboratory, under standardized conditions (22.9 ± 2.0 ºC; 39.3 ± 3.0 % humidity). The study, which was conducted according to the declaration of Helsinki, was approved by the Bioethics Commission of the University of the University of Murcia, and written informed consent was obtained from all participants prior to participation.

Participants

Thirty-three well-trained male cyclists and triathletes volunteered to take part in this study (age 32.4 ± 9.0 yr; height 1.86 ± 0.08 m; body mass 78.6 ± 12.9 kg; VO_{2max} 57.7 ± 6.6 ml·kg^{-1}·min^{-1}; maximal aerobic power (MAP) 399 ± 31 W; cycling training experience 11.2 ± 2.7 years). All participants trained for 6 hours or more per week during a minimum of twelve months preceding the study. Participants were asked to avoid strenuous exercise, caffeine and alcohol for at least 24 hours prior to each testing session.

Testing procedures

A brand new PowerTap P1 power meter (CycleOps, Madison, USA) was compared against an SRM crank-based power meter (scientific model with adjustable 7075 Aluminium crank length; Schoberer Rad Messtechnik, Julich, Germany, ±1% accuracy). For all testing sessions, PP1 were mounted on the SRM cranks with the manufacturer-recommended torque. Additionally, a medium size road bicycle (Giant Defy 3, 2010 Giant Bicycles, Taiwan; Aluminium alloy frame with carbon fibre fork) was fitted with the SRM 172.5 mm crank power meter. This precision strain gauge-based crank and sprocket dynamometer transmitted data to a unit display fixed on the handlebars.

The relationship between the frequency output and the strain gauges and torque is determined during manufacture and considered constant. The validity of this SRM system has been previously demonstrated and therefore taken as the gold standard power meter device (Jones, 1998; Martin et al., 1998; Passfield and Doust, 2000). To minimize the possible influence in the validity and reliability values of the PP1 data, the same bicycle and SRM power meter were used in all testing conditions. A dynamic calibration of the SRM crankset was performed by the manufacturer prior to the beginning of the study.

The rear wheel of the bicycle was removed and attached to a direct drive pedalling unit Cycleops Hammer (Cycleops, Wisconsin, EEUU) (Lillo-Bevia and Pallares, 2017) with 10 speed (11–25 tooth) rear gear ratio and 39–53 tooth front gear ratio. For all tests, the gear ratio 39:15 was selected, and cyclists were not allowed to change it to prevent a potential effect of this variable on pedalling technique. Prior to each testing session, the calibration of the Hammer ergometer was carried out according to the manufacturer's recommendations. In this way, the Hammer can accurately determine the power required to overcome bearing and belt friction, and set the zero offset of strain gauges. Furthermore, the zero offset of the PP1 pedals was set before each testing session. Likewise, the front fork of the bicycle was attached to the accompanying steering apparatus for stability purposes. The bicycle seat height position was matched to the cyclist's own training geometry. Cyclists used their own cycling shoes fitted with Look cleats. The absolute and relative validity of this direct drive device has been recently confirmed (Lillo-Bevia and Pallares, 2017).

Testing protocol

All testing protocols began with a standardized warm-up of 5 minutes at 100 W with a freely chosen cadence. Following this period, the validity and reliability of the devices were assessed in the laboratory during three different testing protocols:

All participants performed three randomized and counterbalanced graded exercises tests, one for each selected fixed cadence (70, 85 and 100 rev·min^{-1}), at six submaximal workloads (100, 150, 200, 250, 300 and 350 W) of 75 seconds duration (Jones, 1998). The three graded exercise tests were separated by 5 min of recovery at 75 W, performed in seated position and with freely chosen cadence. The order of the three cadence levels was randomized to ensure that the validity of the results was not affected by increments on the ergometer break temperature or by the cyclists' fatigue. After 5 min of recovery at 75 W, cyclists performed a 75-second seated free cadence 500 W workload. Finally, they performed a graded exercise test at three sub-maximal workloads (250, 350 and 450 W) of 75 seconds with a freely chosen cadence, in a standing pedalling position. Two minutes of recovery at 75 W with freely chosen cadence were kept between the three workloads tested. The pedalling power output was registered by the PP1 and SRM simultaneously.

