Citation for published version

Kirkland, A.I. and Coleman, Damian A and Wiles, Jonathan and Hopker, James G. (2008) Validity and reliability of the Ergomopro powermeter. International journal of sports medicine, 29 (11). pp. 913-6. ISSN 0172-4622.

DOI

https://doi.org/10.1055/s-2008-1038621

Link to record in KAR

http://kar.kent.ac.uk/36048/

Document Version

Publisher pdf

Validity and Reliability of the Ergomo®pro Powermeter

Authors A. Kirkland[3], D. Coleman[1], J. D. Wiles[1], J. Hopker[2]

Affiliations
[1] Sport Science Tourism and Leisure, Canterbury Christ Church University, Canterbury, United Kingdom
[2] Centre for Sport Studies, University of Kent, Medway, United Kingdom
[3] Physiological Assessment, Scottish Institute of Sport, Stirling, United Kingdom

Key words
- bilateral
- ergometry
- reliability
- SRM

Abstract

The aim of this investigation was to assess the validity and reliability of the Ergomo®pro powermeter. Nine participants completed trials on a Monark ergometer fitted with Ergomo®pro and SRM powermeters simultaneously recording power output. Each participant completed multiple trials at power outputs ranging from 50 to 450 W. The work stages recorded were 60 s in duration and were repeated three times. Participants also completed a single trial on a cycle ergometer designed to assess bilateral contributions to work output (Lode Excaliber Sport PFM). The power output during the trials was significantly different between all three systems, ($p < 0.01$) 231.2 ± 114.2 W, 233.0 ± 112.4 W, 227.8 ± 108.8 W for the Monark, SRM and Ergomo®pro system, respectively. When the bilateral contributions were factored into the analysis, there were no significant differences between the powermeters ($p = 0.58$). The reliability of the Ergomo®pro system (CV%) was 2.31% (95% CI 2.13 – 2.52%) compared to 1.59% (95% CI 1.47 to 1.74%) for the Monark, and 1.37% (95% CI 1.26 – 1.50%) for the SRM powermeter. These results indicate that the Ergomo®pro system has acceptable accuracy under these conditions. However, based on the reliability data, the increased variability of the Ergomo®pro system and bilateral balance issues have to be considered when using this device.

Introduction

Portable power measuring systems are now relatively popular devices used to monitor cycling performance during training and competition [4]. These devices allow researchers an insight into the mechanical demands of elite cycling competition [21 – 23], and have been used to assess power output during laboratory-based research studies [3, 7, 8, 19, 24]. The first commercially available and most commonly used device, the SRM powermeter, has been used for this purpose due to the validity and reliability values reported [14]. Newer devices have also been developed; Gardner et al. [11] and Bertucci et al. [4] assessed the PowerTap (CycleOps, Madison, WI 53711, USA) system and reported that this device can be considered as a valid and reliable device for power output measurements during road cycling and in submaximal laboratory tests. However, other systems have not demonstrated suitable validity scores when scientifically evaluated, Millet et al. [16], and Hurst et al. [13] both concluded the Polar S710 system was not appropriate in conditions where high validity was required, indicating vibration, exercise intensity and pedal cadence all influenced the power output values.

In addition to the validity of power measurement, a low mass for all equipment is desirable, as this reduces the forces acting to slow the cyclist down. Therefore, when compared to lightweight componentry, the additional mass at the rear wheel (PowerTap) or at the chainset (SRM) could potentially limit the use of these devices. Furthermore, contractual sponsorship agreement for wheels or chainsets may also limit their use amongst professional athletes during competition (personal communication).

A new power measuring device (Ergomo®pro) potentially offsets these problems using a bottom bracket set up on the bicycle. This unit has similar "claimed" accuracy to the most accurate SRM 8-strain-guage model, with the benefit of weighing ~ 0.01 kg more than a regular bottom bracket. This system calculates power by measuring the torsional deformation of the bottom bracket bearing shaft of the bicycle.

accepted after revision
April 4, 2008

Bibliography
DOI 10.1055/s-2008-1038621
Published online May 29, 2008
Int J Sports Med 2008; 29: 913 – 916 © Georg Thieme Verlag KG Stuttgart · New York · ISSN 0172-4622

Correspondence
Damian Coleman, PhD
Canterbury Christ Church University
Sport Science Tourism and Leisure
North Holmes Road
CT1 1QU Canterbury
United Kingdom
Phone: + 44 0 12 27 78 26 39
Fax: + 44 0 12 27 47 04 42
damian.coleman@canterbury.ac.uk

The aim of this study is therefore to determine the validity and reliability of the power output recorded using the Ergomo®pro, using the previously validated SRM powermeter and a Monark Weight cycle ergometer for comparison.

