



Relationship Between the Critical Power Test and a 20-min Functional Threshold Power Test in Cycling

Bettina Karsten¹, Luca Petrigna², Andreas Klose³, Antonino Bianco², Nathan Townsend⁴ and Christoph Triska^{5,6*}

¹ European University of Applied Sciences (EUFH), Berlin, Germany, ² Sport and Exercise Sciences Research Unit, University of Palermo, Palermo, Italy, ³ Institut für Sportwissenschaft, Westfälische Wilhelms-Universität Münster, Münster, Germany, ⁴ College of Health and Life Sciences, Hamad Bin Khalifa University, Doha, Qatar, ⁵ Institute of Sport Science, Centre for Sport Science and University Sports, University of Vienna, Vienna, Austria, ⁶ Leistungssport Austria, High Performance Unit, Brunn am Gebirge, Austria

OPEN ACCESS

Edited by:

Juan Manuel Murias,
University of Calgary, Canada

Reviewed by:

Danilo Iannetta,
University of Calgary, Canada
Felipe Mattioni Maturana,
Tübingen University Hospital,
Germany

Nicholas Jamnick,
Victoria University, Australia

*Correspondence:

Christoph Triska
christoph.triska@univie.ac.at

Specialty section:

This article was submitted to
Exercise Physiology,
a section of the journal
Frontiers in Physiology

Received: 01 October 2020

Accepted: 30 December 2020

Published: 22 January 2021

Citation:

Karsten B, Petrigna L, Klose A, Bianco A, Townsend N and Triska C (2021) Relationship Between the Critical Power Test and a 20-min Functional Threshold Power Test in Cycling. *Front. Physiol.* 11:613151. doi: 10.3389/fphys.2020.613151

To investigate the agreement between critical power (CP) and functional threshold power (FTP), 17 trained cyclists and triathletes (mean \pm SD: age 31 ± 9 years, body mass 80 ± 10 kg, maximal aerobic power 350 ± 56 W, peak oxygen consumption 51 ± 10 mL \cdot min⁻¹ \cdot kg⁻¹) performed a maximal incremental ramp test, a single-visit CP test and a 20-min time trial (TT) test in randomized order on three different days. CP was determined using a time-trial (TT) protocol of three durations (12, 7, and 3 min) interspersed by 30 min passive rest. FTP was calculated as 95% of 20-min mean power achieved during the TT. Differences between means were examined using magnitude-based inferences and a paired-samples *t*-test. Effect sizes are reported as Cohen's *d*. Agreement between CP and FTP was assessed using the 95% limits of agreement (LoA) method and Pearson correlation coefficient. There was a 91.7% probability that CP (256 ± 50 W) was higher than FTP (249 ± 44 W). Indeed, CP was significantly higher compared to FTP ($P = 0.041$) which was associated with a trivial effect size ($d = 0.04$). The mean bias between CP and FTP was 7 ± 13 W and LoA were -19 to 33 W. Even though strong correlations exist between CP and FTP ($r = 0.969$; $P < 0.001$), the chance of meaningful differences in terms of performance (1% smallest worthwhile change), were greater than 90%. With relatively large ranges for LoA between variables, these values generally should not be used interchangeably. Caution should consequently be exercised when choosing between FTP and CP for the purposes of performance analysis.

Keywords: power-duration relationship, exercise tolerance, fatigue threshold, cycling performance, functional threshold power

INTRODUCTION

Sport scientists, athletes, and coaches intuitively understand that as exercise intensity increases, a point is reached where a maximal metabolic steady state occurs, beyond which perceptions of effort and physiological perturbations progress more rapidly (for review see: Jones et al., 2019). These perceptions of physical discomfort are associated with mechanisms of peripheral fatigue which ultimately lead to task failure (Hureau et al., 2018). During laboratory testing this threshold is usually identified using lactate landmarks (e.g., lactate turning point, maximal lactate steady state or

ventilatory thresholds). These thresholds are protocol dependent and may not align with a maximal oxidative steady state during constant load exercise, which has thus led to controversy (Jannick et al., 2018; Iannetta et al., 2019; Poole et al., 2020). Conventionally, the maximal lactate steady state (MLSS) was believed to reflect a “true” maximal metabolic steady state (MMSS), however, it was recently demonstrated that an oxidative steady state can be maintained despite gradually increasing blood lactate (Brauer and Smekal, 2020). Furthermore, mathematical modeling of lactate kinetics suggests that a true equilibrium between maximal whole body lactate production and oxidation results in a gradually increasing blood lactate concentration (Beneke, 2003). Therefore, since the so-called critical power (CP) has been shown to lie within the intensity region which distinguishes steady state from non-steady state oxidative metabolism (Poole et al., 1988; de Lucas et al., 2013; Vanhatalo et al., 2016), an emerging consensus recognizes CP to more accurately represent a MMSS than the MLSS (Jones et al., 2019; Galan-Rioja et al., 2020).

Over the past 40 years, the CP concept has been studied extensively within the scientific literature and it has emerged as a simple mathematical model which not only describes the relationship between sustainable power and the development of fatigue during high intensity exercise, but which also provides an estimate of the maximal sustainable metabolic rate (Poole et al., 2016; Galan-Rioja et al., 2020). Nevertheless, there are some methodological considerations which might affect an accurate determination of CP (e.g., day-to-day variability, protocol, methodology, mathematical modeling) (Muniz-Pumares et al., 2019). The work rate at CP is closely associated with performance in endurance events (Kranenburg and Smith, 1996; Florence and Weir, 1997; Joyner and Coyle, 2008; Nimmerichter et al., 2017), and moreover, training above or below this MMSS leads to differences in physiological adaptations and specific performance outcomes (Vanhatalo et al., 2011; Iannetta et al., 2018). Due to the increasing availability of affordable on-the-bike power meters though, field-based methods of threshold assessment have been validated (Karsten et al., 2015; Triska et al., 2015). Additionally, the use of specialized and expensive laboratory equipment is not always justified and it requires technical expertise.

