Validity and Reliability of the Wattbike Cycle Ergometer

Authors J. Hopker[1], S. Myers[2], S. A. Jobson[1], W. Bruce[1], L. Passfield[1]

Affiliations [1] Centre for Sports Studies, University of Kent, Chatham, United Kingdom
[2] University of Chichester, Faculty of Sport Education and Social Sciences, Chichester, United Kingdom

Key words
- cycling
- ergometry
- calibration
- power output
- testing
- talent identification

Abstract

The purpose of this study was to assess the validity and reliability of the Wattbike cycle ergometer against the SRM Powermeter using a dynamic calibration rig (CALRIG) and trained and untrained human participants. Using the CALRIG power outputs of 50–1 250 W were assessed at cadences of 70 and 90 rev.min^{-1}. Validity and reliability data were also obtained from 3 repeated trials in both trained and untrained populations. 4 work rates were used during each trial ranging from 50–300 W. CALRIG data demonstrated significant differences ($P<0.05$) between SRM and Wattbike across the work rates at both cadences. Significant differences existed in recorded power outputs from the SRM and Wattbike during steady state trials (power outputs 50–300 W) in both human populations (156±72 W vs. 153±64 W for SRM and Wattbike respectively; $P<0.05$). The reliability (CV) of the Wattbike in the untrained population was 6.7% (95%CI 4.8–13.2%) compared to 2.2% with the SRM (95%CI 1.5–4.1%). In the trained population the Wattbike CV was 2.6% (95%CI 1.8–5.1%) compared to 1.1% with the SRM (95%CI 0.7–2.0%). These results suggest that when compared to the SRM, the Wattbike has acceptable accuracy. Reliability data suggest coaches and cyclists may need to use some caution when using the Wattbike at low power outputs in a test-retest setting.

Introduction

The Wattbike is a newly developed air-braked ergometer endorsed by British Cycling for talent identification and support for their world class programmes. It calculates power output via the use of a load cell located next to the chain. As the chain runs over the load cell, it calculates the sum of all the forces applied to the chain through the cranks. The Wattbike is designed to be used by both competitive cyclists and the general public exercising within a gym environment. The ergometer also allows 'online' time-trial racing which is used for talent identification. Therefore, the validity and reliability of the Wattbike is of interest to those working in the field, and it is pertinent to investigate these factors across a wide range of power outputs.

The SRM Powermeter has long been considered as the 'gold-standard' power measuring device due to its validity and reliability [1,5,10,12]. The SRM Powermeter is a crankset that calculates power output as the product of torque x angular velocity, where torque is measured via the deformation of strain gauges located between the crank axle and chainrings. Angular velocity is calculated from the cadence of the cyclist.

Variability within an ergometer could arise from both systematic and random error sources. Systematic error refers to a consistent bias or offset in the reading of power provided by an ergometer, whereas random error refers to fluctuations in readings from measurement to measurement [9,15]. Systematic error in static ergometers has previously been investigated with the use of a dynamic calibration rig driving either the crank or bottom bracket [13,16,18]. Random errors affect the repeatability of the ergometer. Large random errors limit the ability to compare repeated tests over time and the ability to track changes in physiological parameters such as exercise thresholds (i.e. ventilatory and lactate), heart rate zones. Random variation also affects the validity of fitness scores, often used in talent identification.

The aims of this study were: 1. To compare the agreement between the Wattbike cycle ergometer and previously validated SRM Powermeter

accepted after revision
June 20, 2010

Bibliography
DOI http://dx.doi.org/10.1055/s-0030-1261968
Published online: July 27, 2010
Int J Sports Med 2010; 31: 731–736 © Georg Thieme Verlag KG Stuttgart · New York
ISSN 0172-4622

Correspondence
James Hopker
University of Kent
Centre for Sport Studies
Chatham Maritime
ME4 4AG Medway
United Kingdom
Tel.: +44/1634/888814
Fax: +44/1634/888809
j.g.hopker@kent.ac.uk

during mechanical trials, human steady state trials and human performance trials; 2. To investigate the reliability of the Wattbike cycle ergometer by performing repeat measures with both trained and untrained participants.