Methods

Study design

Nine competitive male cyclists (mean ± standard deviation 36 ± 9 yr, 1.79 ± 0.08 m, 74.9 ± 10.2 kg) participated in this study. Following university ethical approval, all participants gave written informed consent.

Equipment

Each test was performed on a weight ergometer (Monark 814E, Varberg, Sweden) fitted with an Ergomo®pro powermeter bottom bracket (SG-Sensortechnik GmbH und Co. KG, Mörfelden-Walldorf, Germany) and an SRM research powermeter (SRM, Jülich, Germany). Both devices had been calibrated through first principles by the manufacturer and the SRM had been statically calibrated in our laboratory. Prior to each trial, the SRM powermeter zero offset procedure was conducted.

The Ergomo®pro system calculates power by measuring the torsional deformation of the bottom bracket bearing shaft on the left hand side of the bicycle. This strain is measured by two square-wave generating optical sensors and average torque is determined from the intervals between those signals over one complete revolution of the shaft. Power (P) is calculated by multiplying the power delivered to the left side by two, which is achieved by a bar mounted computer with a wire interface to the sensors using the formula $P = \text{torque} \times \text{angular velocity}$.

Experimental procedures

Participants reported to the laboratory on two occasions in random order. One visit was utilised to assess the bilateral contribution to work output on a cycle ergometer (Excaliber Sport, PFM, Lode, Groningen, NL). Each participant performed an 8-minute cycle trial at an intensity that corresponded to the individuals "heavy" exercise domain [15] which would equate approximately to the typical intensity if racing for approximately 20 minutes. During these trials pedal forces for both right and left pedals were measured via strain gauges in each crank arm at two degree intervals during every pedal revolution. The crank mounted strain gauges were calibrated prior to the study according to the manufacturers' recommendations as well as via a dynamic calibration using known weights and a calibration rig.

The second visit required participants to complete 60-s incremental work stages beginning at a power output value of 50 W. On completion of each work stage, the workload was increased in 25 W increments until the participant could not maintain the required pedal cadence during the test. Each work stage was repeated three times, and between each work stage active recovery (minimum 60 s) was allowed to ensure the completion of the maximum number of trials for each participant, at the higher work rates, longer recovery (up to 120 s) was allowed between work stages to ensure the greatest number of work stages were completed by each participant. The highest work rate for each participant used in the analysis was determined as the highest work stage completed three times; this resulted in work stages from 50 W to 450 W in the analysis. Any incomplete trial was not included in the analysis. In order to effectively maintain these workloads on the Monark cycle ergometer, participants were instructed to maintain a fixed cadence for each work stage, which was dependent upon the load applied to the weights basket of the ergometer.

Statistical analysis

Data files from the SRM and Ergomo software were time aligned. Monark power was determined by multiplying the mean of SRM and Ergomo®pro cadence by the load applied to the basket. The Ergomo®pro data were assessed in two formats, uncorrected (raw data), and corrected (for bilateral imbalance) data.

For all test variables, mean (± s) values were calculated for each method of assessing power output. Prior to all further analysis, data was checked for appropriate test assumptions [9], and heteroscedasticity were assessed using the methods outlined by Nevill [17]. Following checks for distribution, statistical differences between trials ($p < 0.05$) were assessed using a Friedman test. The 95% limits of agreement were calculated to assess the agreement between the three methods of power measurement.

Within-subject variation expressed as a coefficient of variation (CV) was derived by log-transformed two-way analyses of variance as previously described [1]. The methods of Tate and Klett [20] were then used to ascertain optimal confidence intervals for a normal distribution.

Results

In total, 137 workloads were included in the reliability analysis, and 411 (137 × 3) workloads were included in the comparison of the powermeters. The Monark power output was calculated from the mean pedal cadence from the SRM and Ergomo®pro power measuring devices, the mean pedal cadences across all workloads were 80.81 ± 14.88 and 80.54 ± 14.87 $\text{rev} \cdot \text{min}^{-1}$ for the SRM and Ergomo®pro system, respectively. This difference was statistically significant ($p < 0.01$), with 95% limits of agreement of – 0.15 to 0.71 $\text{rev} \cdot \text{min}^{-1}$. Overall, the power output values recorded for the Monark, SRM, and Ergomo®pro systems were 231.2 ± 114.2 W, 233.0 ± 112.4 W, 227.8 ± 108.8 W, respectively. These differences were significant ($p < 0.01$). Data from the Lode ergometry system demonstrated that from the sample of nine cyclists the contribution of work from each limb was 48.89 ± 3.6% (range 43.99 – 55.08%) from the left limb, and 51.11 ± 3.6% (range 44.92 to 56.01%) from the right limb. The Ergomo®pro power values were then corrected establishing a mean power output value of 232.7 ± 111.1 W for the 411 work stages. The corrected data for Ergomo®pro was significantly higher than the Monark system, however there was no significant difference between the SRM and the Ergomo®pro corrected data ($p = 0.58$).