More recently, practical methods of threshold assessments have emerged such as field-based CP testing (Karsten et al., 2015; Triska et al., 2015; Muniz-Pumares et al., 2019), and the Functional Threshold Power (FTP). The FTP is defined as “the highest power that a rider can maintain in a quasi-steady state without fatiguing for approximately 1 hour” (Allen and Coggan, 2010), and has become widely popular amongst recreational and competitive cyclists for the purpose of aerobic capacity assessment and training prescription. However, to-date there is controversy as to whether FTP is related to CP or parameters of other threshold concepts (e.g., lactate landmarks or ventilatory thresholds) (Denham et al., 2017; Borszcz et al., 2018; Valenzuela et al., 2018; Sorensen et al., 2019). For example, Valenzuela et al. (2018) found non-significant differences between FTP and the second anaerobic threshold (AnT2) using the D_{\max} method. Using measures of gas exchange, Denham et al. (2017) found the relative maximal oxygen uptake ($\dot{V}O_{2\max}$) to be significantly correlated with FTP. Also, Borszcz et al. (2018) assessed the

agreement between FTP using a 20 min time trial (TT) and a 60 min TT, and the PO associated with the individual anaerobic threshold (IANt; defined as a $1.5 \text{ mmol}\cdot\text{L}^{-1}$ increase above the point of a minimum ratio between blood lactate concentration and work rate). The authors concluded that despite strong correlations, the limits of agreement between the FTP estimates and IANt were too wide to be used interchangeably. A major limitation of some of these aforementioned studies is that FTP was compared to parameters obtained from incremental exercise, which are known to be protocol dependent and may not align with indices like CP or MLSS obtained during sustained constant work rate exercise (e.g., Faude et al., 2009; Jannick et al., 2018).

Critical power was originally defined as an exercise intensity that could be sustained for a “very long time” (Monod and Scherrer, 1965). CP can be determined using maximal self-paced TT efforts. These have been, when compared to the traditional constant power time-to-exhaustion approach, shown to be valid and reliable (Galbraith et al., 2014; Triska et al., 2017, 2020; Karsten et al., 2018). Only a limited number of published studies exist in the scientific literature which examine the relationship between CP derived from different protocols and FTP (MacInnis et al., 2018; Morgan et al., 2019). Therefore, the aim of the present laboratory-based study was to compare CP as an index of MMSS with FTP. We chose a test protocol for CP assessment that is not different from the traditional constant-work rate approach (Triska et al., 2017; Karsten et al., 2018) and also a widely used and recommended 20 min TT for FTP (Denham et al., 2017; Borszcz et al., 2018; Valenzuela et al., 2018), in a cohort of moderately trained cyclists.

MATERIALS AND METHODS

Participants

Participants in this laboratory-based study were 17 moderately trained cyclists and triathletes (mean \pm SD: age 31 ± 9 years, body mass 80 ± 10 kg, maximal aerobic power [MAP] 350 ± 56 W, peak oxygen consumption [$\dot{V}O_{2\text{peak}}$]; $51 \pm 10 \text{ mL}\cdot\text{min}^{-1}\cdot\text{kg}^{-1}$). All procedures performed were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments. Informed consent was obtained from all participants after information of the nature and any risks associated with this study were provided.

Experimental Design

During visit one, $\dot{V}O_{2\text{peak}}$ and MAP were determined during an incremental test. During visit two and three, participants performed either a CP or FTP test in randomized order. All subjects had previous experience at conducting TTs, and they were instructed to give a maximum effort for each test. Participants refrained from heavy exercise in the 24 h prior to testing, and food and caffeine for 3 h prior to testing. For all three visits participants were instructed to arrive at the laboratory in a fully rested and hydrated state. All testing was conducted on a Cyclus2 ergometer (RBM Electronics, Leipzig, Germany), which enables the participant to use their own personal racing bicycle.

Peak Oxygen Uptake Test

After a standardized warm-up at 150 W and for 5 min, participants completed an incremental step test until volitional exhaustion. After 3-min baseline at ~80 W, the test commenced at an intensity of 100 W with a step-like increase of 20 W min⁻¹. Participants self-selected cadence throughout, and when this decreased by more than 10 rev·min⁻¹ for 10 s despite strong verbal encouragement, the test was terminated. Pulmonary gas exchange was measured breath-by-breath using a Cortex MetaLyzer 3B gas analyzer (Cortex Biophysik, Leipzig, Germany), and heart rate (HR) was monitored via the ergometer. Outliers were excluded from further analysis after visual inspection by two independent researchers. MAP was calculated using the following equation:

$$\text{MAP} = P_L + (t/60 \times P_I) \quad (1)$$

where P_L represents the last completed stage (W), t is time for the incomplete stage (s) and P_I is the step increment (W). $\dot{V}O_{2\text{peak}}$ was taken as the highest 30-s rolling-average during the incremental test. The mean \pm SD duration of the peak oxygen test was 13.5 ± 2.8 min.