Following the recommendation of Jones (1998), only power outputs and cadence values from the 10th to the 70th second of each 75-second step were analysed, to allow the Cycleops Hammer enough time to stabilise the assigned breaking workload. During each test, power output (W) and cadence (rev·min^{-1}) of PP1 were recorded at a frequency of 1 Hz using a Garmin 1000 cycling computer (Garmin International Inc., Olathe, KS, USA). Additionally, power output and cadence of the SRM crankset were recorded at a frequency of 1 Hz using the Power Control V.

Statistical analysis

Standard statistical methods were used for the calculation of means, standard deviations (SD), coefficient of variation (CV) and standard error of the mean (SEM). Data were assessed for heteroscedasticity by plotting the predicted vs. the residual values for power and cadence measurements. The Kolmogorov-Smirnov test and complementary analyses of normality were used. The SRM and PP1 power output and cadence data were not normally distributed. Thus, the analysis of differences between the mean of power outputs and cadences values of each device were assessed with a non-parametric Mann-Whitney U test. Spearman's rank order correlation coefficients were calculated comparing the power outputs values of the SRM and the PP1 power meters during every graded exercise test. Additionally, given the fact that a high correlation does not necessarily imply that there is good agreement between any two methods, Bland-Altman plots were used to assess and display the agreement and systematic difference among the SRM and PP1 power outputs values (Bland and Altman, 1999). The power outputs differences were drawn in relation to the mean values and 95% of the differences, which were expected to lie between the two limits of agreement (LoA). LoA was defined as mean bias ± 2 standard deviation (SD) (Atkinson and Nevill, 1998). Statistical significance for all tests was regarded as $p < 0.05$. The recorded data were downloaded from the previously described units and further analysed using publicly available software (Golden Cheetah, version 3.4) and Microsoft Excel 2016 (Microsoft Software). Analyses were performed using GraphPad Prism 6.0 (GraphPad Software, Inc., CA, USA), SPSS software version 19.0 (SPSS, Chicago, IL) and Microsoft Excel 2016 (Microsoft Corp, Redmond, WA, USA).

Results

Validity

No significant differences were detected in power output values between SRM scientific model and PP1 pedals at 70 rev·min^{-1} in seated position for workloads ranged between 150 W to 300 W ($p > 0.05$). However, in the rest of the workloads, cadences, and pedalling positions assessed, significantly lower values were detected in the PP1 compared to the SRM power meter ($p < 0.05$) (Table 1). Nevertheless, high levels of Spearman's correlation coefficients were detected between the power output values recorded by the PP1 and the SRM devices in seated position (rho ≥ 0.987; $p < 0.001$), independently from the cadence condition (70, 85 and 100 rev·min^{-1}). However, for standing pedalling a slightly weaker correlation coefficient was found (rho = 0.927; $p < 0.001$) (Table 1 and Figure 1). Confirming the means difference data analysis, the Bland-Altman analysis revealed low bias, but not negligible, between the power output values of the SRM power meter and PP1 pedals for all seated tests. Specifically, the PP1 pedals underestimated the power output data obtained by the SRM device in a directly proportional manner to the cyclist's pedalling cadence (bias = -2.4 W (LoA -12.1 to 7.3) at 70 rev·min^{-1},

Table 1. Results from the validity and reliability analysis.

