High Agreement between Laboratory and Field Estimates of Critical Power in Cycling

B. Karsten1, S. A. Jobson2, J. Hopker3, A. Jimenez4, C. Beedie5

1 Life and Sports Science, University of Greenwich, Chatham Maritime, United Kingdom
2 Department of Sports Studies, University of Winchester, United Kingdom
3 Centre for Sport Studies, University of Kent, Chatham, United Kingdom
4 College of Sports and Exercise Science, Victoria University, Melbourne, Australia
5 Department of Sport & Exercise Sciences, Aberystwyth University, Aberystwyth, United Kingdom

Abstract

The purpose of this study was to investigate the level of agreement between laboratory-based estimates of critical power (CP) and results taken from a novel field test. Subjects were fourteen trained cyclists (age 40±7 yrs; body mass 70.2±6.5 kg; \( \text{V}_\text{O}_2\max \) 3.8±0.5L·min\(^{-1}\)). Laboratory-based CP was estimated from 3 constant work-rate tests at 80%, 100% and 105% of maximal aerobic power (MAP). Field-based CP was estimated from 3 all-out tests performed on an outdoor velodrome over fixed durations of 3, 7 and 12 min. Using the linear work limit (\( W_{\text{lim}} \)) vs. time limit (\( T_{\text{lim}} \)) relation for the estimation of CP1 values and the inverse time (1/t) vs. power (P) models for the estimation of CP2 values, field-based CP1 and CP2 values did not significantly differ from laboratory-based values (234±24.4 W vs. 234±25.5 W (CP1); P<0.001; limits of agreement [LOA], −10.98–10.8 W and 236±29.1 W vs. 235±24.1 W (CP2); P<0.001; [LOA], −13.88–17.3 W). Mean prediction errors for laboratory and field estimates were 2.2% (CP) and 27% (W). Data suggest that employing all-out field tests lasting 3, 7 and 12 min has potential utility in the estimation of CP.

Introduction

Performance tests are a key component in the training of athletes, providing markers of performance that can be used as an indicator of training status. One such marker is critical power (CP). In the hyperbolic relationship between power output and time to fatigue, CP represents the highest prolonged sustainable work rate, while the curvature constant (W) relates to a finite amount of energy which can be performed above CP [13, 14, 32]. Poole et al. [34] characterized the physiological response to exercise performed at CP, showing that CP represents the highest power output at which \( \text{V}_\text{O}_2 \) and blood lactate stabilize. Theoretically CP is sensitive to changes in performance capacity which are likely to occur frequently during athletic training, and therefore provides a useful indicator of training status.

In cycling CP is traditionally estimated in laboratory conditions by using time to exhaustion (TTE) trials at fixed intensities [8,25,35]. An estimation of maximal aerobic power (MAP) is required to calculate the intensity in question. The total number of trials required to estimate CP ranges between 3 and 5 [7, 15, 23, 24, 35, 42], although it is usual for at least 3 trials to be performed, especially in non-elite athletes. Given this, laboratory estimation of CP can be time-consuming and potentially disruptive to an athlete’s training programme.