Critical Power Testing

Critical power was determined using maximal self-paced TT efforts over the durations of 12, 7, and 3 min with a 30 min passive rest between efforts. The protocol started with a 5 min warm-up phase at 100 W immediately followed by a switch of the ergometer into TT mode, where resistance increases or decreases as a function of cadence and pedal force. During the TT, participants were allowed to self-pace via use of a virtual gear changer mounted to the handlebars. Feedback of elapsed time and strong encouragement was provided throughout and participants were asked to produce the highest average PO possible. Heart rate (HR) was measured continuously and rate of perceived exertion (RPE) was recorded immediately at the end of trials. HR within 10 beats of age-predicted HR maximum and RPE values above 18 were taken as an indicator for a maximal effort and accepted as a successful test.

Critical power and its related maximum work above CP (W') were determined using two linear and one hyperbolic model:

$$\text{linear work vs. time : } W_{\text{lim}} = W' + CP \times t \quad (\text{Moritani et al., 1981}) \quad (2)$$

$$\text{linear PO vs. inverse of time : } PO = W' \times (1/t) + CP \quad (\text{Whippet et al., 1982}) \quad (3)$$

$$\text{hyperbolic PO vs. time : } t = W' / (PO - CP) \quad (\text{Hill, 1993}) \quad (4)$$

The inverse time linear model ($P = W'/t + CP$) provided the lowest combined standard error of the estimate (SEE) (i.e., sum of SEE% of CP and W') and was consequently used. The

linear PO vs. inverse of time model provided the lowest combined SEE for all participants ($n = 17$) and was therefore used for further analysis. This model was also used to calculate predicted maximal 20 min PO ($p20_{\text{MMP}}$) values (e.g., $15000 \times (1/1200) + 300$).

Functional Threshold Power Testing

Functional threshold power was estimated from a single 20 min TT effort similar to recent research (Denham et al., 2017; Borszcz et al., 2018; Valenzuela et al., 2018) and described elsewhere (Allen and Coggan, 2010). For consistency with the CP protocol, this TT also commenced with a 5 min warm-up at 100 W. Throughout the 20 min TT, participants were allowed to self-pace, and elapsed time feedback was provided as per CP testing. HR was measured continuously and RPE was recorded immediately at the end of trials. HR within 10 beats of age-predicted HR maximum and RPE values above 18 were taken as an indicator for a maximal effort and accepted as a successful test. FTP was calculated as 95% of the 20 min maximal measured PO (20_{MMP}) obtained during the TT (Allen and Coggan, 2010; Denham et al., 2017; Borszcz et al., 2018; Valenzuela et al., 2018).

Statistics

Data were first examined for normality using the Shapiro-Wilk test. Differences between means were assessed using magnitude based inferences (Batterham and Hopkins, 2006) where the “smallest worthwhile change” in PO considered to be meaningful in terms of practical significance, was set at 1% of the mean CP estimate using an Microsoft Excel spreadsheet (Paton and Hopkins, 2001; Batterham and Hopkins, 2006). The agreement between variables was assessed using 95% limits of agreement (LoA) (Bland and Altman, 1986) using GraphPad Prism (version 6.00 for Mac; GraphPad Software, La Jolla CA, www.graphpad.com). Pearson product moment correlation was used to provide an estimate of strength of association between variables, and linear regression was used to calculate the SEE associated with prediction of 20_{MMP} and FTP from CP, and also $p20_{\text{MMP}}$ and pFTP, respectively. Lin’s concordance coefficient and the intraclass correlation coefficient (ICC) was used to evaluate the agreement between methods assessed by a Microsoft Excel spreadsheet. To assess differences between CP, FTP, and 20_{MMP} a repeated-measures analysis of variance was used. Partial squared eta (η_p^2) was used to calculate for effect sizes (small $\eta_p^2 = 0.02$, moderate $\eta_p^2 = 0.13$, and large $\eta_p^2 = 0.26$). Effect sizes of the *post hoc* tests are reported as Cohen’s d calculated as the quotient of mean differences and variance (small $d = 0.2$; moderate $d = 0.5$; large $d = 0.8$). The typical error of the estimate and the coefficient of variation (%) (Hopkins, 2000) were used to assess validity, where CP was taken as the criterion variable and FTP as practical variable assessed by an Microsoft Excel spreadsheet. Statistical significance was set at $P < 0.05$ and raw data is reported as mean \pm SD. All analyses were conducted using SPSS statistical software package 27 (IBM SPSS statistics, SPSS Inc., Chicago, United States) unless stated otherwise.

RESULTS

All data were normally distributed ($P > 0.05$). Descriptive data is presented in **Table 1** and individual CP and W' results from linear and hyperbolic models including combined SEE are depicted in **Table 2**. Significant main effects and a *large* effect size were found between CP, FTP, and 20_{MMP} ($F_{2,32} = 13.029$; $P < 0.001$; $\eta_p^2 = 0.45$). Bonferroni *post hoc* procedures revealed significant differences and *large* effect sizes only between FTP and 20_{MMP} ($P < 0.001$; $d = 2.44$), but not between CP and FTP ($P = 0.122$; $d = 0.04$), and between CP and 20_{MMP} ($P = 0.200$; $d = 0.04$). CP and FTP were significantly correlated ($r = 0.969$; $P < 0.001$) (**Figure 1A**). There was a 91.7% probability that CP was higher than FTP (likely). The mean bias and the 95% LoA between CP and FTP were 7 ± 13 W (95% LoA: -19 to 33 W) (**Figure 1B**). The Lin's concordance coefficient between CP and FTP was $r = 0.950$ (95% confidence limits, [CL]: 0.877 to 0.980), the ICC was $r = 0.976$ (95% CL: 0.915 to 0.991) and the typical error in raw unit was 13 W (95% CL: 9 to 20 W) and expressed as a coefficient of variation (%) was 5.6% (95% CL: 4.1 to 8.8%). Mean peak HR and mean RPE values for the 20 min TT, 12 min TT, 7 and 3 min TT were as follows: 184 ± 12 b min^{-1} and 19 ± 1 ; 184 ± 13 b min^{-1} and 19 ± 1 ; 181 ± 12 b min^{-1} and 19 ± 1 ; 178 ± 14 b min^{-1} and 19 ± 1 ($P > 0.05$).