Methods

Each test was performed on the Wattbike ergometer (Wattbike Ltd, Nottingham, UK), from which the original crankset was removed and the bottom bracket replaced to allow an SRM Powermeter (Science model, SRM, Jülich, Germany) to be fitted. The SRM Powermeter (SRM) fitted had the same size crank arm (170 mm) chainring (48 teeth) as the original Wattbike crankset. The Wattbike chain was used and was configured to follow the same alignment to ensure the correct functioning of the Wattbike load cell. This set-up allowed data to be collected simultaneously from the SRM and the Wattbike. At the start of the study the SRM had been calibrated through a first principles approach by specialists at British Cycling according to the methods of Wooles et al. [17]. Prior to each trial the zero was calibrated for both the Wattbike ergometer and SRM in accordance with manufacturers' recommendations. Wattbike cadence sensors were left in situ. Power output was recorded and averaged over 5 s intervals. The Wattbike calculates power output by measuring the chain tension over a load cell (sampled at 100 Hz). The Wattbike calculates power output using the formula:

$P[W] = (F[N] * l[m])/t[s]$

Where P[W] – Power output per revolution, F[N] – Average force per crank revolution, l[m] = 0.17 m as a crank length, t[s] – time taken to complete a crank revolution. The Wattbike measures angular velocity twice per crank revolution.

The study was completed in 2 parts. Firstly, a motorised calibration rig (Vacumed Ergometer Calibrator Model 17 801, Ventura, CA) was used to drive the ergometer at crank cadences of 70 and 90 rev.min^{-1}. The mechanical calibrator was used to facilitate the application of a constant turning force (cadence) to the crank, rather than to set a power output. A range of power outputs (50–1 250 W) were achieved by manually varying the resistance settings on the Wattbike rather than using the dynamic calibration rig. Data was concurrently recorded using both the SRM and the Wattbike. Power output was increased by 50 W every 3 min, however only data in the final minute of each stage was used for analysis purposes. Cadences of both 70 and 90 rev.min^{-1} were used at power outputs up to 700 W. Above this power output only cadences of 90 rev.min^{-1} could produce the values required.

In the second part of the study, 10 trained cyclists (mean ± SD: 36 ± 7 yr; 1.78 ± 0.08 m; 72 ± 6 kg, Maximum Power Output (MPO) 412 ± 36 W; $\dot{V}O_{2max}$ 60.4 ± 7.2 mL · kg^{-1} · min^{-1}) and 10 untrained individuals (mean ± SD: 25 ± 5 yr; 1.79 ± 0.06 m; 76 ± 8 kg, MPO 323 ± 33 W; $\dot{V}O_{2max}$ 50.8 ± 7.5 mL · kg^{-1} · min^{-1}) volunteered to take part in the study. The untrained individuals cycled regularly as part of a fitness regimen, but had no training or competitive experience. All participants gave written informed consent before taking part in this study which had local ethics committee approval. The study was also conducted in accordance with the ethical standards of the International Journal of Sports Medicine [6]. Throughout the study, participants served as their own control, maintaining their normal diet and daily activity patterns. All were instructed not to train within the 24 h prior to testing.

Each participant attended the laboratory on 4 separate occasions. The first visit was to perform a test of maximal aerobic power. This test assessed maximal power output (MPO), maximal oxygen uptake (VO_{2max}) and maximal heart rate. Participants completed a 10-minute warm-up at the starting power for the test (100 W). After the warm-up the required work rate increased by 25 W each minute. Participants maintained their freely chosen cadence and continued cycling until volitional exhaustion. On the remaining 3 visits participants completed 4 submaximal work rates. Untrained participants completed work rates of 50, 100, 150 and 200 W at a cadence of 70 rev.min^{-1}, whilst trained participants cycled at work rates of 150, 200, 250 and 300 W at a cadence of 90 rev.min^{-1}. Each submaximal work rate was maintained for 6 min, with 4 min rest between stages. Work rates were applied in a random order each visit.

After a short rest, participants completed a 5 min performance trial. The participants began from a standing start and were required to sustain the highest average power over 5 min. Mean power output and cadence were recorded for each performance trial. All trials were completed in an air conditioned laboratory.

Statistical analysis

For all test variables, mean (± SD) values were calculated for each method of assessing power output. Data were subsequently assessed for the normality of distribution and heteroscedasticity [14]. Statistical differences in power output and cadence between the Wattbike and SRM were assessed using Wilcoxon Signed Rank Tests across the 3 conditions (mechanical trials, human steady state and performance trials). The 95 % limits of agreement were calculated to assess the agreement between the Wattbike and SRM across the conditions [4]. To assess random error within the Wattbike, the within-subject variation, expressed as a coefficient of variation (CV), was derived from log-transformed data [8]. The 95 % confidence intervals were calculated for each CV. A significant difference was set at $P < 0.05$.

