

Accuracy of Cycling Power Meters against a Mathematical Model of Treadmill Cycling

Authors

Thomas Maier, Lucas Schmid, Beat Müller, Thomas Steiner, Jon Peter Wehrlin

Affiliation

Section for Elite Sport, Swiss Federal Institute of Sport, Magglingen, Switzerland

Key words

dynamic calibration, validity, reliability, measurement error

accepted after revision 19.01.2017

Bibliography

DOI <http://dx.doi.org/10.1055/s-0043-102945>

Published online: 8.5.2017

Int J Sports Med 2017; 38: 456–461

© Georg Thieme Verlag KG Stuttgart · New York

ISSN 0172-4622

Correspondence

Thomas Maier, MSc

Section for Elite Sport

Swiss Federal Institute of Sport

Hauptstrasse 247

2532, Magglingen

Switzerland

Tel.: +41/58/467 63 35, Fax: +41/58/467 64 05

thomas.maier@baspo.admin.ch

ABSTRACT

The aim of this study was to compare the accuracy among a high number of current mobile cycling power meters used by elite and recreational cyclists against a first principle-based mathematical model of treadmill cycling. 54 power meters from 9 manufacturers used by 32 cyclists were calibrated. While the cyclist coasted downhill on a motorised treadmill, a back-pulling system was adjusted to counter the downhill force. The system was then loaded 3 times with 4 different masses while the cyclist pedalled to keep his position. The mean deviation (trueness) to the model and coefficient of variation (precision) were analysed. The mean deviations of the power meters were $-0.9 \pm 3.2\%$ (mean \pm SD) with 6 power meters deviating by more than $\pm 5\%$. The coefficients of variation of the power meters were $1.2 \pm 0.9\%$ (mean \pm SD), with Stages varying more than SRM ($p < 0.001$) and PowerTap ($p < 0.001$). In conclusion, current power meters used by elite and recreational cyclists vary considerably in their trueness; precision is generally high but differs between manufacturers. Calibrating and adjusting the trueness of every power meter against a first principle-based reference is advised for accurate measurements.

Introduction

Mobile cycling power meters are used extensively in various cycling disciplines to monitor training [15], to conduct field-based performance tests [22], to analyse competitions [26], or to evaluate equipment changes [17]. These applications demand accurate power output measurements, where accuracy is defined by ISO 5725 (International Organization for Standardization, Switzerland) as the combination of trueness and precision.

First principle-based calibration procedures have long been developed for cycle ergometers [25], and the dynamic calibration rig was established as a gold-standard device for factory calibrations, alongside static calibrations. It measures the required torque while propelling the crank axle of an ergometer [19, 27]. This method has also been applied to mobile power meters. Devices from SRM and PowerTap have been shown to deliver precise measurements, although the initial error of individual devices varied [9]. The accuracy of an SRM power meter has further been confirmed for constant power trials but has been shown to be decreased at high power outputs [1].

However, the dynamic calibration rig cannot be used to calibrate all power meter systems. It fails if the power meter measures the

power output of the cyclist before its transmission through the crank axle, as in pedal- or crank arm-based power meters. Additionally, the dynamic calibration rig forces the power meter to operate in an artificial setting with a constant torque on the crank axle, differing from the oscillating torque profile of a cyclist alternately pushing the pedals. It would be preferable to calibrate a power meter during actual use by a cyclist.

Therefore, numerous studies have compared power meters by installing them simultaneously on the same bike and comparing their measurements. The results varied tremendously with differences from 1.2 [5] to 16.5% [8], depending on the power meters used, the range of measured power output, and the calibration protocols [20]. Apart from SRM and PowerTap, power meters from Quarq [20] and Stages [13] were calibrated in this way. However, this experimental setup lacks a first principle-based reference against which to compare the power meters.

Power output during cycling on a motorised treadmill has been shown to be highly reliable [6]. Additionally, the required power output can be calculated with a mathematical model because most resistive forces can be directly quantified [7, 14, 18, 21]. The rolling resistance of the tyres is usually the only unknown, potentially hin-

dering accurate calculations [6, 11]. With a known or controlled rolling resistance, calculated power output during treadmill cycling seems promising as a first principle-based reference against which to calibrate power meters.