▶ **Figs. 1** to **3** demonstrate the agreement between the three methods of assessing power output; the figures also demonstrate the homoscedastic nature of the error associated with the comparisons between the three systems. The 95% limits of agreement were – 10.54 to 6.97 when comparing Monark power output against SRM; for the uncorrected Ergomo®pro data, the 95% limits of agreement were – 23.13 to 29.90 for the Monark comparison, and – 17.81 to 28.15 for SRM comparison.

For the assessment of reliability, three repeated trials were conducted to establish the random error associated with each piece of equipment. The CV% was 1.59% (95% CI 1.47 to 1.74%) for the

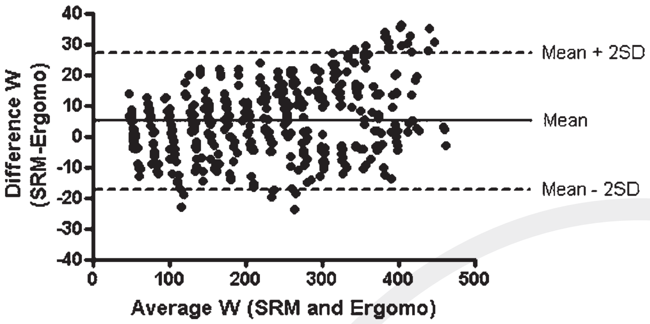

Fig. 1 Agreement between SRM and Ergomo measures.

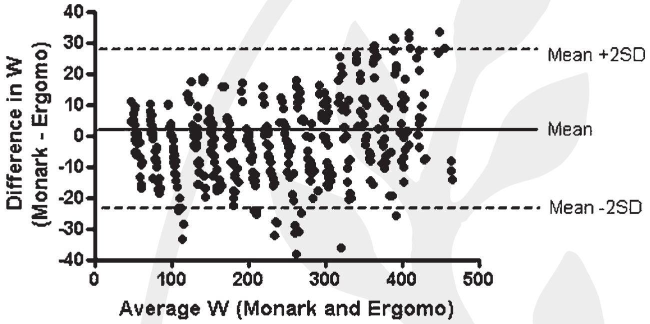

Fig. 2 Agreement between Monark and Ergomo measures.

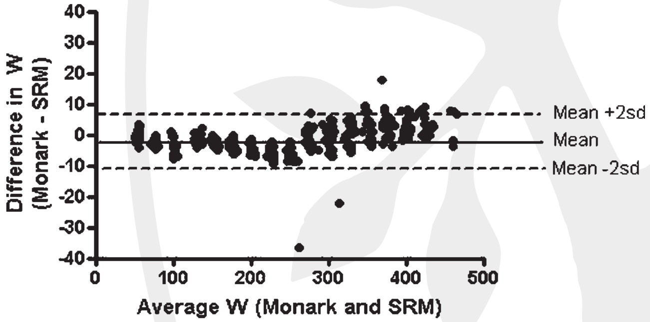

Fig. 3 Agreement between Monark and SRM measures.

Monark, 1.37% (95% CI 1.26 – 1.5%) for the SRM and 2.31% (95% CI 2.13 – 2.52%) for the Ergomo®pro system.

Discussion

The aim of this investigation was to assess the Ergomo®pro power measuring device under controlled laboratory conditions. The simultaneous data collection using the SRM system and the calculation of work achieved on the Monark cycle ergometer suggests that the raw Ergomo®pro system underreads the power output as obtained and predicted by the SRM and Monark systems. There are two potential sources of error that could have contributed to the discrepancy between these pieces of equipment; firstly the pedal cadence is statistically different between the two portable ergometry systems. The limits of agreement for pedal cadence indicate that for a pedal cadence of 80 rev·min^{-1} recorded on the Ergomo®pro system the cadence on the SRM is likely to be 79.85 to 80.71 rev·min^{-1}. Assuming the SRM system measured 230 W at a known cadence of 80 rev·min^{-1}, the lower cadence recorded by the Ergomo®pro system could account for up to ~ 2 W difference between the ergometry systems. Indeed, the method of calculating the Monark power in this study and the comparative SRM values could be adjusted by half of this amount (~ 1 W) due to the lower cadence recorded by the Ergomo®pro system.