DISCUSSION

The primary finding of this study was that for moderately trained cyclist, mean CP was non-significantly higher than FTP (i.e., 95% of 20_{MMP}). The probability of a meaningful difference was $>90\%$, however, the effect size was only of a *trivial* order. The results also revealed wide LoA between CP and FTP (± 24 W), a large bias (-12 W) and large intraindividual variations (± 12 W). Nevertheless, the variables were strongly correlated ($r = 0.969$). According to Cohen, the typical error of the estimate between CP and FTP can be interpreted as small (13 W; 5.6%). This, however, is arguably above the 5%, which is the commonly accepted upper limit in sport science research (Hopkins, 2000). Our results

are supported by Borszcz et al. (2018) who reported notably larger LoA between the IAnT and FTP and 20_{MMP} , respectively, than reported in the present study. Furthermore, our results are consistent with Valenzuela et al. (2018), who identified poor LoA and a significant difference between FTP and the AnT2 in moderately trained athletes. Interestingly, however, LoA between FTP and the AnT2 were stronger and non-significantly different in well trained athletes (Valenzuela et al., 2018).

We observed a mean difference of 7 ± 13 W between CP and FTP. Given that a 1% difference in PO (~ 2.6 W in our study) is considered a "smallest worthwhile change," the disparity between CP and FTP equates to a high probability, that a meaningful difference exists in practice (Paton and Hopkins, 2001, 2006; Batterham and Hopkins, 2006). Even though using different methods, Borszcz et al. (2018) also demonstrated wide LoA between the IAnT and FTP determined from a 20 min TT (-62 to 60 W). Notably in this study is the difference of 5 W (2%) between measured 20_{MMP} (231 W) and predicted 20 min PO (236 W). This might raise questions about the application rule of 5% from the 20 min FTP values as postulated by Allen and Coggan (2010). Although Valenzuela et al. (2018) did not find significant differences and a trivial effect size between FTP and the AnT2, their results also demonstrated wide LoA. Arguably, these findings question the underlying physiology of the FTP concept. With respect to the present study it is noteworthy that Borszcz et al. (2018) and Valenzuela et al. (2018) did not use CP as criterion. According to the correlation between CP and FTP only 6.1% of the variance between CP and FTP ($R^2 = 0.939$) are explained by other factors. Nevertheless, these 6.1% are suggested to notably question the interchangeable use of CP and FTP. Collectively, these findings do not support the assertion by Allen and Coggan (2010) that FTP always corresponds to the highest the PO maintainable in a quasi-steady state. Importantly, the identified differences and error between CP and FTP values in the present study suggest that these estimates of threshold power should not be used interchangeably.

Interestingly, there were no significant differences, apart from a *trivial* effect size, between CP and 20_{MMP} . However, previous research has shown that time-to-fatigue at CP equals ~ 23 min in both untrained (Poole et al., 1988) and trained cyclists (de Lucas et al., 2013). Conversely, CP has been found to reside ~ 20 W above MLSS intensity (Pringle and Jones, 2002) which indicates a clear difference between CP and MLSS (Jones et al., 2019; Galan-Rioja et al., 2020). It is, however, noteworthy that other authors found small differences (~ 1 W) between CP and MLSS (Keir et al., 2015). These different results are suggested to be due to the fact that both CP and MLSS are protocol depended which is a clear limitation. We therefore suggest that CP tends to reside above MLSS. Therefore, a 20 min TT would more closely align with CP and 95% of a 20 min TT would more closely align with MLSS where time-to-fatigue is 55 ± 8.5 min (Baron et al., 2008). This is also confirmed by the suggestions of Jones et al. (2019) that MLSS and CP cannot be used interchangeably as a boundary between the heavy and the severe intensity exercise domain. As a result of these findings, it is suggested that FTP (calculated as 95% of 20_{MMP}) is generally lower than CP.

TABLE 1 | Performance measures obtained from the tests (mean \pm SD).

Measure	Group ($n = 17$)
CP (W)	256 ± 50
CP SEE (W)	5 ± 5
CP SEE (%)	2.9 ± 3.0
W' (kJ)	16.97 ± 6.907
CP SEE (kJ)	1.9 ± 1.5
CP SEE (%)	13.8 ± 16.5
20_{MMP} (W)	262 ± 46
$p20_{MMP}$ (W)	270 ± 51
FTP (W)	249 ± 44
$pFTP$ (W)	261 ± 50

CP, critical power; SEE, standard error of the estimate; W' , maximal work accomplished above CP; R^2 , coefficient of determination; 20_{MMP} , 20 min maximal mean power; $p20_{MMP}$, predicted 20 min maximal mean power; FTP, functional threshold power; $pFTP$, predicted functional threshold power.

TABLE 2 | Individual CP and W' for the linear models (power vs. 1/time and work vs. time) and the hyperbolic model.