Results

Wattbike vs. SRM comparative data

Mechanical trials

14 separate power outputs (50–700 W) were used with the calibration rig driving the ergometer at a set cadence of 70 rev.min^{-1}. Wattbike power output was significantly different ($P < 0.05$) from that recorded on the SRM at each power output using a cadence of 70 rev.min^{-1}. However, a strong correlation between power output recorded by the Wattbike and SRM system ($r = 0.99$; $P > 0.001$) was found (**Fig. 1a**). The 95 % limits of agreement between the Wattbike and SRM were −12 to 14 W (**Fig. 1b**). Cadence was not significantly different between the 2 systems across the power outputs (mean difference = 0 ± 1 rev.min^{-1}; $P > 0.05$). 24 separate power outputs (100–1 250 W) were used at a cadence of 90 rev.min^{-1}. Wattbike power output was significantly different from the SRM across the range of set wattages ($P < 0.05$), except at powers of 100, 550 and 600 W ($P > 0.05$).

Fig. 2a demonstrates a near perfect correlation ($r = 0.99$; $P < 0.01$) between the Wattbike and SRM power outputs across the range of work rates examined at 90 rev.min^{-1}. However the Bland-Altman plot clearly illustrates the differences in recorded power outputs between the 2 devices across the same range (**Fig. 2b**). The 95 % limits of agreement were −13 to 27 W. Mean cadence was not significantly different between the 2 systems (mean difference = 1 ± 1 rev.min^{-1}; $P > 0.05$).

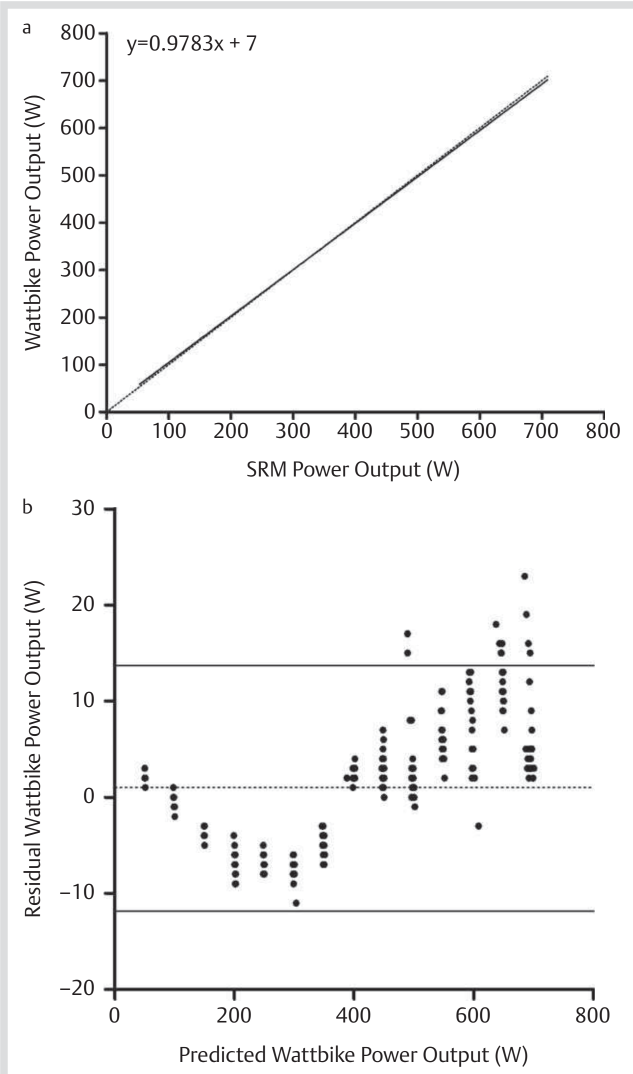

Fig. 1 **a.** Regression of SRM power (watts) on Wattbike power (watts) at 70 rev.min^{-1}. Power measured by the SRM was related to the power recorded by the Wattbike. The resultant regression line is displayed as a solid line. The hashed line displays the 'line of unity'. **b.** Bland-Altman plot of the difference in power output between the SRM and Wattbike systems at 70 rev.min^{-1}; 331 separate 5 s data points have been used. Dashed line is the mean bias (+1 W) and solid lines are the 95 % limits of agreement.