While the estimation of CP in cycling has traditionally been lab-based, other sports have used field-based estimates of the related phenomenon of Critical Velocity (CV). In swimming, Wakayoshi [47] and Dekkerle [9] suggested that the field estimation of CV in swimmers requires only 2 performances (200m and 400m). In running Kranenburg and Smith [28] estimated CV using all-out cycling tests on a treadmill that induced exhaustion within 3, 7 and 12 mins, and employed 3 set distances, each run on an indoor track, to estimate field CV. The authors reported that this field-based method of estimating CV proved to be robust, and that field CV was significantly related to 10 km race speed. Again in running, Galbraith et al. [16] developed a field test to determine critical speed also using set distances yielding finishing times between 2 and 12 min. Both studies used three trials of durations between 3 and 12 min, and used trained subjects. Hiyanne et al. [20] estimated CV using all-out cycling tests over distances of 2, 4 and 6 km resulting in testing times between 1 and 10 min.
Data suggest that laboratory and field tests might produce different findings. For example, Jobson et al. [27] reported higher power output values in the field than in the laboratory at given VO₂ values, while Bertucci et al. [3] found an increased gross efficiency and cycling economy in the field when compared to the laboratory. While conditions in the laboratory are more controllable, providing greater reliability, field tests have the advantage of providing greater ecological validity [21,30]. Such validity might be a function of many factors. For example, field tests allow the athlete to perform in an environment consistent with that in which they usually compete, permitting previously acquired effort regulation skills to be employed, therefore reducing the need for habituation to laboratory protocols. Field tests are also relatively unconstrained by the mechanical limitations often imposed by laboratory equipment. Contrast for example cycling on a velodrome with riding a mechanically stable ergometer. In the former the bike moves laterally under the rider, and the rider is likely to have developed a handling technique that both controls for this and in doing so optimises the contribution to forward motion of various synergistic and stabilising components of the skeletal- and neuro-muscular systems. These components are less likely to be employed in all but the most ecologically valid laboratory settings. These factors are especially pertinent if the performance in question is measured over conditions in which they usually compete, permitting previously acquired e

Methods

Subjects

12 male and 2 female cyclists were recruited from local cycling clubs (mean ± SD: age 40 ± 6 yrs; body mass 70.2 ± 6.5 kg; VO₂max 3.8 ± 0.5 L · min⁻¹; MAP 311 ± 32.5 W). All subjects had been involved in regular cycling training and competition for a minimum 2 years. The study was performed in accordance with the ethical standards of the International Journal of Sports Medicine [18] and approved by the University Ethics Committee. Prior to providing written informed consent and participation, cyclists were fully informed of the nature and risks of the study.

Protocol

The study used an intra-subject design. During the first laboratory session maximal oxygen consumption (VO₂max) and MAP values were established. Subjects then performed 3 laboratory-based ergometer CP tests and 3 field CP tests all randomised (below). Subjects were not informed of their performance outcomes until the completion of the study. To minimise training effects each subject completed all 7 sessions within 21 days. A minimum of 24 h rest was required between individual tests [8,36].

Maximal oxygen uptake test protocol

Following a standardised warm-up (5 min cycling at 100 W, followed by 3 min unloaded phase), subjects completed a progressive, incremental laboratory exercise test to exhaustion. Expired gases were collected continuously throughout the test using a MetaMax gas analyser (MetaMax 3B, Cortex Biophysik, Leipzig, Germany). Blood lactate was analysed using the Biosen C line analyser (EFK Diagnostics, Barleben, Germany), and heart rate was continuously monitored using the PowerTap. The maximal test commenced at a work rate of 150 W for male and 120 W for female cyclists (females have lower absolute peak power values than males [44] and a maximal test should last somewhere between 8 and 10 min for moderately trained to trained subjects [48]). Thereafter, intensity increased at a step rate of 20 W · min⁻¹ using power values obtained from the PowerTap. Consistent
with previous CP research [46] cyclists’ preferred cadence was used throughout the test. The test was terminated when cadence dropped by more than 10 rev·min⁻¹ for more than 10 s. MAP was calculated as the highest mean power values (W) over the final 60 s. VO₂max was calculated as the highest mean oxygen consumption over a 30-s period.

Laboratory-based critical power tests

Cyclists completed 3 tests to exhaustion on the equipment described above. All tests were performed on separate days at work rates equivalent to 80%, 100% and 105% of MAP. After a 5-min warm-up at a work rate of 100 W, the test resistance was set and cyclists were instructed to maintain their preferred cadence for as long as possible. Heart rate (HR), P and cadence were logged continuously by the PowerTap. Consistent with previous CP research [46] strong verbal encouragement was provided throughout the tests. Capillary fingertip blood samples were collected at rest, immediately post-test and 3 min post-test and analysed for lactate concentration. Consistent with published guidelines [4] a post-test blood lactate concentration of ≥8 mmol·l⁻¹ or HR within 10 beats of age-predicted HR maximum was taken as an indicator for attainment of VO₂max and accepted as a successful test.

