Journal of Science and Medicine in Sport 21 (2018) 742–747

Contents lists available at ScienceDirect

Journal of Science and Medicine in Sport

journal homepage: www.elsevier.com/locate/jsams

Original research

# Critical power: How different protocols and models affect its determination

Felipe Mattioni Maturana[a], Federico Y. Fontana[b], Silvia Pogliaghi[b], Louis Passfield[a,c], Juan M. Murias[a,*]

[a] Faculty of Kinesiology, University of Calgary, Canada
[b] Department of Neurosciences, Biomedicine and Movement Sciences, University of Verona, Italy
[c] School of Sport & Exercise Sciences, University of Kent, United Kingdom

ARTICLE INFO

*Article history:*
Received 20 July 2017
Received in revised form 11 November 2017
Accepted 22 November 2017
Available online 29 November 2017

*Keywords:*
Power–time relationship
Time-to-exhaustion
Linear model
Nonlinear model
Exercise intensity domains

ABSTRACT

In cycling, critical power (CP) and work above CP (W') can be estimated through linear and nonlinear models. Despite the concept of CP representing the upper boundary of sustainable exercise, overestimations may be made as the models possess inherent limitations and the protocol design is not always appropriate.

*Objectives:* To measure and compare CP and W' through the exponential ($CP_{exp}$), 3-parameter hyperbolic ($CP_{3\text{-}hyp}$), 2-parameter hyperbolic ($CP_{2\text{-}hyp}$), linear ($CP_{linear}$), and linear 1/time ($CP_{1/time}$) models, using different combinations of TTE trials of different durations (approximately 1–20 min).

*Design:* Repeated measures.

*Methods:* Thirteen healthy young cyclists (26 ± 3 years; 69.0 ± 9.2 kg; 174 ± 10 cm; 60.4 ± 5.9 mL $kg^{-1}$ $min^{-1}$) performed five TTE trials on separate days. CP and W' were modeled using two, three, four, and/or five trials. All models were compared against a criterion method ($CP_{3\text{-}hyp}$ with five trials; confirmed using the leaving-one-out cross-validation analysis) using smallest worthwhile change (SWC) and concordance correlation coefficient (CCC) analyses.

*Results:* CP was considerably overestimated when only trials lasting less than 10 min were included, independent of the mathematical model used. Following CCC analysis, a number of alternative methods were able to predict our criterion method with almost a perfect agreement. However, the application of other common approaches resulted in an overestimation of CP and underestimation of W', typically these methods only included TTE trials lasting less than 12 min.

*Conclusions:* Estimations from $CP_{3\text{-}hyp}$ were found to be the most accurate, independently of TTE range. Models that include two trials between 12 and 20 min provide good agreement with the criterion method (for both CP and W').



## 1. Introduction

Since first introduced by Monod and Scherrer[1] as the maximal capacity of a muscle, or muscle group to perform work for a prolonged period of time, the concept of critical power (CP) has been widely used as it presents a useful approximation of the endurance capacity of an individual.[2–4] Typically, CP is determined from a series of 5 time-to-exhaustion trials (TTE) conducted at severe exercise intensities.[5–7] However, several studies suggest that estimates of CP can vary and are influenced by the test protocol design. Factors such as the particular model used, and the duration of the TTE trials can change the CP calculated from the model.[8–11] Researchers use varying models to estimate CP, which are derived from a range of two to seven TTE trials that are not standardized in terms of their duration. Although, we note that the most commonly used method is probably one employing four to five trials and fitted with the two-parameter hyperbolic model ($CP_{2\text{-}hyp}$).[7]

Different studies have focused on either the influence of changing the mathematical model, or the number of repetitions on the derived value for CP. For example, Gaesser et al.,[8] Bull et al.,[10] and Bergstrom et al.[11] investigated the influence of different mathematical models such as exponential ($CP_{exp}$), three parameter hyperbolic ($CP_{3\text{-}hyp}$), $CP_{2\text{-}hyp}$, linear ($CP_{linear}$), and linear 1/time ($CP_{1/time}$) on the determination of CP and the work above CP (W'). These three

* Corresponding author.
*E-mail address:* jmmurias@ucalgary.ca (J.M. Murias).