			POWER OUTPUT								CADENCE			
			SRM (W)		PT P1 (W)		SEM (W)	Rho Spearman value	Bland Altman		SRM (rpm)		PT P1 (rpm)	
			Mean ±SD	CV	Mean ±SD	CV			Bias (W)	SD Bias (W)	Mean ±SD	CV	Mean ±SD	CV
70 CAD	SITTING	100 W	99±6	5.6%	97±4*	4.2%	0.7	0.989#	-2.4	4.8 LoA (-12.1 to 7.3)	70.4±1.0	1.4%	71.7±1.1	1.5%
		150 W	150±5	3.4%	148±5	3.0%	0.8				70.7±0.9	1.3%	70.7±1.0	1.5%
		200 W	200±5	2.4%	198±4	2.1%	0.7				70.6±1.1	1.5%	70.9±1.1	1.5%
		250 W	251±5	2.0%	248±5	1.9%	0.8				70.7±1.0	1.4%	70.8±1.0	1.3%
		300 W	303±5	1.5%	300±5	1.6%	0.8				70.4±0.9	1.3%	70.9±0.9	1.3%
		350 W	356±4	1.2%	352±5*	1.4%	0.9				70.0±1.0	1.5%	70.6±1.0	1.5%
85 CAD	SITTING	100 W	101±6	5.9%	96±6*	5.7%	1.0	0.987#	-5.3	6.1 LoA (-17.6 to 7.0)	84.7±0.8	0.9%	85.0±0.8	0.9%
		150 W	149±6	4.0%	145±5*	3.7%	0.9				84.7±0.8	0.9%	84.8±0.8	0.9%
		200 W	201±6	2.7%	196±5*	2.7%	0.9				84.8±0.9	1.1%	85.0±0.9	1.1%
		250 W	252±5	1.9%	246±5*	2.2%	0.9				84.8±1.1	1.3%	85.0±1.1	1.8%
		300 W	303±6	1.8%	298±6*	2.0%	1.1				84.9±1.2	1.4%	85.1±1.2	1.4%
		350 W	355±5	1.5%	349±6*	1.7%	1.0				84.9±1.0	1.2%	85.1±1.0	1.7%
100 CAD	SITTING	100 W	96±8	8.6%	91±7*	7.2%	1.1	0.999#	-7.3	7.9 LoA (-23.1 to 8.4)	98.9±1.3	1.3%	99.7±1.2	1.3%
		150 W	145±7	4.9%	139±5*	3.9%	0.9				98.9±1.4	1.4%	99.2±1.5	1.5%
		200 W	197±8	4.1%	191±7*	3.7%	1.2				99.6±1.2	1.2%	99.1±1.2	1.3%
		250 W	248±7	2.9%	241±7*	2.8%	1.2				99.6±1.3	1.3%	99.7±1.3	1.3%
		300 W	298±7	2.4%	291±7*	2.4%	1.2				99.5±1.5	1.6%	99.8±1.6	1.6%
		350 W	352±5	1.9%	342±8*	2.3%	1.3				99.8±1.9	1.9%	99.7±1.9	1.9%
FC	STAND	250 W	253±7	2.6%	241±5*	2.2%	0.9	0.927#	-9.0	5.3 LoA (-19.7 to 1.7)	75.9±6.1	8.0%	74.9±11.0	14.7%
		350 W	352±6	1.8%	345±5*	1.5%	0.9				74.8±9.1	12.1%	73.5±12.9	17.6%
		450 W	455±8	1.7%	446±6*	1.2%	1.0				69.6±7.7	11.1%	68.5±10.7	15.7%
FC	SIT	500 W	499±9	1.8%	492±11*	2.2%	1.9		-7.0	3.5 LoA (-14.1 to 0.0)	90.0±10.1	11.2%	89.8±10.5	11.7%

STAND = Standing; SIT= Sitting; CAD = Cadence; FC = Free cadence; SD = Standard Deviation; CV = Coefficient of variation; rho Spearman = Spearman correlation coefficient; LoA = Limits of Agreement; * Significant differences compared to the SRM device; # significant Spearman correlation coefficient; ($p < 0.05$).

-5.3 W (LoA -17.6 to 7.0) at 85 rev·min^{-1} and -7.3 W (LoA -23.1 to 8.4) at 100 rev·min^{-1}, rho = 0.999). A slightly greater underestimation was found for standing tests (bias = -9.0 W (LoA -19.7 to 1.7)) (Table 1 and Figure 2).

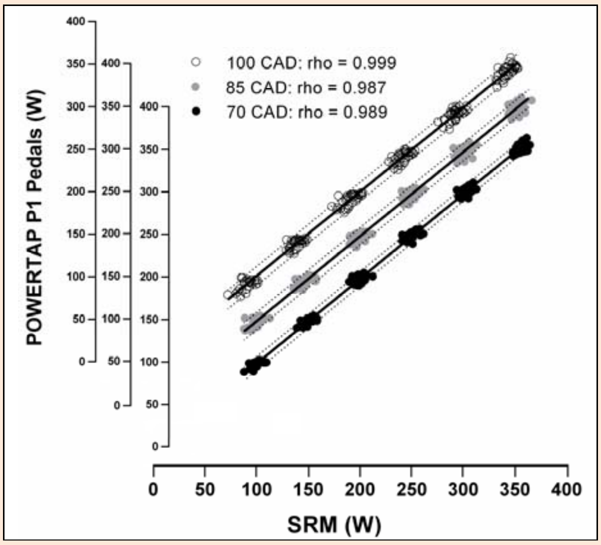

Figure 1. Spearman´s Correlation Coefficient of the PowerTap PP1 pedals under three different cadences, during the submaximal graded exercises tests, compared to the scientific SRM power meter at 70, 85 and 100 rev·min^{-1}.