No.	linear model (power vs.-1/time)			linear model (work vs. time)			hyperbolic model		
	CP	W'	combined SEE	CP	W'	combined SEE	CP	W'	combined SEE
1	224	13.850	11.9%	228	12.559	20.8%	232	9.679	28.4%
2	222	11.574	18.1%	226	9.891	33.7%	233	5.435	56.4%
3	312	21.347	4.4%	314	20.632	7.1%	316	19.240	8.4%
4	245	8.614	32.3%	238	10.965	42.7%	233	14.027	37.8%
5	317	13.312	3.7%	316	13.715	5.8%	315	14.419	6.2%
6	237	16.467	8.3%	234	17.516	12.5%	231	19.255	12.8%
7	190	23.930	8.2%	194	22.639	12.8%	198	20.016	15.3%
8	301	17.012	11.9%	296	18.614	17.7%	292	21.161	17.6%
9	296	12.125	11.1%	293	13.243	16.8%	290	15.026	16.8%
10	311	24.481	8.3%	307	25.991	12.4%	303	28.502	12.7%
11	193	20.300	23.8%	185	23.642	32.4%	176	28.515	30.5%
12	244	24.146	10.6%	239	25.955	15.5%	234	28.909	15.6%
13	304	26.966	14.1%	297	29.721	20.3%	289	34.055	19.9%
14	220	3.358	24.7%	222	2.609	53.5%	220	3.420	177.5%
15	187	9.480	71.2%	171	15.911	87.5%	158	22.094	84.1%
16	221	26.289	6.3%	218	27.395	9.2%	215	29.290	9.5%
17	333	15.301	1.5%	333	15.486	2.4%	332	15.817	2.6%
mean	256	16.974	15.9%	254	18.028	23.7%	251	19.345	32.5%
SD	50	6.907	16.4%	51	7.276	21.3%	52	8.697	42.6%

CP, critical power; W' , maximal work accomplished above CP; SEE, standard error of the estimate.

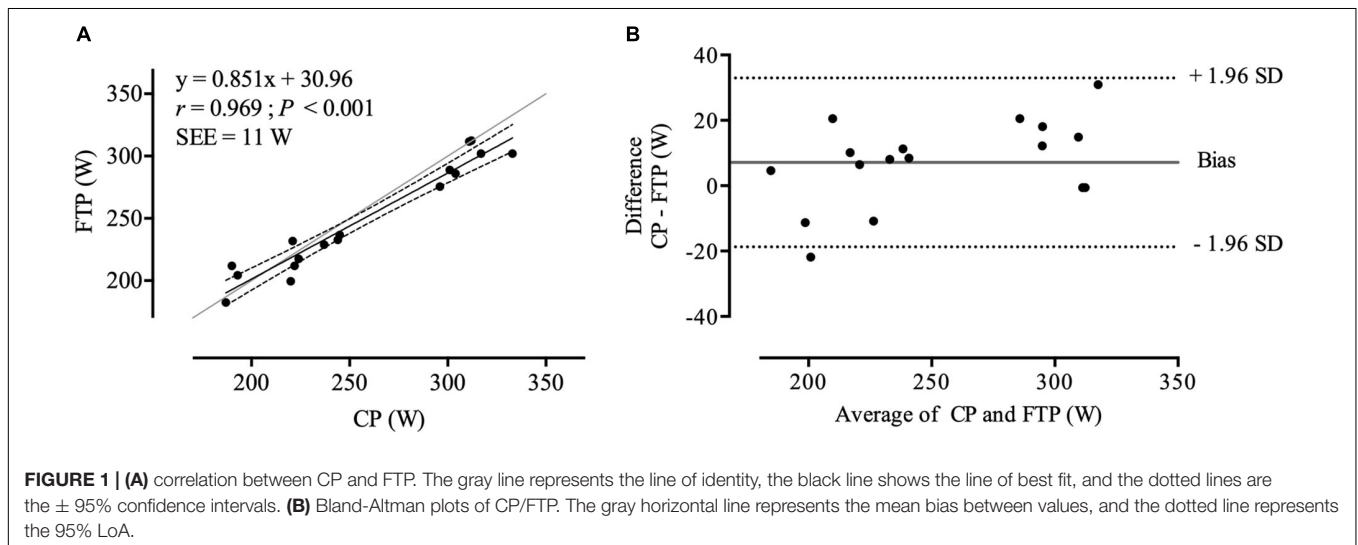


FIGURE 1 | (A) correlation between CP and FTP. The gray line represents the line of identity, the black line shows the line of best fit, and the dotted lines are the \pm 95% confidence intervals. (B) Bland-Altman plots of CP/FTP. The gray horizontal line represents the mean bias between values, and the dotted line represents the 95% LoA.

The difference in CP and FTP reported here raises another important question, that of which estimate of threshold power more closely aligns with the underlying physiological determinants. A recent definition of CP is as follows: “In contrast to historical definitions, CP is now considered to represent the greatest metabolic rate that results in wholly-oxidative energy provision” (Poole et al., 2016). If the ATP demand is not supplied by “wholly-oxidative” metabolism, the additional anaerobic contribution leads to a decline in [PCr] and increasing $\dot{V}O_2$ uptake (Korzeniewski and Rossiter, 2015). Thus, the most direct non-invasive method of validation is to measure *in vivo* muscle metabolism slightly above and below the

estimate of threshold. To-date, only one study has conducted these measurements using ^{31}P magnetic resonance spectroscopy during single-leg knee extension exercise (Jones et al., 2008). These authors reported a progressive loss of [PCr], [P_i], and [H⁺] homeostasis at work rate \sim 10% above CP, but attainment of a steady state when the work rate was 10% below CP. The most direct non-invasive method of physiological validation for whole-body exercise is to measure $\dot{V}O_2$ uptake. Several studies have reported the occurrence of a $\dot{V}O_2$ steady state corresponding to a work rate at, or slightly below CP, whereas non-steady state $\dot{V}O_2$ were observed slightly above CP (Poole et al., 1988; de Lucas et al., 2013; Murgatroyd et al., 2014; Vanhatalo et al., 2016).