So far, only one study calibrated more than one device of a power meter system simultaneously [9]. Therefore, limited generalizable data regarding the accuracy of different power meters are available. Moreover, no studies reported calibrations of power meters from Quarq and Stages using a first principle-based reference. The aim of this study was to compare the accuracy among a high number of current power meters, used by elite and recreational cyclists, against a first principle-based mathematical model of treadmill cycling.

Method

Power meters and cyclists

A total of 54 power meters were calibrated (► **Table 1**). Thereof, 47 power meters were currently used by 32 cyclists (19 elite cyclists from the National Team and 13 recreational cyclists), who volunteered to participate in the study. The 7 remaining power meters were currently used in a sports science lab (4 SRM Science, 2 PowerTap G3 and 1 PowerTap SL). All power meters were installed on mountain bikes or road bikes of the participating cyclists.

Most of the cyclists were accustomed to treadmill cycling from previous performance tests or training, whereas 9 cyclists were only accustomed to the similar task of riding on cycling rollers.

All cyclists received written and oral information about the study aim and the procedures. Written informed consent was obtained. The study was accepted by the institutional review board and meets the ethical standards of the International Journal of Sports Medicine [10].

Study design

All calibrations were conducted between September 2015 and August 2016.

After consenting to participate, the cyclists completed the calibration protocol. Cyclists not accustomed to treadmill cycling completed habituation training before. 11 cyclists calibrated multiple power meters, but no recalibration of a power meter was included

in this study. For each cyclist, at most one Stages power meter was included in the study.

Calibration protocol

Preceding the calibration, the correct bike setup was checked, followed by a free 10–15 min warm-up by the cyclist. The slope of the treadmill was then set to -1° , and the back-pulling system was prepared (► **Fig. 1**). The cyclist was instructed to minimise lateral movements during the calibration and to ride with a freely chosen but constant pedalling cadence.

Subsequently, the following protocol was repeated 3 times:

1. The 0-offset of the power meter was reset when applicable.
2. The back-pulling system was adjusted with a small mass (m_1) to counter the downhill force while the cyclist coasted (without pedalling) downhill at a speed of $6 \text{ m} \cdot \text{s}^{-1}$. This resulted in a stable position of the cyclist with no forward or backward movement relative to the treadmill border.
3. Using the same speed of $6 \text{ m} \cdot \text{s}^{-1}$ the back-pulling system was additionally loaded with 4 different masses (m_2), and the cyclist had to pedal to keep his position (any displacement $< 0.2 \text{ m}$). The power output of the cyclist was measured and averaged for 1 min for each m_2 after the cyclist had approximately 15 s to adjust to the new load.

The protocol resulted in a total of 12 measurements (3 repetitions, 4 masses for m_2).

Bike setup

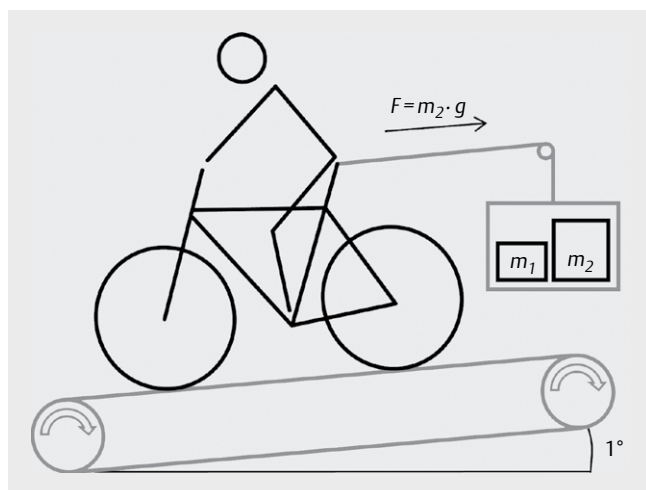
The drivetrain and wheel-bearings of each bike were checked for unusual friction by slowly rotating the crank and each wheel. Concerning the power meter, the correct installation (e. g., installation torque, cadence magnet placement), manufacturer-specific calibration steps (e. g., slope setting on the recording device, pedalling routine with some pedal-based systems), and signal transmission to the recording device were controlled. The recording device was set to a measuring frequency of 1 Hz, and automatic adjustment of the 0-offset was disabled when applicable.