Reliability

The mean CV for the sitting graded exercise tests were 2.7% vs. 2.4%, 3.0% vs. 3.0% and 4.1% vs. 3.7% for the SRM compared with the PP1 at 70, 85 and 100 rev·min^{-1}, respectively. These values were considerably lower if the 100 W workload was excluded (2.1 vs. 2.0%, 2.4 vs. 2.5% and 3.2 vs. 3.0%). The mean CV for standing pedalling tests of both devices (SRM vs. PP1) were 2.0% vs. 1.6%, respectively, while CV for the high workload (i.e., 500 W) in seated position remained very low (1.8% vs. 2.2%). The SEM for the PP1 remained at very low values for all testing conditions (ranging between 0.7 W and 1.9 W) (Table 1).

Discussion

The main finding of this study is that the PP1 is a highly valid and reliable tool for testing and training purposes in cycling under all assessed workloads (100 W to 500 W), cadences (70, 85 and 100 rev·min^{-1}) and pedalling positions (seated and standing). To our knowledge, this is the first study that validates the PP1, which is a portable power meter with some important advantages with respect to other portable devices such as the use of the cyclist's own bicycle, maintaining the usual riding position and the wheelset and the crankset of the bicycle, the reduced extra weight compared to other high performance portable power meters (installed at crankset or hub), and finally the ease of installation, which allows exchanging it between various bicycles.

Nevertheless, it is important to be conscious that this portable power meter slightly underestimated the power output data in a directly proportional manner to the pedalling cadence (from -2.4 W at 70 rev·min^{-1} to -7.3 W at 100 rev·min^{-1}), independently of the cycling workload or pedalling position. This fact could be due to the strain gauges' sensitivity, or due to the signal processing (amplification, filtering, analog to digital conversion and data analysis).

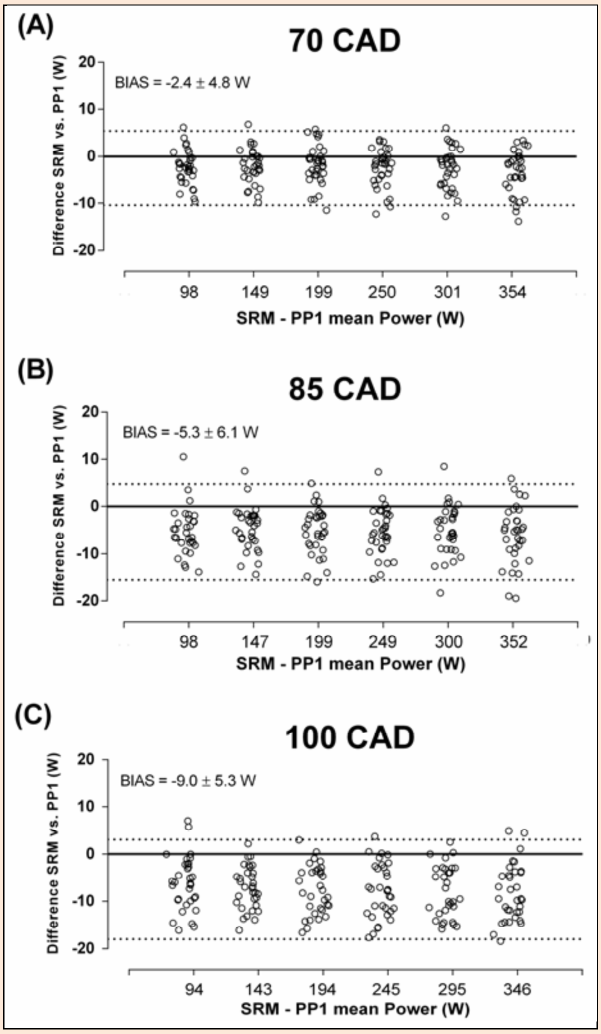

Figure 2. Bland-Altman plots of the PowerTap PP1 pedals, assessed during the submaximal graded exercises tests, compared to the scientific SRM power meter at 70 (A), 85 (B) and 100 (C) rev·min^{-1}.