https://doi.org/10.1016/j.jsams.2017.11.015
1440-2440/

studies found $CP_{3\text{-hyp}}$ and $CP_{exp}$ result in different estimations of CP. Bishop et al.[9] asked their participants to perform five TTE trials ranging from 1 to 10 min in duration in order to evaluate the influence of the length of TTE trials on CP parameter predictions. Using data from only three of the five trials CP was modelled with $CP_{linear}$ and $CP_{2\text{-hyp}}$. Bishop et al. found that a significant difference in modelled CP when the three shortest trials (i.e., $CP_{1,2,3}$), the three longest trials (i.e., $CP_{3,4,5}$), or the first, the third, and the fifth trials (i.e., $CP_{1,3,5}$) were selected. Consequently, the authors suggested that TTE trials of widely varying duration should be used to minimize the influence of shorter trials when modelling CP. However, this investigation did not fit the data from all five TTE trials, and was also limited modelling CP using TTE rides of less than 10 min, about half the longest duration recommended by Morton.[7] Moreover, the aforementioned studies lacked comparisons of the effects of using different mathematical methods and range of TTE trials on W' outcomes.

Given the variety of approaches used in the literature and the effects of different models and combinations of TTE trials, and the lack of a complete comparison of estimations of CP and W', the present investigation aimed to examine the effect of number and range of TTE trials, and equation model on CP and W'. Specifically, we modelled CP and W' using combinations of two to five TTE trials with a variety of different mathematical approaches ($CP_{exp}$, $CP_{3\text{-hyp}}$, $CP_{2\text{-hyp}}$, $CP_{linear}$, and $CP_{1/time}$).

## 2. Methods

Thirteen healthy young participants (9 men and 4 women; mean ± SD values: age, 26 ± 3 year; body mass, 69.0 ± 9.2 kg; height, 174 ± 10 cm) volunteered and gave written informed consent to participate in the study. All participants had previous recreational or competitive cycling experience at the provincial level. Participants were nonsmokers, with no musculoskeletal or cardiorespiratory conditions. The full testing protocol was completed in 3 ± 1 weeks and consisted of: (i) a preliminary maximal ramp incremental test for determination of maximal $\dot{V}O_2$ ($\dot{V}O_{2max}$), and peak power output ($PO_{peak}$); and (ii) five TTE trials for estimation of CP. All procedures were conducted in an environmentally controlled laboratory (i.e. temperature ~21 °C, relative humidity ~36%), at a similar time of the day for each participant, with each test performed on separate days, with a minimum interval of 24 h and a maximum interval of 72 h (most typically 48 h) between tests to ensure appropriate recovery between trials. Participants were instructed to keep their water and carbohydrate intake consistent throughout the protocol, and they were requested not to engage in vigorous physical activity for 24 h prior to each test. Participants were asked not to consume caffeine less than 12 h prior to the test. This study was approved by the Conjoint Health Research Ethics Board of the University of Calgary. The results from $CP_{2\text{-hyp}}$ using five TTE trials have been published as part of a separate study comparing CP with the maximal lactate steady-state.[5]

All exercise tests were performed on an electromagnetically braked cycle ergometer (Velotron Dynafit Pro, Racer Mate, Seattle, WA, USA). Breath-by-breath pulmonary gas exchange, ventilation and heart rate (HR) were continuously measured using a metabolic cart (Quark CPET, COSMED, Rome, Italy), as previously described.[12] Calibration was performed before each test as recommended by the manufacturer. Breath-by-breath $\dot{V}O_2$ data were edited as follows: data points that were 3 SD from the local mean were considered outliers and then removed[13]; trials were time-aligned to the onset of exercise (i.e. time zero representing the onset of the ramp incremental exercise), and averaged into 30-s time bins. $\dot{V}O_{2max}$ was considered as the highest 30-s $\dot{V}O_2$ average throughout the ramp incremental test. $PO_{peak}$ was established as the highest power output achieved at the end of the ramp incremental test.