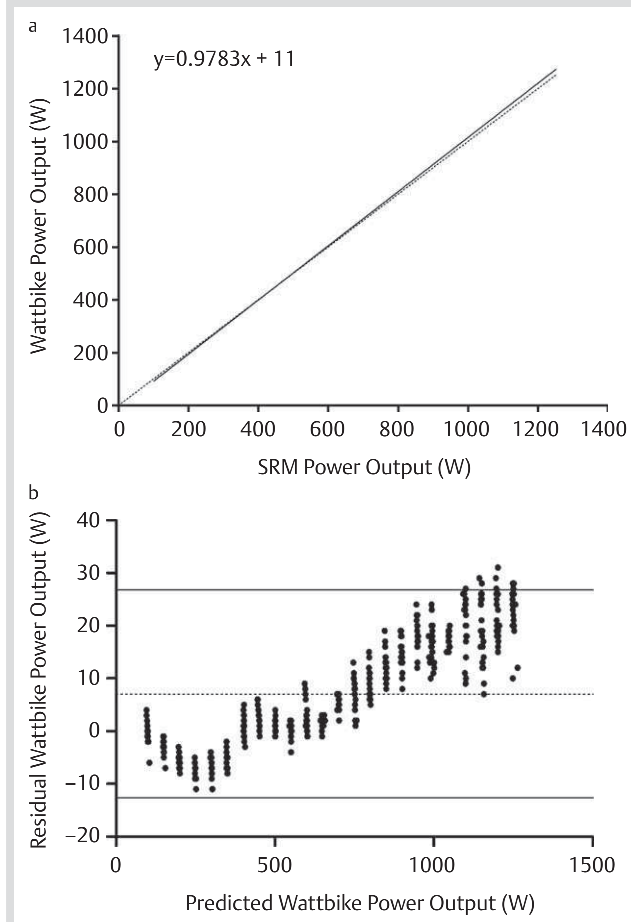

Fig. 2 **a.** Regression of SRM power (watts) on Wattbike power (watts) at 90 rev.min^{-1}. Power measured by the SRM was related to the power recorded by the Wattbike. The resultant regression line is displayed as a solid line. The hashed line displays the 'line of unity'. **b.** Bland-Altman plot of the difference in power output between the SRM and Wattbike systems at 90 rev.min^{-1}; 557 separate 5 s data points have been used. Dashed line is the mean bias (+7 W) and solid lines are the 95 % limits of agreement.

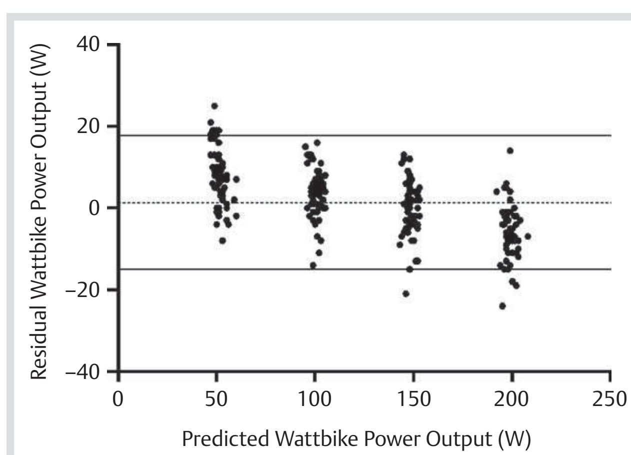

Fig. 3 Bland-Altman plot of the difference in power output between the SRM and Wattbike systems in an untrained population; n = 240 power outputs. Dashed line is the mean bias (+1 W) and solid lines are the 95 % limits of agreement.

Wattbike versus SRM steady-state human trials

Significant differences existed in recorded power outputs from the SRM and Wattbike during steady state trials (power outputs 50–300 W) in both trained and untrained populations (156 ± 72 W vs. 153 ± 64 W for SRM and Wattbike respectively; $P < 0.05$). Specifically, in the untrained population at the lower power outputs of 50 and 100 W the Wattbike recorded power outputs that were significantly higher than the SRM (Mean difference Wattbike – SRM: 8 ± 7 W at 50 W $P < 0.01$; 4 ± 6 W at 100 W $P < 0.01$). There was no significant difference between the 2 systems at 150 W (0 ± 7 W; $P > 0.05$). However, at 200 W the Wattbike power output was significantly lower than the SRM (−6 ± 7 W $P < 0.01$). ▶ **Fig. 3** shows the agreement between the measurement of power outputs from the Wattbike and SRM. The 95 % limits of agreement between the Wattbike and SRM for the untrained participants were −22 to 6 W at 50 W, −15 to 8 W at 100 W, −13 to 13 W at 150 W and −7 to 19 W at 200 W.