Field-based critical power tests

Subjects were tested over fixed times of 3, 7 and 12 min rather than over set distances [34] on an outdoor velodrome. Tests were completed on separate days and in randomised order. Tests started from a standing position, and subjects were instructed to ride around the velodrome as fast as possible in each test. Feedback regarding remaining time, as well as verbal encouragement, was provided throughout the tests. Capillary fingertip blood samples were taken at rest, immediately post-test and 3 min post-test. A post-test blood [lactate] of ≥8 mmol·l⁻¹ or HR within 10 beats of age-predicted HR maximum was taken as an indicator for attainment of VO₂max and accepted as a successful all-out test [5].

Control of environmental factors

As suggested above, environmental conditions are a major concern in field testing. Consistent with the data reported by Swain [43], it was initially decided that field testing would not take place in wind speeds above 6.6 m·s⁻¹, or in rain or otherwise wet conditions. The latter scenario was relatively straightforward to address. However, wind speed so frequently exceeded the 6.6 m·s⁻¹ level that cancelling tests on the basis of this criterion would have extended data collection beyond the 21-day criterion and might have introduced other sources of error (e.g., training/de-training effects). Cancelling on the basis of wind speed – which would have led to several tests being abandoned once underway – would likely have led to subjects dropping out of the study. Therefore testing went ahead irrespective of measured wind speed, and this issue and decision are discussed further below.

Calculation of critical power and W′

Linear regression was used to provide an estimate of CP and W′ from the results of the laboratory and the field trials using the work-time model [P=W′+(CP·t)] are consequently termed CP1 and W1 and using the power-1/time model [P=(W′t)+CP] are consequently termed CP2 and W2.

Statistical methods

The distribution of each variable was examined with the Shapiro-Wilk’s normality test. Pearson product moment correlation analysis was used to provide an indication of the strength of any relationship between field- and laboratory-derived CP1 and CP2 and W1 and W2. Agreement between laboratory and field CP1 and CP2 and W1 and W2 was assessed using a paired samples t-test and Limits of Agreement (LOA; [1,6]). Paired samples t-tests were conducted to identify any differences in laboratory and field based CP TTE trials, in maximal lactate concentration, and maximal HR for each equivalent test (80% and 12 min, 100% and 7 min, 105% and 3 min) and for differences between relative percentages of MAP achieved during the laboratory- and field-based CP1 and CP2 tests. Additionally, linear regression was used to estimate error associated with predicting field CP and W′ values [22]. Statistical significance was accepted at P < 0.05. Results are reported as mean ± SD unless otherwise stated.

Results

No significant differences were observed between field-based and laboratory-based CP1 (234±24.4 W vs. 234±25.5 W respectively; t (13) = 0.97, p = 0.345) and CP2 (235±24.1 W vs. 236±29.1 W respectively; t (13) = 0.81, P = 0.435). Data recorded in the 2 environments were highly correlated (r = 0.976; P < 0.05 (CP1) and r = 0.973; P < 0.05 (CP2)). Mean difference between laboratory- and field-based values for CP1 was 0.17±7.2 W (95% CI, −3.14−16.61; limits of agreement [LOA], −10.98−10.8 W) and for CP2 it was 2±7.72 W (95% CI, −22.28−25.35; [LOA], −13.88−17.3 W) (Fig. 1). Significant differences were observed between laboratory- and field-based W1 (12.2±2.7 kJ vs. 17.3±5.4 kJ, respectively, t (13) = −3.98, P = 0.02) and W2 (11.6±2.7 kJ vs. 16.5±4.8 kJ, respectively; t (13) = −3.93, P = 0.02). The mean difference in W1 was −5.1±4.8 kJ (95% CI, −7.86−9.14; [LOA], −14.5–4.3 kJ) and in W2 it was −4.9±4.7 kJ (95% CI, −7.58–8.94; [LOA], −14.0−4.2 kJ) (Fig. 2). The predication error associated with the laboratory-based and field-based estimates of CP/W′ was 1.9% (CP1), 2.5% (CP2) and for W it was 26.3% (W1) and 27.6% (W2). Analysis of [blood] lactate (mmol·l⁻¹) revealed significantly higher concentrations for field-based testing when comparing the 100% TTE trial vs. the 7 min test (t (13) = −2.12, p = 0.035) and the 105% TTE trial vs. the 3 min test (t (13) = −2.36, p = 0.009), while the 80% TTE trial vs. the 12 min test did not result in a statistically significant but low p-value (0.054) (Table 1). Table 2 illustrates the differences in mean initial 10 s and 30 s power values for field and laboratory-based tests. Table 3 shows mean durations (±SD) and mean distances (±SD) for laboratory and field tests respectively. Ferguson et al. [11,12] in their CP research added another TTE trial if individual SE values for a CP estimate fell above or below that of 3 W. Interestingly, individual SE values of ±3 W in the present study fit well for the linear work-time model of laboratory and field-based CP estimates but lie above (–8 W) of the recommended value in the power-1/time-power model.