Laboratory based ergometers (e.g., SRM, Lode, Velotron, Wattbike) are still considered the "gold standard" power meters due to their high levels of validity and reliability (Abbiss et al., 2009; Earnest et al., 2005; Hopker et al., 2010; Hopkins et al., 1999; Jones, 1998; Paton and Hopkins, 2001; Reiser et al., 2000; Wainwright et al., 2017). Thus, for a cycle trainer or power meter to be useful in a research setting it must have similar qualities of measurement. Different researchers have tested the validity of other mobile ergometers such as Tacx Fortius (Peiffer and Losco, 2011), KICKR Power Trainer (Zadow et al., 2017; Zadow et al., 2016), LeMond Revolution (Novak et al., 2015), and Elite Axiom Powertrain (Bertucci et al., 2005a), as well as other mobile power meters, including PowerTap Hub (Bertucci et al., 2005b; Bouillod et al., 2016; Gardner et al., 2004) and Garmin Vector (Bouillod et al., 2016; Nimmerichter et al., 2017; Novak and Dascombe, 2016). It should be noted that the SRM, as the reference power meter, is also affected by some measurement error. Jones (1998) reported extremely low variability (± 0.3% and ± 1.0% for two different 20 strain gauge, and ± 1.8% for a 4 strain gauge models), while the accuracy claimed by the manufacturer of these devices is also very high (± 0.5 % and ± 2.5 %, for the 20 and 4 strain gauge, respectively). Additionally, most of these validation studies have used the SRM scientific model comprising 20 strain gauges (Bertucci et al., 2005a, Duc et al., 2007; Jones, 1998) or the SRM professional model (4 strain gauge) (Gardner et al., 2004; Hurst and Atkins, 2006) as the gold standard devices.

Despite the fact that, according to the data collected in the present study, there are small but significant differences between the mean power output values obtained by the PP1 pedals and the SRM scientific model, there are highly significant, "near perfect", relationships (rho $\geq$ 0.987; $p < 0.001$) from 100 W to 350 W with seated position at low, medium and high cadences. The previous concordance is reduced for standing, freely chosen cadence pedalling (rho = 0.927; $p < 0.001$).

It is also important to note that this study has found very small bias and SD of bias in the agreement between the SRM and PP1 power output data, as well as between SRM and PP1 cadence (from -2.4 ± 4.8 W to -9.0 ± 5.3 W), both for the standing and seated pedalling positions, even though it is known that standing pedalling causes lateral sways and affects the biomechanics of pedalling (Stone and Hull, 1993). These results are consistent but progressive. When used in the laboratory and compared to the SRM crankset, similar mean and SD biases, as well as the 95% limits of agreement data, were reported for other mobile power meters, such as Garmin Vector Pedals (Bouillod et al., 2016; Nimmerichter et al., 2017) (0.6 ± 6.2 W, 11.6 to 12.7 W; -11.6 to 12.7 W, -3.7 to 9.5 W), PowerTap Hub (Bertucci et al., 2005b) (2.9 ± 3.3 W; -3.7 to 9.5 W), and Look Keo Power Pedal (Sparks et al., 2015) (4.6 ± 0.4 W; -15.9 to 13.9 W). Bouillod, et al. (2016) found higher mean and SD biases when the SRM crankset was compared with the Stages (-13.7 ± 12.4 W, -37.9 W to 10.6 W).

Paton and Hopkins (2001) suggested that in elite athletes, a magnitude lower than 2% is required to detect changes in performance from an ergogenic or training intervention. Besides, Hopkins (2000) suggested that an 84% confidence interval is a more reasonable threshold than the traditional 95% interval when attempting to detect changes in athletic performance. Based on a workload of 350 W, changes of $\geq$ 2% (7.0 W) and $\geq$ 1% (3.5 W) would be required to be confident (84%) that a trained cyclist had changed power output because of a training intervention. When compared to the SRM, the mean error of the PP1 shows that, in the present data, it falls within this range. Based on the current study's evaluation of the PP1, a mean error of ~2% compared to the SRM would be acceptable for talent identification purposes. These results suggest that the PP1 power meter is sufficiently accurate to track performance changes over time, and thus would serve as an acceptable training tool. Regarding reliability (Table 1), when we compare the PP1 with other mobile power meters from previous studies, mean CVs are similar to these findings. Bertucci et al. (2005b) reported a CV from 1.7 to 2.7% for the PowerTap Hub and from 1.2 to 2.0% for the SRM crankset over a workload range of 100 W to 420 W. The mean CVs reported in laboratory and field trials by Nimmerichter et al. (2017) were 0.95 vs. 1.00% and 2.82 vs. 3.05%, for the SRM and a Garmin device, respectively. These results mean that CV for the SRM and the PP1 in the current study concur with the reliability data from previous studies.