Treadmill

A motorised treadmill ($3 \times 4 \text{ m}$, Poma, Germany) was used with a belt suitable for cycling. The exact speed of the treadmill was calibrated before and after the study by measuring the length of the

► **Table 1** Power meters calibrated in this study.

n	Manufacturer	Country	Models (n)	Position
12	SRM	Germany	Science (4), Dura Ace (5), FSA (1), XX1 (2)	Crank spider
10	PowerTap	USA	P1 (4), G3 (4), GS (1), SL (1)	Pedals (P1), wheel hub
11	Quarq	USA	XX1 (8), SRAM Red (2), Elsa (1)	Crank spider
13	Stages Cycling	USA	XTR (6), Rival (2), Dura Ace (2), Carbon (1), Ultegra (1), XT (1)	Crank arm (left only)
3	Verve Cycling	Australia	InfoCrank (3)	Crank arm (left and right)
2	power2max	Germany	FSA (1), Ultegra (1)	Crank spider
1	Garmin	USA	Vector (1)	Pedals
1	Polar	Finland	Kéo Power (1)	Pedals
1	Rotor	Spain	Power (1)	Crank arm (left and right)



► **Fig. 1** Treadmill setup for the calibration protocol with the back-pulling system. m_1 = mass to counter downhill force, m_2 = mass for calibration measurements, F = back-pulling force for calibration measurements, g = standard gravity.

treadmill belt and counting the revolutions per time interval with different speed settings.

Back-pulling system

A back-pulling system as described previously [6, 14] was used (► **Fig. 1**). A rope was attached to the saddle railing of the bike, was guided over a freely rotating pulley behind the treadmill and connected to a basket. The vertical and lateral position of the pulley was matched with the relative position of the saddle railing to the treadmill belt. To define the lateral position of the cyclist a visual marker was placed in front of the treadmill.

Small masses (1 – 250 g) were used to adjust m_1 . To adjust m_2 , mass plates (1 – 2 kg) were used after their exact mass was measured with a calibrated precision scale (ICS425k, Mettler-Toledo, Switzerland). For the calibration protocol, m_2 was set to 3, 4, 5, and 6 kg to result in required power outputs of approximately 180, 240, 300, and 360 W. For 6 cyclists, m_2 was changed to 2, 3, 4, and 5 kg to lower the intensity of the calibration protocol. The power outputs used in the calibration protocol represent a compromise between covering various intensities and suiting cyclists of different levels.

Mathematical model

The required power output of the cyclist for each m_2 was calculated with a mathematical model. In step 2 of the calibration protocol, the cyclist is coasting downhill without pedalling. By adjusting m_1 , he then achieves force equilibrium. In step 3, the only force the cyclist has to overcome is the gravitational force of the additional mass m_2 and the frictional resistance of the drivetrain.

Limited scientific literature exists regarding the efficiency of bicycle drivetrains [24]. After personal communication with an expert in this field (J. Smith, Friction Facts, www.friction-facts.com), a constant and a power output-dependent part were considered for the drivetrain loss. The constant part covers the frictional loss-

es in the derailleur pulleys (~3 W), whereas the power output-dependent part covers the frictional losses in the loaded upper part of the chain (~1.5 %).

Therefore, the final model for the calculated power output (P_{calc}) was $P_{\text{calc}} = m_2 \cdot g \cdot v / 0.985 + 3 \text{ W}$ (where g = standard gravity and v = speed of the treadmill). For the power meters located in the rear hub (PowerTap G3, GS, SL), the drivetrain loss was excluded from the calculations.