Significant differences also existed between power outputs in the trained population with Wattbike power output being significantly lower than SRM power output at all work rates (−4±6W at 150W; −10±5W at 200W; −14±4W at 250W; −20±6W at 300W). The 95% LoA for the trained population between the Wattbike and SRM were −16 to 8W at 150W, −20 to 1 at 200W, −22 to −6 at 250W and −31 to −9 at 300W (**Fig. 4**).

Wattbike cadence and SRM cadence were significantly different at all work rates in both untrained and trained populations (mean cadence across all work rates; untrained: 70±3 vs. 69±2 rev.min^{-1}; $P<0.01$; 95% LoA, −4.22 to 2.14 rev.min^{-1}; trained: 91±1 vs. 90±2 rev.min^{-1}; $P<0.01$; 95% LoA, −3.27 to 1.33 rev.min^{-1}).

Reliability data

3 repeated trials were used to establish the random error associated with the Wattbike and SRM. Across all work rates used the Wattbike demonstrated greater random variability than the SRM within both the untrained (**Table 1**) and trained (**Table 2**) populations. Repeated measures ANOVA identified significant differences across the repeated trials in the Wattbike power output of untrained cyclists at 50W and 100W ($P<0.01$). No significant differences were found within the data of the trained cyclists ($P=0.87$). No differences were found between the repeated trials of the SRM power output data in either untrained ($P=0.14$) or trained groups ($P=0.26$).

Recorded cadence on the Wattbike showed a random variation (CV) across all work rates of 3.9% (95% CI 3.3–5.2%) vs. 2.9% for the SRM (95% CI 2.4–3.8%) in the untrained population and 1.3% (95% CI 1.1–1.8) vs. 1.6% (95% CI 1.3–2.1%) for the SRM in the trained population.

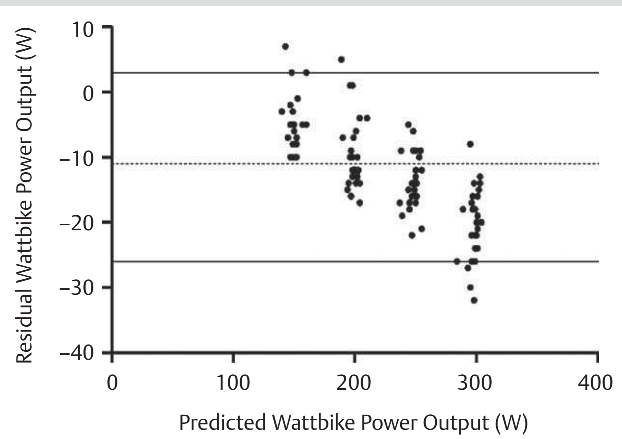

Fig. 4 Bland-Altman plot of the difference in power output between the SRM and Wattbike systems in a trained population; n=120 power outputs. Dashed line is the mean bias (−11W) and solid lines are the 95% limits of agreement.

Performance trial

A significant difference was found between Wattbike and SRM power outputs across the repeated 5-min performance trials in both untrained and trained populations. Mean power output of the untrained population was significantly lower from the Wattbike compared to the SRM (234±30W vs. 239±35W respectively; $P<0.01$; 95% LoA −21 to 11W). No significant differences existed in recorded cadence between the 2 systems (93±5 vs. 92±4 rev.min^{-1}; $P=0.13$; 95% −6 to 8 rev.min^{-1}). With the higher power outputs used by the trained cyclists the mean differences between the 2 systems were greater (Wattbike: 310±32W vs. SRM: 339±38W; $P=0.03$; LoA, −4 to 62W). Recorded cadence was again not significantly different (105±8 vs. 104±3 rev.min^{-1}; $P=0.29$).

Discussion

The purpose of this investigation was to assess the validity and reliability of the Wattbike cycle ergometer. The simultaneous data collected in both mechanical and human trials suggests that the Wattbike recording of power output is significantly different from that of the SRM across a range of power outputs up to and including 1250W. Specifically, at power outputs up to ~550W the Wattbike tends to under predict the SRM. At power outputs above ~550W the Wattbike records power output higher than the SRM. These findings were consistent during both mechanical and human trials.