The second potential source of error that could contribute to a discrepancy between the ergometry systems is the measurement of the torque on the single-sided Ergomo®pro system. It has previously been reported that the dominant limb provides the greatest torque values during 40-km time trial performance [5], and obviously if any bilateral imbalance were present in an individual it may influence the power output values obtained from a single-sided measurement device. In the current study, assessment of bilateral contribution allowed the experimenters to identify if this could have resulted in any discrepancy between the power measuring systems. Five of the participants had right leg dominance, resulting in a mean left to right leg balance of 48.89 ± 3.6% and 51.11 ± 3.6%. The difference between the Ergomo®pro and SRM system could therefore be accounted for by the multiplication (× 2) of the raw left limb contribution to the overall work attained.

The three repeated reliability trials indicated a greater variability in the recorded power output values on the Ergomo®pro system. The variability was outside the confidence interval of both the other methods of establishing power output, suggesting that the Ergomo®pro system is not as reliable at measuring power output compared to the other systems. One of the major uses of collecting data on the variability of power measuring devices is to use the numbers obtained to inform future work with these devices. Atkinson and Nevill [2] specify that the reliability researcher should specify how reliability analysis influences the interpretation of individual responses. This is particularly pertinent to the sport scientist using a powermeter for scientific support services with an individual athlete. Atkinson and Nevill [2] report that the International Standards Organisation (ISO) advocates using the 95% limits of agreement to indicate the limits that are represented by measurement error, and if changes are outside these limits then the changes are likely to be real. According to Hopkins [12], using 95% limits of agreement provides limits as they are too stringent for a decision limit, he indicates that for elite athletes smaller changes are probably detectable using half of these limits which offers 84% confidence or odds of 5 to 1 that a change has taken place. Based on the calculations of Hopkins [12], for a power output measurement of 230 W, changes of > 3.2% (~ 7 W) and > 1.9% (~ 4 W) would be required to be certain (84%) a change in an individual's power had taken place for the Ergomo®pro and SRM systems, respectively.

The typical use of the portable power measuring device is for the monitoring of training responses of cyclists. The results from this study suggest the two devices could detect relatively small changes in most cyclists mechanical work output. In terms of training with either powermeter, a training "zone" rather than a fixed number would be prescribed in a similar manner to heart rate prescription [25]. The coach would have to be aware of the limits presented in prescribing these zones to minimise overlap if prescribing using thresholds, exercise domains or proportions of maximal capacity. These limits also have to be considered when applying progressive overload to the cyclist.

The data presented on the reliability of these devices can also help to inform the sample size requirement for a particular ex-

Table 1 Sample size calculations based on the methods described by Hopkins 2000 [12], using the 95 % confidence interval for the CV. The smallest worthwhile change (*d*) was derived as 0.2 of the between subject variation using the control trial of Folland et al. (in press) [10] as an example. During this trial, the mean power was 322 W for 16.1 km, with the between subject standard deviation of 15 W. Therefore *d* was calculated to be 3 W or 0.93 %

	95 % CI for the CV%	Participant numbers required for a simple test-retest experiment	Participant numbers required experimental and control group
Ergomo®pro	2.13 – 2.52	42 – 59	168 – 235
SRM	1.26 – 1.5	15 – 21	59 – 84
Monark	1.47 – 1.74	20 – 28	80 – 112

periment as advocated by Hopkins [12]. For a crossover or simple test-retest, study the number of participants required is based around precision defined by 95 % confidence limits (deriving a power 0.8). Hopkins [12] calculates sample size as $n = 8s^2/d^2$, where n is the sample size, *s* is the typical error and *d* is the smallest worthwhile effect. If *d* is presented as a proportion (%) of mean group score then CV% can be inserted for *s*. These figures are changed to $n = 32s^2/d^2$ for a study using an experimental and control group. Using 0.2 of the between subject variation as the smallest worthwhile change (*d*) [6], the sample sizes required for these studies are presented in ▶ **Table 1**.

The analysis in ▶ **Table 1** demonstrates that there would be added costs in terms of resources and participants required if the Ergomo®pro system were to be used in comparison to those established on the SRM and Monark calculations. However, caution is required in the generation of the smallest worthwhile change as other authors indicate there are other approaches to obtaining this number rather than strictly relying on Cohen's 0.2 units. Some authors have selected practically important changes based around prior knowledge of the parameters under investigation [18].