Data analysis

Power output measurements were analysed with cycling performance software (Golden Cheetaah 3.1, www.goldencheetaah.org) and visually inspected for interruptions in the signal transmission. The relative deviations of the 12 measured power outputs (P_{meas}) to the calculated power outputs were derived ($P_{\text{meas}}/P_{\text{calc}} - 1$) and were used for further calculations. Accuracy is defined by ISO 5725 as the combination of trueness (mean deviation to the reference value) and precision (variability of repeated measurements). Accordingly, trueness was quantified with the mean deviation, and precision was quantified with the coefficient of variation (CV) for each power meter.

Power meters were grouped by manufacturers. SRM, PowerTap, Quarq, and Stages were compared with non-parametric Kruskal-Wallis tests ($\alpha = 0.05$). In case of significant main effects, pairwise post-hoc Mann-Whitney U tests with Bonferroni corrections were applied. Intensity related effects were analysed with Friedman tests and pairwise post-hoc Wilcoxon signed-rank tests with Bonferroni corrections.

Data analysis was conducted with a statistical software package (R 3.2.2, R Core Team, Austria). Values are presented as mean \pm standard deviation.

Results

Trueness

The mean deviations of the power meters were $-0.9 \pm 3.2\%$ (► **Table 2**), which was not significantly different from 0 % ($p = 0.08$). 6 (11 %) power meters deviated by more than $\pm 5\%$ (1 power2max, 1 Quarq, 4 Stages, ► **Fig. 2**). There was a significant main effect of manufacturer ($p = 0.03$), but no pairwise comparison reached significance. Overall, the mean deviation was 0.9 % lower with the lightest m_2 compared to the 3 heavier loads ($p < 0.001$).

Precision

The coefficients of variation of the power meters were $1.2 \pm 0.9\%$ (► **Table 2**), with 5 (9 %) power meters having a CV greater than 2.5 % (1 Quarq, 1 Polar, 3 Stages, ► **Fig. 3**). There was a significant main effect of manufacturer ($p < 0.001$). Post-hoc tests revealed a higher CV of Stages power meters ($2.0 \pm 1.4\%$) compared to SRM ($0.8 \pm 0.4\%$, $p < 0.001$) and PowerTap ($0.8 \pm 0.2\%$, $p < 0.001$). Overall, the CV was 0.4 % higher with the lightest m_2 compared to the 3 heavier loads ($p < 0.001$) and the CV was 0.2 % lower with the heaviest m_2 compared to the second ($p = 0.007$) and the third load ($p = 0.004$).

► **Table 2** Trueness and precision of power meters by manufacturer.

n	Manufacturer	Mean deviation (%)	Coefficient of variation (%)	Cadence (RPM)
12	SRM	-0.5 ± 2.4	0.8 ± 0.4	83 ± 14
10	PowerTap	0.9 ± 2.1	0.8 ± 0.2	87 ± 5
11	Quarq	0.5 ± 3.0	1.3 ± 0.8	87 ± 6
13	Stages Cycling	-2.9 ± 3.9	2.0 ± 1.4 *	89 ± 6
3	Verve Cycling	-1.7 ± 1.1	0.6 ± 0.4	88 ± 3
2	power2max	-4.8 ± 3.4	1.5 ± 0.4	87 ± 16
1	Garmin	-2.0	1.6	86
1	Polar	-3.9	2.6	93
1	Rotor	2.1	0.4	84
54	All	-0.9 ± 3.2	1.2 ± 0.9	87 ± 8

Values are presented as mean ± standard deviation (if n > 1).
* Different from SRM and PowerTap p < 0.05

Discussion

In this study, a high number of current power meters used by elite and recreational cyclists was calibrated against a first principle-based mathematical model of treadmill cycling. The overall mean deviation of the power meters was not different from 0% and demonstrates the general agreement between the measured values and the mathematical model.