	Wattbike			SRM		
Power output	**CV(%)**	**Lower CI (%)**	**Upper CI (%)**	**CV(%)**	**Lower CI (%)**	**Upper CI (%)**
50W	12.0	8.6	24.5	4.8	3.4	9.2
100W	6.2	4.4	12.0	1.6	1.1	2.9
150W	4.8	3.4	9.2	1.3	0.9	2.4
200W	3.8	2.6	7.2	1.0	0.7	1.9
mean	6.7	4.8	13.2	2.2	1.5	4.1

Table 1 Reliability indices for untrained participants at steady state power output. Coefficient of variation (CV) of log transformed data. Lower and upper 95% confidence intervals of the CV are also shown.

	Wattbike			SRM		
Power output	**CV (%)**	**Lower CI (%)**	**Upper CI (%)**	**CV(%)**	**Lower CI (%)**	**Upper CI (%)**
150W	3.0	2.1	5.9	1.4	0.9	2.6
200W	3.2	2.2	6.4	1.0	0.7	2.0
250W	2.2	1.5	4.4	1.2	0.8	2.3
300W	1.8	1.2	3.5	0.6	0.4	1.2
mean	2.6	1.8	5.1	1.1	0.7	2.0

Table 2 Reliability indices for trained participants at steady-state power output. Coefficient of variation (CV) of log transformed data. Lower and upper 95% confidence intervals of the CV are also shown.

When being driven mechanically by a dynamic calibration rig at cadences of 70 and 90 rev.min^{-1} there were significant differences between the Wattbike and SRM systems across the majority of power outputs measured. Our results indicate that the Wattbike has an agreement of ±1.7%, in relation to the SRM, between power outputs of 50 and 700 W at 70 rev.min^{-1}. At 90 rev.min^{-1} and power outputs between 100 and 1 250 W the Wattbike has an agreement of ±1.4% with the SRM.

Significant differences were also present between the Wattbike and SRM during steady state trials conducted by both trained cyclists and untrained individuals. At the low power outputs of 50 and 100 W, the Wattbike recorded significantly higher values than the SRM. As power output increased above 150 W, the Wattbike under reported power output compared to the SRM power meter. This was also the case during the 5-min performance trials, where the Wattbike significantly ($P<0.05$) under reported SRM power output in both untrained and trained populations (−5 W and −29 W, respectively).

Fig. 5 illustrates the relationship between data collected from the mechanical trials, trained cyclists and untrained individuals. The between system differences are remarkably similar despite the very different methods used to evaluate the Wattbike.

Results for the SRM in the current study are in agreement with reliability data from previous studies. Specifically, mean reliability of the SRM across power outputs from 50–200 W in untrained cyclists was 2.2% and 1.1% for trained cyclists across power outputs 150–300 W. Kirkland et al. [11] found a mean CV of 2.31% across 137 separate power outputs in a group of trained cyclists. Similarly, Bertucci et al. [3] showed a mean CV of 1.7% for an SRM (scientific model) crank set over power outputs 100–420 W; again this is comparable to the findings of the current study.

Error within the Wattbike system could also be obtained from differences in pedal cadence. Data collected from both systems during mechanical trials indicate that there were no significant differences (mean difference 0–1 rev.min^{-1}) in the recorded cadence between the two systems. However, recorded cadence was significantly different during the steady-state trials in both trained and untrained populations. It could be speculated that

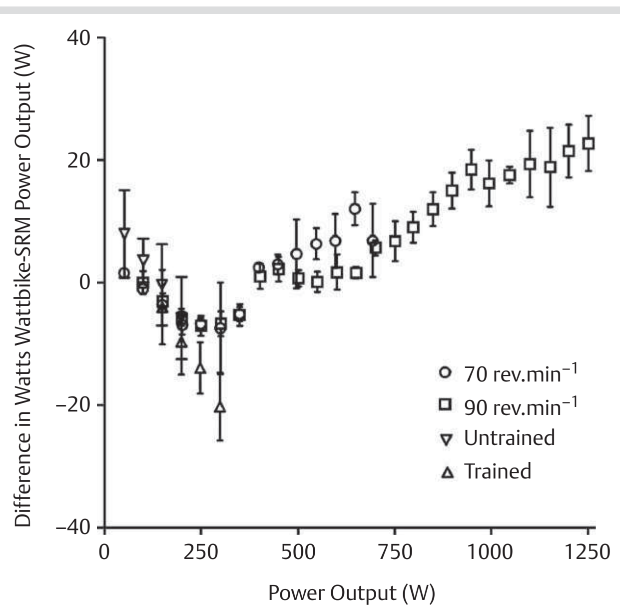

Fig. 5 Combined data from mechanical and human trials across power outputs of 50–1 250 W.