Discussion

A mean difference between laboratory- and field-derived estimates of CP of 0.2 ± 5.7 W, suggests that field testing might provide...
CP1/2 of 1.9% and 2.5% respectively, the discrepancy between an average SEE value for laboratory-based CP1/2 vs. field-based CP1/2 and field-based CP using the Bland Altman test for the relation and bias (solid line) ±95% limits of agreement (dashed lines) between laboratory-based CP1/2 and field-based CP1/2 a and b.

Using a magnitude-based analysis, Paton and Hopkins [33], Kranenburg and Smith [28], and Galbraith et al. [16].

Table 1

<table>
<thead>
<tr>
<th>Test</th>
<th>Mean of W1 laboratory and W1 field (kJ)</th>
<th>Mean of W2 laboratory and W2 field (kJ)</th>
<th>Difference in power output (W1 laboratory - W1 field)</th>
<th>Difference in power output (W2 laboratory - W2 field)</th>
<th>Critical Power 1 laboratory (W)</th>
<th>Critical Power 2 laboratory (W)</th>
<th>Critical Power 1 field (W)</th>
<th>Critical Power 2 field (W)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100% MAP 3 min</td>
<td>12.26 (±2.29)</td>
<td>14.22 (±2.98)</td>
<td>0.009*</td>
<td>−3.34</td>
<td>13.55 (±3.30)</td>
<td>15.19 (±2.75)</td>
<td>0.021</td>
<td>−0.14</td>
</tr>
<tr>
<td>80% MAP 7 min</td>
<td>13.55 (±1.99)</td>
<td>15.19 (±2.75)</td>
<td>0.035*</td>
<td>−3.14</td>
<td>13.84 (±3.30)</td>
<td>14.95 (±3.09)</td>
<td>0.054</td>
<td>−0.21</td>
</tr>
</tbody>
</table>

Values are mean (± SD)

* significantly different from the mean 105% constant work load lactate values (P<0.05)

<table>
<thead>
<tr>
<th>Test</th>
<th>Mean of W1 laboratory and W1 field (kJ)</th>
<th>Mean of W2 laboratory and W2 field (kJ)</th>
<th>Difference in power output (W1 laboratory - W1 field)</th>
<th>Difference in power output (W2 laboratory - W2 field)</th>
<th>Critical Power 1 laboratory (W)</th>
<th>Critical Power 2 laboratory (W)</th>
<th>Critical Power 1 field (W)</th>
<th>Critical Power 2 field (W)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100% MAP 3 min</td>
<td>12.26 (±2.29)</td>
<td>14.22 (±2.98)</td>
<td>0.009*</td>
<td>−3.34</td>
<td>13.55 (±3.30)</td>
<td>15.19 (±2.75)</td>
<td>0.021</td>
<td>−0.14</td>
</tr>
<tr>
<td>80% MAP 7 min</td>
<td>13.55 (±1.99)</td>
<td>15.19 (±2.75)</td>
<td>0.035*</td>
<td>−3.14</td>
<td>13.84 (±3.30)</td>
<td>14.95 (±3.09)</td>
<td>0.054</td>
<td>−0.21</td>
</tr>
</tbody>
</table>

Values are mean (± SD)

* significantly different from the mean 150% constant work load lactate values (P<0.05)

Table 1 Group maximal [blood] lactate (mmol·l⁻¹) results, p-values and confidence intervals of the difference for each test.