Trueness

While the overall mean deviation was not different from 0%, individual power meters deviated substantially, with a concerning number of 6 power meters deviating by more than 5%. Therefore, trueness seems to vary considerably between individual devices, even when they are from the same manufacturer. Presumably, this could be caused by inadequate factory calibrations or later shifts in the inherent torque-to-signal characteristics of the power meters during the use by the cyclists.

So far, only one study calibrated a high number of power meters. Gardner et al. [9] showed that 12 of 19 SRM power meters deviated by more than 2% but remained stable after the first calibration over a period of 11 months. Their finding of differing trueness in individual power meters from the same manufacturer is in line with the current study. The numerous studies analysing the agreement between simultaneously installed power meters are difficult to interpret [1, 3–5, 8, 13, 20, 23] because generalisations from individual power meters to their respective manufacturers are highly limited, as Gardner et al. [9] and the current study illustrate. However, direct system-to-system comparisons revealed differences comparable to the current study or higher (1–2% between PowerTap and SRM [5]; 0–12% between PowerTap, Quarq and Stages [20]; 2–17% between PowerTap, SRM and Ergomo®Pro [8]).

Precision

The low CV of all power meters of 1.2 ± 0.9% indicates high general precision of most power meters, but manufacturer-dependant differences exist. Power meters from SRM and PowerTap were more precise compared to devices from Stages.

This comparison, however, is not completely legitimate because the power meters from Stages calibrated in this study measured only torque in the left crank arm, with the assumption of the right side being equal. Thus, the derived trueness and precision in the current study always depended on the power meter itself and the riding style of the cyclist (left-right balance). A varying left-right balance during the calibration would increase the variability of the measured power outputs and, therefore, lower the precision. Kirkland et al. [16] reported a contribution of 48.9 ± 3.6% from the left leg in a group of 9 competitive cyclists for similar power outputs. The variability of 3.6% illustrates how the accuracy of the Stages power meter could be strongly influenced by the cyclist himself, apart from technical measurement error.

The precision values of the current study are in line with previous results from numerous SRM and PowerTap power meters (CV < 2%) [9]. Some studies that tested single devices reported slightly higher CV values (1–3% for PowerTap and SRM [5]; 2–3% for PowerTap, Quarq and Stages [20]; 2% for PowerTap and SRM and 4% for Ergomo®Pro [8]).

When interpreting the precision of the calibrated power meters, it is important to consider that any variation induced by the experimental setup would have decreased the precision. Nonetheless, as the overall CV was only 1.2%, the potential variation had to be very small, underpinning the reliability of the experimental setup used in the current study.

Accuracy

Because the precision of the power meters in the current study was generally high, the individual devices with good trueness can be judged as accurate. For the other devices, the torque-to-signal characteristic would have to be adjusted for accurate measurements of power output, which is often not possible.

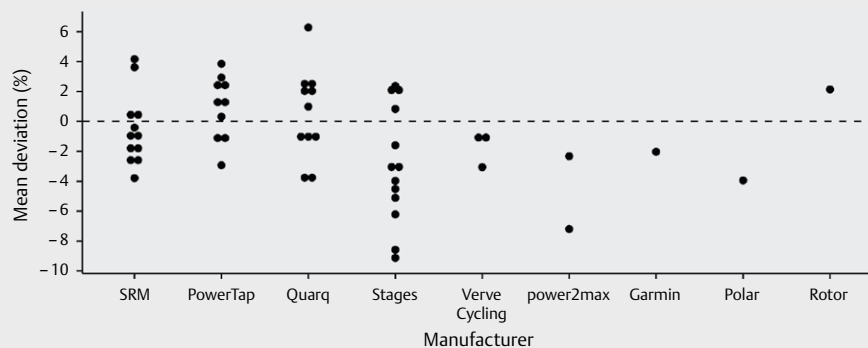
Relevance for training and testing

Measurement error, apart from biological variation, deteriorates the test-retest reliability of performance tests. Comparing power output between systems that do not measure true introduces a systematic bias and tracking values over time with an imprecise system may hide a true signal (e. g., improvement in performance) in the noise [2, 12].