In each case, the limit of tolerance was reached markedly sooner at the work rate slightly above CP. Collectively, these studies provide evidence that estimates of CP correspond to an intensity which demarcates steady state from non-steady state oxidative metabolism. It should be noted, however, that estimates of CP are protocol dependent (e.g., Hill and Smith, 1994; Bishop et al., 1998; Mattioni Maturana et al., 2018; Triska et al., 2018), thus only those protocols which have been physiologically validated should be considered to represent threshold intensity. In the current study we used a protocol that has been shown to derive CP and W' estimates with a low SEE and thus a high accuracy (Triska et al., 2017; Karsten et al., 2018).

One of the differences between CP and FTP is that the concept of CP also incorporates W' . During the TT used for the determination of CP, energy contribution is not “wholly-oxidative,” but also includes an energy contribution derived from expenditure of W' . For example, in the present study, mean W' was 16.6 kJ which equates to an additional 4.6 W (averaged over 60 min) above the maximal PO estimated to be “wholly-oxidative.” Since anaerobic energy provision contributes to W' (Poole et al., 2016), then a valid determination of a threshold using a single maximum effort (such as FTP), requires that either no expenditure of W' occurs (PO at or below CP), or if expenditure of W' occurs, then due to some other mechanism a corresponding decrease in average power must also occur. This has to be done to offset the additional energy contribution arising from expenditure of W' . In fact, it should be expected that predicted MMP over a given duration i.e.: $P = W' (1/t) + CP$, where t = the duration of the task, should be close to the actual performance power. Further analysis revealed that p_{20MMP} and 20_{MMP} were significantly different (mean difference: 8 ± 10 W; $P = 0.005$; $d = 0.07$). The discrepancy in mean values here is potentially explained by emergence of fatigue mechanisms that are not dominant during the short duration TTs used to estimate CP. For example, central fatigue has been shown to increase during a 20 and 40 km TTs compared to a 4 km TT (Thomas et al., 2015), and also during a time to exhaustion task lasting ≈ 11 min as compared to one lasting ≈ 3 min (Thomas et al., 2016), thus for TTs lasting even only 20 min, actual sustainable power may diverge from linearity as predicted by the 2-parameter CP model. Also, cycling efficiency at a 60% maximum minute PO intensity has been found to decrease during 2 h cycling below CP (Hopker et al., 2017), and W' has been demonstrated to markedly decrease with glycogen depletion (Miura et al., 2000). Thus, it may be possible for both CP and W' to decline during prolonged higher intensity endurance exercise (i.e., short duration TTs), which suggests the domain of validity of the 2-parameter CP model should be limited to durations less than 20 min.

There are certain key limitations to the present study. In the absence of pulmonary gas exchange and measurements

slightly above and below both CP and FTP, we were unable to validate these estimates according to the physiological criteria which best describes the threshold phenomenon. Thus, further studies are required to establish the physiological validity of the FTP concept. Moreover, the prediction of performance PO beyond 20 min based on the CP model can be questioned. This is a limitation of the power-duration relationship modeling procedure. Furthermore, we did not conduct familiarization trials but used only experienced cyclists or triathletes. Finally, it is currently unclear if different warm-up protocols influence FTP and this should consequently be addressed in further research.

The “smallest worthwhile change” in PO during a sustained TT which approximates threshold intensity is approximately 1% in elite athletes (Paton and Hopkins, 2006). In the present study, such a difference equates to a $\sim 92\%$ chance that CP estimated from the 2-parameter inverse-time model is greater than FTP estimated via the 95% of 20_{MMP} method. Furthermore, the LoA between CP and FTP were wide and therefore, these estimates of threshold intensity should not be used interchangeably.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the University of Greenwich. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

BK conceptualized the study, conducted testing, and wrote parts of manuscript. LP and AB wrote parts of the manuscripts. AK conducted testing and wrote parts of the manuscripts. NT conducted statistical analysis and wrote parts of the manuscript. CT analyzed data and wrote parts of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

Open access funding provided by the University of Vienna.

REFERENCES

- Allen, H., and Coggan, A. (2010). *Training and Racing With a Power Meter*. Boulder: Velo Press.
- Baron, B., Noakes, T. D., Dekerle, J., Moullan, F., Robin, S., Matran, R., et al. (2008). Why does exercise terminate at the maximal lactate steady state intensity? *Br. J. Sports Med.* 42, 828–833. doi: 10.1136/bjism.2007.040444

- Batterham, A. M., and Hopkins, W. G. (2006). Making meaningful inferences about magnitudes. *Int. J. Sports Physiol. Perform.* 1, 50–57. doi: 10.1123/ijssp.1.1.50
- Beneke, R. (2003). Maximal lactate steady state concentration (MLSS): experimental and modelling approaches. *Eur. J. Appl. Physiol.* 88, 361–369. doi: 10.1007/s00421-002-0713-2
- Bishop, D., Jenkins, D. G., and Howard, A. (1998). The critical power function is dependent on the duration of the predictive exercise