this slight discrepancy in cadence might be because the two systems are not synchronised at their point of measurement. For example, the SRM only measures angular velocity once per crank revolution, compared to 2 samples from the Wattbike. Averaging angular velocity from 2 samples per revolution will not necessarily provide the same angular velocity, and thus cadence, as a single measurement taken once over the same period. When removing cadence from the power output equation and considering torque alone, the difference between the two systems remains. For example, calculated Wattbike torque, across the range of power outputs at 90 rev.min^{-1}, continues to both under and over predict the SRM (100 W: 10.54 vs. 10.64 Nm; 300 W: 30.81 vs. 30.87 Nm; 600 W: 64.00 vs. 65.05 Nm; 900 W: 98.27 vs. 97.48 Nm; 1 200 W: 112.89 vs. 112.27 Nm, from the Wattbike and SRM respectively).

It could be suggested that by changing the original configuration of the Wattbike (by removing the crank set and replacing it with an SRM crank set) we may have altered the factory calibration. However, the original chain was used and the SRM chain ring size and crank length replicated the standard Wattbike configuration. We also zero adjusted the Wattbike as recommended by the manufacturer prior to each trial. If the calibration was affected we would expect to see a uniform error across the range of work rates examined. However, as can be seen in **Fig. 5**, this was clearly not the case. The Wattbike reports both above and below that of the SRM power output across 3 separate conditions. This variability could not be accounted for by calibration errors in the Wattbike.

Differences in recorded power output between the two systems might be due to the way they measure force. The Wattbike calculates force via the use of chain tension over a load cell, whereas the SRM measures torque from strain gauge deformation in the crank. It is possible that friction between the chain and load cell may cause temperature related changes within the Wattbike load cell and, therefore, its measurement of force. The SRM would not be affected by these frictional forces within the chain as it directly measures torque from force applied to the pedals. It might also be the case that the differences in the recorded power output between the two systems are simply because they are situated at different places within the drive chain. Data comparing the SRM and Power Tap powermeters (located at the bottom bracket and rear hub of the bicycle, respectively), suggest the different location may lead to values approximately 2% lower in the Power Tap because of transmission losses in the chain and sprocket drive mechanism [5]. The Wattbike load cell is located part way between these 2 points and thus may be partially affected by drive chain losses.

3 repeated steady state trials were used to calculate reliability coefficients for the Wattbike and SRM in both trained and untrained populations. These trials indicated a greater variability in power output recorded from the Wattbike during concurrent data collection. It can be seen from **Tables 1, 2** that the mean CV% for the Wattbike was outside of the upper confidence interval for the SRM on all but one of the power outputs (trained 250 W), suggesting that it is not as reliable at measuring power output. However, the low CV values for power output from the Wattbike still indicate that it can provide reliable power output measurements. The mean CV for Wattbike power output in trained cyclists was 2.6%. This is in line with reliability coefficients reported for other commercially available power measuring devices [3,11]. Some variability in power output during these trials may have been due to the participants not being able

to exactly maintain the required power output, rather than there being a difference in the calibration on a trial-to-trial basis. To accommodate for this 'human variation' we used the mechanical calibration rig as discussed above, thereby ensuring an exact cadence was maintained throughout the mechanical trials.

It is important for coaches and sport scientists working with cyclists to be confident that the power measuring device they are using is accurate. Hopkins and colleagues have suggested that, in elite athletes, the required detectable change in performance from an ergogenic or training intervention should be of a magnitude of less than 2% [7,8]. When compared to the SRM, the mean error of the Wattbike ergometer demonstrated in our data falls within this range. However, this is largely due to the nature of the S-shaped relationship between Wattbike and SRM power readings. The largest error in power output when driven by the mechanical CALRIG at 90 rev.min^{-1} and 150 W was 4.5%. This degree of precision may be unacceptable for testing the elite population, especially when absolute magnitudes are considered. For example, the mean error of the Wattbike power output compared to the SRM at 1 250 W was only 1.8%, however, in absolute terms the Wattbike measures higher than the SRM by 23 W. Even so, this level of agreement is comparable with that of other commercially available power measuring devices [1,2,5,11].