**Table 1**
Group mean percent $PO_{peak}$, duration, absolute PO, and mechanical work during the five time-to-exhaustion trials for estimation of critical power.

| Trial | $\%PO_{peak}$ | Duration (min) | Absolute PO (W) | Mechanical Work (kJ) |
|---|---|---|---|---|
| 1 | 110 ± 0 | 1.7 ± 0.4 | 413 ± 60 | 42.7 ± 13.4 |
| 2 | 95 ± 0 | 3.2 ± 0.8 | 354 ± 59 | 66.5 ± 16.6 |
| 3 | 80 ± 0 | 7.1 ± 1.8 | 303 ± 50 | 127.9 ± 31.2 |
| 4 | 75 ± 4 | 12.5 ± 1.9 | 281 ± 47 | 209.8 ± 44.0 |
| 5 | 72 ± 4 | 19.4 ± 3.4 | 271 ± 46 | 312.1 ± 57.5 |

For the ramp incremental test, the baseline consisted of participants cycling at 50 W for 4 min, as suggested by Boone and Bourgois,[14] followed by either 1 W every 2 s (30 W $min^{-1}$) (men) or 1 W every 2.4 s (25 W $min^{-1}$) (women) increase in PO.

For the estimation of CP, each participant performed five constant-power output trials to exhaustion which ranged from approximately 1–20 min, as recommended by Morton.[7] The first three TTE trials were performed at 80, 95 and 110% of $PO_{peak}$ (as determined from the preliminary ramp incremental test). The order of the tests was randomly assigned. Subsequently, the other two power outputs were determined to generate an even distribution of TTE between the five trials. Each test was preceded by a 4-min baseline at 20 W, followed by a square-wave transition to the predetermined PO.

For all TTE trials, participants cycled at their preferred pedal cadence (range, 70–105 rpm), which was determined during the preliminary ramp incremental test. The moment of exhaustion was deemed to occur when participants failed to maintain the cadence within 5 rpm of their preferred rate for more than 5 s despite strong verbal encouragement. Participants were blinded to the elapsed time, but they received visual feedback on their pedal cadence.

CP was modelled as follows:

| | | |
|---|---|---|
| i. | $CP_{exp} \rightarrow PO = CP + (P_{max} - CP) * \exp(-t/\tau)$ | Hopkins et al.[15] |
| ii. | $CP_{3\text{-hyp}} \rightarrow t = (W'/PO - CP) + (W'/CP - P_{max})$ | Morton[16] |
| iii. | $CP_{2\text{-hyp}} \rightarrow t = W'/(PO - CP)$ | Hill[17] |
| iv. | $CP_{linear} \rightarrow W_{lim} = W' + CP * t$ | Moritani et al.[18] |
| v. | $CP_{1/time} \rightarrow PO = W' * (1/t) + CP$ | Whipp et al.[19] |

where $P_{max}$ is the maximal instantaneous power (in watts), $\tau$ an undefined time constant, and $W_{lim}$ is the work done (i.e., PO * t) in each predictive trial (in Joules).

When the model was fitted using four trials, two combinations were used: trials 1–4 and trials 2–5. Using three trials, four combinations were performed: trials 1, 2, 3; trials 1, 3, 5; trials 2, 3, 4; and trials 3, 4, 5. Finally, when using two trials in the linear models, four combinations were tested: trials 1 and 2; trials 1 and 5; trials 3 and 4; and trials 4 and 5. Importantly, not every possible combination was reported to avoid superfluous comparisons that would not add predicting value to the model. Instead, we selected the combination of methods that would result in a wide combination of TTE, as well as those often used in the literature. See Table 1 for details on the exercise intensities and durations of the aforementioned TTE trials.

All data editing, processing, and modeling were performed using OriginLab version 9.2 (OriginLab, Northampton, MA).

Data are presented as means ± SD. 90% confidence intervals were calculated and used as a measure of uncertainty (the likely limit of the true value in the population[20]) around each CP and W' values derived from the different methods proposed. Differences between methods were quantified by calculating chances that the true value of a difference was substantial or greater than the smallest worthwhile change (see below). To perform these calculations, we assumed that a substantial difference (in either direction, positive or negative) was larger than 8 W (3.2%) and 1500 J (6.5%) (these are calculated as a constant factor (0.2) multiplied by the between-