Cycling technique and type of ergometer can affect cycling efficiency (Arkesteijn et al., 2013). In our opinion the inclusion of cyclists as participants adds more ecological validity to the real use of the pedals. From this point of view, the reliable results of the current research confirm that this biological variability does not affect the validity of the power output data, nor the cadence, of this power meter. What is more, the pedals and cyclists were tested with three different and representative cadences to analyse if the cadence affects the reliability of the power output and cadence. Besides, the number of participants and their fitness level (i.e., well-trained cyclists) are consistent with other published research studies assessing the reliability and validity of cycle ergometers (Pallares et al., 2016; Passfield and Doust, 2000; Wainwright et al., 2017).

It is important to note that PP1 pedals have some limitations in their use, in spite of the practical advantages they offer. As the SRM power meters are checked for validity and reliability against a first principles dynamic calibration rig, the PP1 pedals cannot be easily checked by this method because of the difficulty of applying a known force dynamically to the pedals. The application of the torque at the bottom bracket will not cause any deflection within pedals axles. On the other hand, a static calibration cannot be performed either due to the fact that the PP1 pedals will not transmit any data to a recording device if a cadence reading is not available (Bini and Hume, 2014). As stated earlier, the best current method to assess the variability of the PP1 pedals was to compare them with a scientific model SRM crankset, which has been shown to be accurate and reliable. Additionally, the slope of the power curve cannot be adjusted, meaning that PP1 will always be limited by the factory calibration. Accordingly, PP1 pedals should be checked regularly against a calibrated scientific SRM crankset. If this process is done regularly, PP1 pedals provide an acceptable method of power output measurement and their use in detecting changes in performance and monitoring external training power output is supported. Since the tests were developed with workloads up to 500 W, additional research must be done to test the reliability and validity of the PP1 for sprint cycling tests above 500 W. Also, further research is needed to evaluate this power meter system in field conditions (Bouillod et al., 2016).

Conclusions

This study confirms that, despite the slight but consistent underestimation found at 85 and 100 rev·min^{-1} (2–7 W), which slightly depends on both pedalling cadence and cyclists' position (i.e., seated vs. standing), this new PP1 is a valid, reliable and accurate mobile power meter, compared with the recognized gold standard scientific SRM, to measure power output and cadence in cyclists. This new portable power meter provides an alternative to more expensive laboratory ergometers, allowing cyclists to use their own bicycle for testing, training or competition purposes. Current results demonstrate that the PP1 provides valid readings of power output from 100 to 500 W, in either seated or standing positions, at cadences of 70, 85 and 100 rev·min^{-1}, or even at freely chosen cadences.

Acknowledgements

The authors wish to thank the participants for their invaluable contribution to the study. The experiments comply with the current laws of Spain. There are no conflicts of interest.

References

Abbiss, C.R., Quod, M.J., Levin, G., Martin, D.T. and Laursen, P.B. (2009) Accuracy of the Velotron Ergometer and SRM Power Meter. *International Journal of Sports Medicine* **30,** 107-112.

Arkesteijn, M., Jobson, S.A., Hopker, J. and Passfield, L. (2013) Effect of Gradient on Cycling Gross Efficiency and Technique. *Medicine and Science in Sports and Exercise* **45,** 920-926.

Astorino, T.A. and Cottrell, T. (2012) Reliability and Validity of the Velotron Racermate Cycle Ergometer to Measure Anaerobic Power. *International Journal of Sports Medicine* **33,** 205-210.

Atkinson, G. and Nevill, A.M. (1998) Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Medicine* **26,** 217-238.

Bertucci, W., Duc, S., Villerius, V. and Grappe, F. (2005a) Validity and reliability of the axiom powertrain cycle ergometer when compared with an SRM powermeter. *International Journal of Sports Medicine* **26,** 59-65.

Bertucci, W., Duc, S., Villerius, V., Pernin, J.N. and Grappe, F. (2005b) Validity and reliability of the PowerTap mobile cycling powermeter when compared with the SRM device. *International Journal of Sports Medicine* **26,** 868-873.

Bini, R.R. and Hume, P.A. (2014) Assessment of Bilateral Asymmetry in Cycling Using a Commercial Instrumented Crank System and Instrumented Pedals. *International Journal of Sports Physiology and Performance* **9,** 876-881.