Agreement between ergometers is important and needs to be addressed for the accurate assessment between Wattbikes at the same, and different laboratory locations. Research is required to assess inter-Wattbike agreement especially as one of its principle uses is for talent identification, with bikes located in schools, gyms and cycling centres across the world. Based on the current study's evaluation of one Wattbike, a mean error of <2% compared to the SRM would be acceptable for talent identification purposes. Although it is important to acknowledge that this error might be greater at very low or high power outputs which is supported by the reliability results from the current study illustrating a ~3% between trial variability. The Wattbike comparative data is also strengthened by the reliability results from the current study which illustrate ~3% between trial variability. Hopkins [8] 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 power output of 300 W, changes of >2% (5 W) and >1% (3 W) would be required to be confident (84%) that a trained cyclist had changed power as a result of a training intervention for the Wattbike and SRM systems, respectively. For an untrained population at a power output of 200 W, changes of >2% (~5 W) and >1% (~2%) would be required. These results suggest that the Wattbike is sufficiently accurate to track performance changes over time and thus would serve as an acceptable training tool for both trained and untrained populations.

Conclusion

In conclusion, the results of the present study suggest that the Wattbike ergometer provides close agreement to power output measurements across a range of power outputs compared to the SRM power meter. However, the level of agreement varied according to the level of power output. Data from the mechanical tests suggests that the Wattbike is less accurate at high power outputs (>700 W). Both the SRM and Wattbike have been shown to have good reliability (>50 W) in repeated trials undertaken by both trained and untrained participants.

Acknowledgement

We wish to thank Professor Lars McNaughton for the loan of the mechanical calibration rig used in this study.

References

1 *Abbis CR, Quod MJ, Levin G, Martin DT, Laursen PB.* Accuracy of the Velotron ergometer and SRM power meter. Int J Sports Med 2009; 30: 107–112
2 *Balmer J, Davison RCR, Bird SR.* Peak power predicts performance power during an outdoor 16.1-km cycling time trial. Med Sci Sports Exerc 2000; 32: 1485–1490
3 *Bertucci W, Duc S, Villerius V, Pernin JN, Grappe F.* Validity and reliability of the PowerTap mobile cycling Powermeter when compared with the SRM device. Int J Sports Med 2005; 26: 868–873
4 *Bland JM, Altman DG.* Statistical methods for assessing agreement between 2 methods of clinical measurement. Lancet 1986; 1: 307–310
5 *Gardner AS, Stephens S, Martin DT, Lawton E, Lee H, Jenkins D.* Accuracy of SRM and Power Tap power monitoring systems for bicycling. Med Sci Sports Exerc 2004; 36: 1252–1258
6 *Harriss DJ, Atkinson G.* International Journal of Sports Medicine – Ethical Standards in Sport and Exercise Science Research. Int J Sports Med 2009; 30: 701–702
7 *Hopkins W, Hawley JA, Burke LM.* Design and analysis of research on sport performance enhancement. Med Sci Sports Exerc 1999; 31: 472–485
8 *Hopkins W.* Measures of reliability in sports medicine and science. Sports Med 2000; 30: 1–15
9 *Hopkins WG, Schabort EJ, Hawley JA.* Reliability of power in physical performance tests. Sports Med 2001; 31: 211–234
10 *Jones SM, Passfield L.* The dynamic calibration of bicycle power measuring cranks. In: Haake SJ (ed.). The Engineering of Sport. Oxford: Blackwell; 1998; 265–274
11 *Kirkland A, Coleman D, Wiles JD, Hopker J.* Validity and reliability of the Ergomo pro Powermeter. Int J Sports Med 2008; 29: 913–916
12 *Martin JC, Milliken DL, Cobb JE, McFadden KL, Coggan AR.* Validation of a mathematical model for road cycling power. J Appl Biomech 1998; 14: 276–291
13 *Maxwell BF, Withers RT, Ilsley AH, Wakim MJ, Woods GF, Day L.* Dynamic calibration of mechanically, air- and electromagnetically braked cycle ergometers. Eur J Appl Physiol 1998; 78: 346–352
14 *Nevill AM.* Why the analysis of performance variables recorded on a ratio scale will invariably benefit from a log transformation (editorial). J Sports Sci 1997; 15: 457–458
15 *Paton C, Hopkins W.* Tests of cycling performance. Sports Med 2001; 31: 489–496
16 *Russell JC, Dale JB.* Dynamic torquemeter calibration of bicycle ergometers. J Appl Physiol 1986; 61: 1217–1220
17 *Wooles AL, Robinson AJ, Keen PS.* A static method for obtaining a calibration factor for SRM bicycle power cranks. Sports Eng 2005; 8: 137–144
18 *Wilmore JH, Constable SH, Stanforth PR, Buono MJ, Tsao YW, Roby FB, Lowdon BJ, Ratliff RA.* Mechanical and physiological calibration of 4 cycle ergometers. Med Sci Sports Exerc 1982; 14: 322–325