- tests chosen. *Int. J. Sports Med.* 19, 125–129. doi: 10.1055/s-2007-971894
- Bland, J. M., and Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1, 307–310. doi: 10.1016/s0140-6736(86)90837-8
- Borszcz, F. K., Tramontin, A. F., Bossi, A. H., Carminatti, L. J., and Costa, V. P. (2018). Functional threshold power in cyclists: validity of the concept and physiological responses. *Int. J. Sports Med.* 39, 737–742. doi: 10.1055/s-0044-101546
- Brauer, E. K., and Smekal, G. (2020). VO₂ steady state at and just above maximum lactate steady state intensity. *Int. J. Sports Med.* 41, 574–581. doi: 10.1055/a-1100-7253
- de Lucas, R. D., de Souza, K. M., Costa, V. P., Grossi, T., and Guglielmo, L. G. A. (2013). Time to exhaustion at and above critical power in trained cyclists: the relationship between heavy and severe intensity domains. *Sci. Sports* 28, e9–e14.
- Denham, J., Scott-Hamilton, J., Hagstrom, A. D., and Gray, A. J. (2017). Cycling power outputs predict functional threshold power and maximum oxygen uptake. *J. Strength Cond. Res.* 34, 3489–3497. doi: 10.1519/JSC.0000000000002253
- Faude, O., Kindermann, W., and Meyer, T. (2009). Lactate threshold concepts: how valid are they? *Sports Med.* 39, 469–490. doi: 10.2165/00007256-200939060-00003
- Florence, S., and Weir, J. P. (1997). Relationship of critical velocity to marathon running performance. *Eur. J. Appl. Physiol. Occup. Physiol.* 75, 274–278. doi: 10.1007/s004210050160
- Galan-Rioja, M. A., Gonzalez-Mohino, F., Poole, D. C., and Gonzalez-Rave, J. M. (2020). Relative proximity of critical power and metabolic/ventilatory thresholds: systematic review and meta-analysis. *Sports Med.* 50, 1771–1783. doi: 10.1007/s40279-020-01314-8
- Galbraith, A., Hopker, J., Lelliott, S., Diddams, L., and Passfield, L. (2014). A single-visit field test of critical speed. *Int. J. Sports Physiol. Perform.* 9, 931–935. doi: 10.1123/ijsp.2013-0507
- Hill, D. W. (1993). The critical power concept. A review. *Sports Med.* 16, 237–254. doi: 10.2165/00007256-199316040-00003
- Hill, D. W., and Smith, J. C. (1994). A method to ensure the accuracy of estimates of anaerobic capacity derived using the critical power concept. *J. Sports Med. Phys. Fitness* 34, 23–37.
- Hopker, J. G., O'Grady, C., and Pageaux, B. (2017). Prolonged constant load cycling exercise is associated with reduced gross efficiency and increased muscle oxygen uptake. *Scand. J. Med. Sci. Sports* 27, 408–417. doi: 10.1111/sms.12673
- Hopkins, W. G. (2000). *A New View on Statistics Internet Society for Sport Science*. Available online at: <http://www.sportsci.org/resource/stats/> (accessed November 22, 2020).
- Hureau, T. J., Romer, L. M., and Amann, M. (2018). The 'sensory tolerance limit': a hypothetical construct determining exercise performance? *Eur. J. Sport Sci.* 18, 13–24. doi: 10.1080/17461391.2016.1252428
- Iannetta, D., de Almeida Azevedo, R., Keir, D. A., and Murias, J. M. (2019). Establishing the Vo₂ versus constant-work-rate relationship from ramp-incremental exercise: simple strategies for an unsolved problem. *J. Appl. Physiol.* 127, 1519–1527. doi: 10.1152/jappphysiol.00508.2019
- Iannetta, D., Inglis, E. C., Fullerton, C., Passfield, L., and Murias, J. M. (2018). Metabolic and performance-related consequences of exercising at and slightly above MLSS. *Scand. J. Med. Sci. Sports* 28, 2481–2493. doi: 10.1111/sms.13280
- Jamnick, N. A., Botella, J., Pyne, D. B., and Bishop, D. J. (2018). Manipulating graded exercise test variables affects the validity of the lactate threshold and VO₂max. *PLoS One* 13:e0199794. doi: 10.1371/journal.pone.0199794
- Jones, A. M., Burnley, M., Black, M. I., Poole, D. C., and Vanhatalo, A. (2019). The maximal metabolic steady state: redefining the 'gold standard'. *Physiol. Rep.* 7:e14098. doi: 10.14814/phy2.14098
- Jones, A. M., Wilkerson, D. P., DiMenna, F., Fulford, J., and Poole, D. C. (2008). Muscle metabolic responses to exercise above and below the "critical power" assessed using 31P-MRS. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 294, R585–R593. doi: 10.1152/ajpregu.00731.2007
- Joyner, M. J., and Coyle, E. F. (2008). Endurance exercise performance: the physiology of champions. *J. Physiol.* 586, 35–44. doi: 10.1113/jphysiol.2007.143834
- Karsten, B., Baker, J., Naclerio, F., Klose, A., Bianco, A., and Nimmerichter, A. (2018). Time trials versus time-to-exhaustion tests: effects on critical power, W', and Oxygen-Uptake Kinetics. *Int. J. Sports Physiol. Perform.* 13, 183–188. doi: 10.1123/ijsp.2016-0761
- Karsten, B., Jobson, S. A., Hopker, J., Stevens, L., and Beedie, C. (2015). Validity and reliability of critical power field testing. *Eur. J. Appl. Physiol.* 115, 197–204. doi: 10.1007/s00421-014-3001-z
- Keir, D. A., Fontana, F. Y., Robertson, T. C., Murias, J. M., Paterson, D. H., Kowalchuk, J. M., et al. (2015). Exercise intensity thresholds: identifying the boundaries of sustainable performance. *Med. Sci. Sports Exerc.* 47, 1932–1940. doi: 10.1249/MSS.0000000000000613
- Korzeniewski, B., and Rossiter, H. B. (2015). Each-step activation of oxidative phosphorylation is necessary to explain muscle metabolic kinetic responses to exercise and recovery in humans. *J. Physiol.* 593, 5255–5268. doi: 10.1113/jp271299
- Kranenburg, K. J., and Smith, D. J. (1996). Comparison of critical speed determined from track running and treadmill tests in elite runners. *Med. Sci. Sports Exerc.* 28, 614–618. doi: 10.1249/00005768-199605000-00013
- MacInnis, M. J., Thomas, A. C. Q., and Phillips, S. M. (2018). The reliability of 4-min and 20-min time trials and their relationships to functional threshold power in trained cyclists. *Int. J. Sports Physiol. Perform.* 14, 38–45. doi: 10.1123/ijsp.2018-0100
- Mattioni Maturana, F., Fontana, F. Y., Pogliaghi, S., Passfield, L., and Murias, J. M. (2018). Critical power: how different protocols and models affect its determination. *J. Sci. Med. Sport* 21, 742–747. doi: 10.1016/j.jsams.2017.11.015
- Miura, A., Sato, H., Sato, H., Whipp, B. J., and Fukuba, Y. (2000). The effect of glycogen depletion on the curvature constant parameter of the power-duration curve for cycle ergometry. *Ergonomics* 43, 133–141. doi: 10.1080/001401300184693
- Monod, H., and Scherrer, J. (1965). The work capacity of a synergic muscular group. *Ergonomics* 8, 329–338. doi: 10.1080/00140136508930810
- Morgan, P. T., Black, M. I., Bailey, S. J., Jones, A. M., and Vanhatalo, A. (2019). Road cycle TT performance: relationship to the power-duration model and association with FTP. *J. Sports Sci.* 37, 902–910. doi: 10.1080/02640414.2018.1535772
- Moritani, T., Nagata, A., deVries, H. A., and Muro, M. (1981). Critical power as a measure of physical work capacity and anaerobic threshold. *Ergonomics* 24, 339–350. doi: 10.1080/00140138108924856
- Muniz-Pumares, D., Karsten, B., Triska, C., and Glaister, M. (2019). Methodological approaches and related challenges associated with the determination of critical power and curvature constant. *J. Strength Cond. Res.* 33, 584–596. doi: 10.1519/JSC.0000000000002977
- Murgatroyd, S. R., Wylde, L. A., Cannon, D. T., Ward, S. A., and Rossiter, H. B. (2014). A 'ramp-sprint' protocol to characterise indices of aerobic function and exercise intensity domains in a single laboratory test. *Eur. J. Appl. Physiol.* 114, 1863–1874. doi: 10.1007/s00421-014-2908-8
- Nimmerichter, A., Novak, N., Triska, C., Prinz, B., and Brees, B. C. (2017). Validity of treadmill-derived critical speed on predicting 5000-meter track-running performance. *J. Strength Cond. Res.* 31, 706–714. doi: 10.1519/JSC.0000000000001529
- Paton, C. D., and Hopkins, W. G. (2001). Tests of cycling performance. *Sports Med.* 31, 489–496. doi: 10.2165/00007256-200131070-00004
- Paton, C. D., and Hopkins, W. G. (2006). Variation in performance of elite cyclists from race to race. *Eur. J. Sport Sci.* 6, 25–31. doi: 10.1080/17461390500422796
- Poole, D. C., Burnley, M., Vanhatalo, A., Rossiter, H. B., and Jones, A. M. (2016). Critical power: an important fatigue threshold in exercise physiology. *Med. Sci. Sports Exerc.* 48, 2320–2334. doi: 10.1249/mss.0000000000000939
- Poole, D. C., Ward, S. A., Gardner, G. W., and Whipp, B. J. (1988). Metabolic and respiratory profile of the upper limit for prolonged exercise in man. *Ergonomics* 31, 1265–1279. doi: 10.1080/00140138808966766
- Poole, D. C., Rossiter, H. B., Brooks, G. A., and Gladden, L. B. (2020). The anaerobic threshold: 50+ years of controversy. *J. Physiol.* doi: 10.1113/JP279963 [Epub ahead of print].
- Pringle, J. S., and Jones, A. M. (2002). Maximal lactate steady state, critical power and EMG during cycling. *Eur. J. Appl. Physiol.* 88, 214–226. doi: 10.1007/s00421-002-0703-4
- Sorensen, A., Aune, T. K., Rangul, V., and Dalen, T. (2019). The validity of functional threshold power and maximal oxygen uptake for cycling