Bland, J.M. and Altman, D.G. (1999) Measuring agreement in method comparison studies. *Statistical Methods in Medical Research* **8,** 135-160.

Bouillod, A., Pinot, J., Soto-Romero, G., Bertucci, W. and Grappe, F. (2016) Validity, Sensitivity, Reproducibility and Robustness of the Powertap, Stages and Garmin Vector Power Meters in Comparison With the SRM Device. *International Journal of Sports Physiology and Performance* 1-26.

Disley, B.X. and Li, F.X. (2014) The effect of Q Factor on gross mechanical efficiency and muscular activation in cycling. *Scandinavian Journal of Medicine & Science in Sports* **24,** 117-121.

Duc, S., Villerius, V., Bertucci, W. and Grappe, F. (2007) Validity and Reproducibility of the Ergomo (R) Pro Power Meter Compared With the SRM and Powertap Power Meters. *International Journal of Sports Physiology and Performance* **2,** 270-281.

Earnest, C.P., Wharton, R.P., Church, T.S. and Lucia, A. (2005) Reliability of the lode excalibur sport ergometer and applicability to computrainer electromagnetically braked cycling training device. *Journal of Strength and Conditioning Research* **19,** 344-348.

Garcia-Lopez, J., Diez-Leal, S., Ogueta-Alday, A., Larrazabal, J. and Rodriguez-Marroyo, J.A. (2016) Differences in pedalling technique between road cyclists of different competitive levels. *Journal of Sports Sciences* **34,** 1619-1626.

Gardner, A.S., Stephens, S., Martin, D.T., Lawton, E., Lee, H. and Jenkins, D. (2004) Accuracy of SRM and power tap power monitoring systems for bicycling. *Medicine and Science in Sports and Exercise* **36,** 1252-1258.

Granier, C., Hausswirth, C., Dorel, S. and Yann, L.M. (2017) Validity And Reliability Of The Stages Cycling Power Meter. *Journal of Strength and Conditioning Research*, 2017-Sep-06. Epub ahead of print.

Hopker, J., Myers, S., Jobson, S.A., Bruce, W. and Passfield, L. (2010) Validity and Reliability of the Wattbike Cycle Ergometer. *International Journal of Sports Medicine* **31,** 731-736.

Hopkins, W.G. (2000) Measures of reliability in sports medicine and science. *Sports Medicine* **30,** 1-15.

Hopkins, W.G., Hawley, J.A. and Burke, L.M. (1999) Design and analysis of research on sport performance enhancement. *Medicine and Science in Sports and Exercise* **31,** 472-485.

Hopkins, W.G., Schabort, E.J. and Hawley, J.A. (2001) Reliability of power in physical performance tests. *Sports Medicine* **31,** 211-234.

Hurst, H.T. and Atkins, S. (2006) Agreement between Polar and SRM mobile ergometer systems during laboratory-based high-intensity, intermittent cycling activity. *Journal of Sports Sciences* **24,** 863-868.

Jeukendrup, A.E., Craig, N.P. and Hawley, J.A. (2000) The bioenergetics of World Class Cycling. *Journal of science and medicine in sport* **3,** 414-433.

Jones, S.M.P., L. (1998) Dynamic calibration of bicycle power measuring cranks. *The Engineering of Sport* 265-274.

Lillo-Bevia, J. and Pallares, J.G. (2017) Validity and Reliability of the Cycleops Hammer Cycle Ergometer. *International Journal of Sports Physiology and Performance* **28,** 1-19.

MacIntosh, B.R., Bryan, S.N., Rishaug, P. and Norris, S.R. (2001) Evaluation of the Monark Wingate Ergometer by direct measurement of resistance and velocity. *Canadian Journal of Applied Physiology-Revue Canadienne De Physiologie Appliquee* **26,** 543-558.

Martin, J.C., Milliken, D.L., Cobb, J.E., McFadden, K.L. and Coggan, A.R. (1998) Validation of a mathematical model for road cycling power. *Journal of Applied Biomechanics* **14,** 276-291.

Maxwell, B.F., Withers, R.T., Ilsley, A.H., Wakim, M.J., Woods, G.F. and Day, L. (1998) Dynamic calibration of mechanically, air- and electromagnetically braked cycle ergometers. *European Journal of Applied Physiology and Occupational Physiology* **78,** 346-352.