- While in designing the study, the research team were optimistic that the field-based estimation of CP held some promise, differences between laboratory-based and field-based CP1/2 and field-based CP using the Bland Altman test for the relation and bias (solid line) ±95% limits of agreement (dashed lines) between laboratory-based CP1/2 and field-based CP1/2 a and b.

- The 2 measurement methods in the present study are deemed to be acceptable, considering that a group of elite cyclists would have likely produced lower biological variability.

- While in designing the study, the research team were optimistic that the field-based estimation of CP held some promise, differences between laboratory-based and field-based CP1/2 and field-based CP using the Bland Altman test for the relation and bias (solid line) ±95% limits of agreement (dashed lines) between laboratory-based CP1/2 and field-based CP1/2 a and b.

- With an average SEE value for laboratory-based CP1/2 vs. field-based CP 1/2 of 1.9% and 2.5% respectively, the discrepancy between the 2 measurement methods in the present study is deemed to be acceptable, considering that a group of elite cyclists would have likely produced lower biological variability.

- While in designing the study, the research team were optimistic that the field-based estimation of CP held some promise, differences between laboratory-based and field-based CP1/2 and field-based CP using the Bland Altman test for the relation and bias (solid line) ±95% limits of agreement (dashed lines) between laboratory-based CP1/2 and field-based CP1/2 a and b.

- A valid estimate of CP in cycling. Results support those of Quod et al. [37], Kraneburg and Smith [28], and Galbraith et al. [16].

- Using a magnitude-based analysis, Paton and Hopkins [33] identified that a change of 1.7% in performance impacts on the chances of an elite road time trial cyclist winning an event. With an average SEE value for laboratory-based CP1/2 vs. field-based CP 1/2 of 1.9% and 2.5% respectively, the discrepancy between the 2 measurement methods in the present study is deemed to be acceptable, considering that a group of elite cyclists would have likely produced lower biological variability.

- While in designing the study, the research team were optimistic that the field-based estimation of CP held some promise, differences between laboratory-based and field-based CP1/2 and field-based CP using the Bland Altman test for the relation and bias (solid line) ±95% limits of agreement (dashed lines) between laboratory-based CP1/2 and field-based CP1/2 a and b.

- A valid estimate of CP in cycling. Results support those of Quod et al. [37], Kraneburg and Smith [28], and Galbraith et al. [16].

- Using a magnitude-based analysis, Paton and Hopkins [33] identified that a change of 1.7% in performance impacts on the chances of an elite road time trial cyclist winning an event. With an average SEE value for laboratory-based CP1/2 vs. field-based CP 1/2 of 1.9% and 2.5% respectively, the discrepancy between the 2 measurement methods in the present study is deemed to be acceptable, considering that a group of elite cyclists would have likely produced lower biological variability.
used for field testing provided no shelter and wind speeds above the 6 m·s⁻¹ criterion suggested by Swain [43] were frequently observed. Given the linear function between work completed and time, any deviation of this linearity due to unequal headwind and tailwind speeds would have been identified in the individual CP1/CP2 field-based plots (the mean r-value for field-based CP1 was 0.99±0.001 and for field-based CP2 it was 0.99±0.008). Therefore our data do not appear to support those of Swain, and individual SE values reported above appear to support this position. Of course, given the relatively small number of subjects there is the possibility that the findings are due to chance. Therefore results will need to be tested on different, and ideally larger, samples.