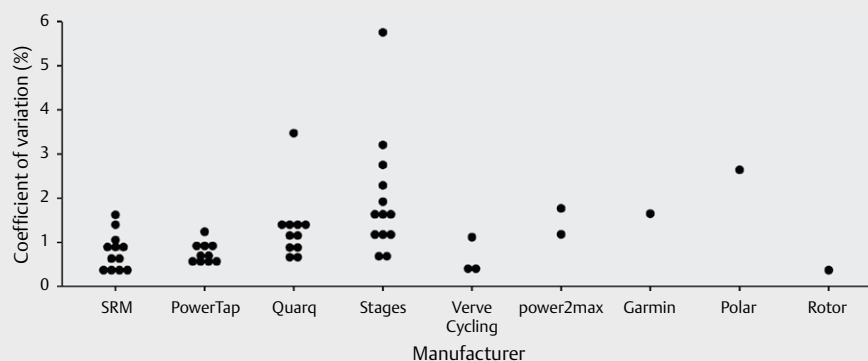
In the current study, trueness varied by 3.2% between individual power meters, whereas the smallest worthwhile change in performance could be lower than 1%, depending on the discipline [2]. Therefore, the differing trueness of the power meters could lead to substantial over- or underestimations of the capabilities of the respective cyclists. The precision values in the current study illustrate the difficulty of identifying small but worthwhile changes. The mean precision of SRM or PowerTap power meters of 0.8% allows identifying a change of 1.1% (0.8% · √2) with 68% confidence in a test-retest scenario [12]. Using a Stages power meter with a precision of 2.0%, the identifiable change increases to 2.8%. To identify smaller changes, multiple (n) tests or measurements could be averaged, which decreases the identifiable change by the factor 1/√n [12].

Limitations

In this study, the power meters were calibrated under laboratory conditions. During field use, the accuracy could further be compro-



► Fig. 2 Trueness of individual power meters.



► Fig. 3 Precision of individual power meters.

mised because of changing ambient temperature or vibrations and impacts from the ground surface or gear shifts, among others. Additionally, only short measurement periods were analysed in contrast to longer recordings that are typically used for training and testing. However, accurate measurements in controlled conditions are presumably necessary for high accuracy under field conditions [20]. Further research in this area could investigate the accuracy of a high number of power meters under field conditions.

The assumed drivetrain loss in the mathematical model used could not be based on direct measurements. Even though this directly influences the values concerning trueness, the influence on the group comparisons and the precision values is limited. The high precision values across the studied range of power outputs further confirm the assumption of drivetrain losses in the current study, but the intensity-related effect on trueness may hint at a difference between the assumed and actual drivetrain loss.

Power meters were calibrated only up to power outputs of approximately 360 W, often below the values of elite cyclists during high-intensity training or testing. It is unclear if the results of the current study are also valid for higher power outputs. Depending on the ability of the cyclists, higher power outputs could be used in the calibration protocol, but the method is not suitable for power outputs occurring in sprints.

Because the pedalling cadence was not controlled during the calibrations, it is not possible to estimate its effects on the accuracy of the power meters.

Conclusion

It can be concluded that the trueness seems to vary considerably between current power meters used by elite and recreational cyclists, even when the devices are from the same manufacturer. However, precision is generally high, apart from devices from Stages that show a lower precision than devices from SRM and PowerTap.

The current study illustrates the value of calibrating and, if possible, adjusting the trueness of every power meter for accurate measurements of power output for training and testing. Calibrating power meters against a first principle-based mathematical model of treadmill cycling is specific and feasible with every current system.

Acknowledgements

This study was supported by the Swiss Federal Office of Sports, Magglingen and Swiss Cycling.

Conflict of interests

The authors declare no conflicts of interest.