Nimmerichter, A., Schnitzer, L., Prinz, B., Simon, D. and Wirth, K. (2017) Validity and Reliability of the Garmin Vector Power Meter in Laboratory and Field Cycling. *International Journal of Sports Medicine* **38(6),** 439-446.

Novak, A.R. and Dascombe, B.J. (2016) Agreement of Power Measures between Garmin Vector and SRM Cycle Power Meters. *Measurement in Physical Education and Exercise Science* **20,** 167-172.

Novak, A.R., Stevens, C.J. and Dascombe, B.J. (2015) Agreement between LeMond Revolution cycle ergometer and SRM power meter during power profile and ramp protocol assessments. *Journal of Science and Cycling* **4,** 37-43.

Pallares, J.G., Moran-Navarro, R., Fernando Ortega, J., Emilio Fernandez-Elias, V. and Mora-Rodriguez, R. (2016) Validity and Reliability of Ventilatory and Blood Lactate Thresholds in Well-Trained Cyclists. *Plos One* **11(9)**, e0163389.

Passfield, L. and Doust, J.H. (2000) Changes in cycling efficiency and performance after endurance exercise. *Medicine and Science in Sports and Exercise* **32,** 1935-1941.

Paton, C.D. and Hopkins, W.G. (2001) Tests of cycling performance. *Sports Medicine* **31,** 489-496.

Peiffer, J.J. and Losco, B. (2011) Reliability/Validity of the Fortius Trainer. *International Journal of Sports Medicine* **32,** 353-356.

Reiser, M., Meyer, T., Kindermann, W. and Daugs, R. (2000) Transferability of workload measurements between three different types of ergometer. *European Journal of Applied Physiology* **82,** 245-249.

Smak, W., Neptune, R.R. and Hull, M.L. (1999) The influence of pedaling rate on bilateral asymmetry in cycling. *Journal of Biomechanics* **32,** 899-906.

Sparks, S.A., Dove, B., Bridge, C.A., Midgley, A.W. and McNaughton, L.R. (2015) Validity and Reliability of the Look Keo Power Pedal System for Measuring Power Output During Incremental and Repeated Sprint Cycling. *International Journal of Sports Physiology and Performance* **10,** 39-45.

Stone, C. and Hull, M.L. (1993) Rider bicycle interaction loads durinig standing treadmill cycling. *Journal of Applied Biomechanics* **9,** 202-218.

Vanpraagh, E., Bedu, M., Roddier, P. and Coudert, J. (1992) A simple calibration method for mecanically braked cycle ergometers. *International Journal of Sports Medicine* **13,** 27-30.

Wainwright, B., Brian Cooke, C. and Paul O'Hara, J. (2017) The validity and reliability of a sample of 10 Wattbike cycle ergometers. *Journal of Sports Sciences* **35(14),** 1451-1458.

Zadow, E.K., Kitic, C.M., Wu, S.S.X. and Fell, J.W. (2017) Reliability of Power Settings of the Wahoo KICKR Power Trainer After 60 Hours of Use. *International International Journal of Sports Physiology and Performance* 1-13.

Zadow, E.K., Kitic, C.M., Wu, S.S.X., Smith, S.T. and Fell, J.W. (2016) Validity of Power Settings of the Wahoo KICKR Power Trainer.

International Journal of Sports Physiology and Performance **11,** 1115-1117.

Key points

- PP1 pedals slightly underestimate power output at medium to high cadences (2 to 7 W).
- PP1 pedals provide valid readings of power output from 100 to 500 W, in either seated or standing positions, at fixed cadences of 70, 85 and 100 rev·min^{-1}, or even at freely chosen cadences.
- These results suggest that the new PP1 pedals is a valid, reliable and accurate mobile power meter to measure power output and cadence in cyclists using their own bicycles.

AUTHOR BIOGRAPHY

Jesús G PALLARÉS
Employment
Professor of Exercise Physiology (Human Performance and Sport Science Laboratory, University of Murcia, Spain)
Degree
PhD
Research interest
Exercise physiology and training; performance analysis; ergogenic aids.
E-mail: jgpallares@um.es

Jose Ramon LILLO-BEVIA
Employment
PhD student at Faculty of Sport Sciences, University of Murcia, Murcia, Spain
Degree
MSc, PhD candidate
Research interests
Exercise physiology and training, performance analysis
E-mail: joseramon.lillo@um.es

✉ **Jesús García Pallarés**
Faculty of Sport Science, Argentina S/N, Santiago de la Ribera, Murcia, Spain