These results indicate that the Ergomo®pro system has acceptable accuracy under laboratory conditions. The differences in the power output values when compared to SRM are accounted for when the bilateral balance of the participants is factored into the analysis. This could, however, potentially limit the use of this system for support and research purposes as an advance ergometry system is required to assess bilateral contributions to the work achieved.

References

1 *Atkinson G, Nevill AM.* Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Med 1998; 26: 217 – 236

2 *Atkinson G, Nevill AM.* Selected issues in the design and analysis of sport performance research. J Sports Sci 2001; 19: 811 – 827

3 *Balmer J, Bird S, Davison RC, Doherty M, Smith P.* Mechanically braked Wingate powers agreement between SRM, corrected and conventional methods of measurement. J Sports Sci 2004; 22: 661 – 667

4 *Bertucci W, Duc S, Villerins V, Pernin JN, Gappe F.* Validity and reliability of the PowerTap mobile cycling powermeter when compared with the SRM Device. Int J Sports Med 2005; 26: 868 – 873

5 *Carpes FP, Rossato M, Faria IE, Bolli Mota C.* Bilateral pedaling asymmetry during a simulated 40-km cycling time-trial. J Sports Med Phys Fitness 2007; 47: 51 – 57

6 *Cohen J.* Statistical Power Analysis for the Behavioural Sciences. 2nd edn. Hillsdale, NJ: Lawence Erlbaum, 1988

7 *Coleman DA, Wiles JD, Davison RC, Smith MF, Swaine IL.* Power output measurement during treadmill cycling. Int J Sports Med 2007; 28: 525 – 530

8 *Coleman DA, Wiles JD, Nunn M, Smith MF.* Reliability of sprint test indices in well-trained cyclists. Int J Sports Med 2005; 26: 383 – 387

9 *Field A.* Discovering Statistics: Using SPSS for Windows. London, UK: Sage, 2000

10 *Folland JP, Stern R, Brickley G.* Sodium phosphate loading improves laboratory cycling time trial performance in trained cyclists. J Sci Med Sport (in press)

11 *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

12 *Hopkins WG.* Measures of reliability in sports medicine and science. Sports Med 2000; 30: 1 – 15

13 *Hurst HT, Atkins S.* Agreement between polar and SRM mobile ergometer systems during laboratory-based high-intensity, intermittent cycling activity. J Sports Sci 2006; 24: 863 – 868

14 *Jones SA, Passfield L.* The dynamic calibration of bicycle power measuring cranks. In: Haake S (ed). The Engineering of Sport. Oxford, UK: Blackwell Science, 1998: 265 – 274

15 *Jones AM, Poole DC.* Oxygen Uptake Kinetics in Sport, Exercise and Medicine. London, UK: Routledge, 2005

16 *Millet GP, Tronche C, Fuster N, Bentley DJ, Candau R.* Validity and reliability of the Polar S710 mobile cycling powermeter. Int J Sports Med 2003; 24: 156 – 161

17 *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

18 *Petersen CJ, Wilson BD, Hopkins WG.* Effects of modified-implement training on fast bowling in cricket. J Sports Sci 2004; 22: 1035 – 1039

19 *Smith MF, Davison RC, Balmer J, Bird SR.* Reliability of mean power recorded during indoor and outdoor self-paced 40 km cycling time-trials. Int J Sports Med 2001; 22: 270 – 274

20 *Tate RF, Klett GW.* Optimal confidence for the variance of a normal distribution. J Am Stat Ass 1959; 54: 674 – 682

21 *Vogt S, Heinrich L, Schumacher YO, Blum A, Roeker K, Dirkhuth HH, Schmid A.* Power output during stage racing in professional road cycling. Med Sci Sports Exerc 2006; 38: 147 – 151

22 *Vogt S, Schumacher YO, Blum A, Roecker K, Dickhuth HH, Schmid A, Heinrich L.* Cycling power output produced during flat and mountain stages in the Giro d'Italia: a case study. J Sports Sci 2007; 25: 1299 – 1305

23 *Vogt S, Schumacher YO, Roecker K, Dickhuth HH, Schoberer U, Schmid A, Heinrich L.* Power output during the Tour de France. Int J Sports Med 2007; 28: 756 – 761

24 *Wiles JD, Coleman DA, Tegerdine M, Swaine IL.* The effects of caffeine ingestion on performance time, speed and power during a laboratory-based 1 km cycling time-trial. J Sports Sci 2006; 24: 1165 – 1171

25 *Wilkinson DM.* Training for middle and long-distance running. In: Fallowfield JL, Wilkinson DM (eds). Improving Sports Performance in Middle and Long Distance Running. A Scientific Approach to Race Preparation. Chichester, UK: Wiley, 1999: 69 – 97