References

- [1] Abbiss 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] Atkinson G, Nevill AM. Selected issues in the design and analysis of sport performance research. *J Sports Sci* 2001; 19: 811–827
- [3] Bertucci W, Crequy S, Chiementin X. Validity and reliability of the G-cog BMX powermeter. *Int J Sports Med* 2013; 34: 538–543
- [4] Bertucci W, Duc S, Villerius V, Grappe F. Validity and reliability of the axiom PowerTrain cycle ergometer when compared with an SRM powermeter. *Int J Sports Med* 2005; 26: 59–65
- [5] 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
- [6] Coleman DA, Wiles JD, Davison RC, Smith MF, Swaine IL. Power output measurement during treadmill cycling. *Int J Sports Med* 2007; 28: 525–530
- [7] Di Prampero PE, Cortili G, Mognoni P, Saibene F. Equation of motion of a cyclist. *J Appl Physiol* 1979; 47: 201–206
- [8] Duc S, Villerius V, Bertucci W, Grappe F. Validity and reproducibility of the ErgomoPro power meter compared with the SRM and PowerTap power meters. *Int J Sports Physiol Perform* 2007; 2: 270–281
- [9] Gardner AS, Stephens S, Martin DT, Lawton E, Lee H, Jenkins D. Accuracy of SRM and PowerTap power monitoring systems for bicycling. *Med Sci Sports Exerc* 2004; 36: 1252–1258
- [10] Harriss DJ, Atkinson G. Ethical standards in sport and exercise science research: 2016 update. *Int J Sports Med* 2015; 36: 1121–1124
- [11] Henchoz Y, Crivelli G, Borrani F, Millet GP. A new method to measure rolling resistance in treadmill cycling. *J Sports Sci* 2010; 28: 1043–1046
- [12] Hopkins WG. Measures of reliability in sports medicine and science. *Sports Med* 2000; 30: 1–15
- [13] Hurst HT, Atkins S, Sinclair J, Metcalfe J. Agreement between the Stages cycling and SRM powermeter systems during field-based off-road climbing. *J Sci Cycling* 2015; 4: 21–27
- [14] Jakeman P. A longitudinal study of exercise metabolism during recovery from viral illness. *Br J Sports Med* 1993; 27: 157–161
- [15] Jobson SA, Passfield L, Atkinson G, Barton G, Scarf P. The analysis and utilization of cycling training data. *Sports Med* 2009; 39: 833–844
- [16] Kirkland A, Coleman D, Wiles J, Hopker J. Validity and reliability of the Ergomo® pro powermeter. *Int J Sports Med* 2008; 29: 913–916
- [17] Lim AC, Homestead EP, Edwards AG, Carver TC, Kram R, Byrnes WC. Measuring changes in aerodynamic/rolling resistances by cycle-mounted power meters. *Med Sci Sports Exerc* 2011; 43: 853–860
- [18] 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
- [19] Maxwell BF, Withers RT, Ilesley 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
- [20] Miller MC, Macdermid PW, Fink PW, Stannard SR. Agreement between PowerTap, Quarq and Stages power meters for cross-country mountain biking. *Sports Technol* 2016. Epub ahead of print: 1–7
- [21] Olds TS, Norton KI, Lowe EL, Olive S, Reay F, Ly S. Modeling road-cycling performance. *J Appl Physiol* 1995; 78: 1596–1611
- [22] Quod MJ, Martin DT, Martin JC, Laursen PB. The power profile predicts road cycling MMP. *Int J Sports Med* 2010; 31: 397–401
- [23] Sparks SA, Dove B, Bridge CA, Midgely AW, McNaughton LR. Validity and reliability of the look Keo power pedal system for measuring power output during incremental and repeated sprint cycling. *Int J Sports Physiol Perform* 2015; 10: 39–45
- [24] Spicer JB, Richardson CJK, Ehrlich MJ, Bernstein JR, Fukuda M, Terada M. Effects of frictional loss on bicycle chain drive efficiency. *J Mech Des* 2001; 123: 598–605
- [25] Stein ES, Rothstein MS, Clements CJ Jr. Calibration of two bicycle ergometers used by the health examination survey. *Vital Health Stat* 2: 1967; 1–10
- [26] 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
- [27] Woods GF, Day L, Withers RT, Ilesley AH, Maxwell BF. The dynamic calibration of cycle ergometers. *Int J Sports Med* 1994; 15: 168–171