Another aspect of the data worthy of discussion concerns the significant differences between laboratory and field-based estimates of W¹ and W². Field-based estimates for W¹ were on average 5.09kJ and for W² 4.89kJ higher than the respective laboratory values. This is accompanied by overall higher blood [lactate] responses for field testing (♦ Table 1) and by a difference in power profiles between laboratory and field. ♦ Table 2 illustrates the initial 10 and 30s of the all tests. Testing in the field began from a standing start with an initial acceleration phase, while constant load testing was performed at a constant cadence with the resistance increasing to the required intensity at the beginning of each TTE trial. This difference in power profile is most pronounced in the shorter field trials (3 and 7 min). It can be speculated that during the acceleration phase in the field subjects utilized a higher portion of type II muscle mass resulting in significantly higher power and blood [lactate] values [40, 45] compared to the constant load tests. The relative rate of field-based W' (kJ) expenditure therefore also seems to be greater when compared to the laboratory testing. Skiba et al. [41] suggested that W' may be primarily a representative of exercising type I and type II muscle mass but that the sum of W' expended at exhaustion is equal to the known total W. If this is true, then the difference between laboratory- and field-based W' estimates might be explained by the difference in environmental or testing conditions (i.e., standing start, acceleration against air resistance or use of body weight during the acceleration phase). We acknowledge this limitation to our field-based approach and recognise that a rolling start with paced lap times might provide a more reliable estimate of W'.

CP has traditionally been estimated in the laboratory. Results of the present study, while suggesting a significant difference in W' between the laboratory and the field, also suggest a high agreement in CP between the same environments. The field estimation of CP may offer a more ecologically valid and less expensive alternative to traditional approaches, making it a more widely available test. However the data presented above are from a small sample, and the authors advise a replication of the study, ideally with a larger subject group.

Acknowledgements

The authors would like to thank all subjects who participated in this study. No financial support was received.

References
3 Bertucci WM, Betik AC, Duc S, Grappe F. Gross efficiency and cycling economy are higher in the field as compared with an Axiom stationary ergometer. J Applied Biomech 2012; 28: 636–644

<table>
<thead>
<tr>
<th>Field Test</th>
<th>Initial 10s P (W)</th>
<th>Laboratory Test</th>
<th>Initial 10s P (W)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test 1</td>
<td>12 min = 532 ± 184 W</td>
<td>Test 1</td>
<td>80% TTE = 179 ± 38 W</td>
</tr>
<tr>
<td>Test 2</td>
<td>7 min = 624 ± 133 W</td>
<td>Test 2</td>
<td>100% TTE = 174 ± 38 W</td>
</tr>
<tr>
<td>Test 3</td>
<td>3 min = 633 ± 148 W</td>
<td>Test 3</td>
<td>105% TTE = 204 ± 34 W</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Field Test</th>
<th>Initial 30s P (W)</th>
<th>Laboratory Test</th>
<th>Initial 30s P (W)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test 1</td>
<td>12 min = 451 ± 132 W</td>
<td>Test 1</td>
<td>80% TTE = 212 ± 45 W</td>
</tr>
<tr>
<td>Test 2</td>
<td>7 min = 496 ± 108 W</td>
<td>Test 2</td>
<td>100% TTE = 230 ± 40 W</td>
</tr>
<tr>
<td>Test 3</td>
<td>3 min = 524 ± 95 W</td>
<td>Test 3</td>
<td>105% TTE = 279 ± 45 W</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mean time elapsed seconds (minutes:seconds)</th>
<th>SD time elapsed seconds (minutes:seconds)</th>
<th>Mean distance covered (metres)</th>
<th>SD distance covered (metres)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lab trials</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80% MAP</td>
<td>725.1 (12:05:1)</td>
<td>141.1 (02:21:1)</td>
<td>–</td>
</tr>
<tr>
<td>100% MAP</td>
<td>239.4 (03:59:4)</td>
<td>48.0 (00:48:0)</td>
<td>–</td>
</tr>
<tr>
<td>105% MAP</td>
<td>152.2 (02:32:2)</td>
<td>29.8 (00:29:8)</td>
<td>–</td>
</tr>
<tr>
<td>Field trials</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 min (180s)</td>
<td>–</td>
<td>1857.7</td>
<td>156.8</td>
</tr>
<tr>
<td>7 min (420s)</td>
<td>–</td>
<td>4118.4</td>
<td>232.6</td>
</tr>
<tr>
<td>12 min (720s)</td>
<td>–</td>
<td>7030.4</td>
<td>260.9</td>
</tr>
</tbody>
</table>

Table 2 Mean initial 10s and 30s P-values (W) for field and laboratory-based tests.

Table 3 Mean durations (+ SD) of laboratory-based TTE trials and field based set-time trials.