- performance in moderately trained cyclists. *Sports* 7:217. doi: 10.3390/sports7100217
- Thomas, K., Elmeua, M., Howatson, G., and Goodall, S. (2016). Intensity-dependent contribution of neuromuscular fatigue after constant-load cycling. *Med. Sci. Sports Exerc.* 48, 1751–1760. doi: 10.1249/mss.0000000000000950
- Thomas, K., Goodall, S., Stone, M., Howatson, G., St Clair Gibson, A., and Ansley, L. (2015). Central and peripheral fatigue in male cyclists after 4-, 20-, and 40-km time trials. *Med. Sci. Sports Exerc.* 47, 537–546. doi: 10.1249/mss.0000000000000448
- Triska, C., Karsten, B., Beedie, C., Koller-Zeisler, B., Nimmerichter, A., and Tschan, H. (2018). Different durations within the method of best practice affect the parameters of the speed-duration relationship. *Eur. J. Sport Sci.* 18, 332–340. doi: 10.1080/17461391.2017.1418025
- Triska, C., Karsten, B., Heidegger, B., Koller-Zeisler, B., Prinz, B., Nimmerichter, A., et al. (2017). Reliability of the parameters of the power-duration relationship using maximal effort time-trials under laboratory conditions. *PLoS One* 12:e0189776. doi: 10.1371/journal.pone.0189776
- Triska, C., Tschan, H., Tazreiter, G., and Nimmerichter, A. (2015). Critical power in laboratory and field conditions using single-visit maximal effort trials. *Int. J. Sports Med.* 36, 1063–1068. doi: 10.1055/s-0035-1549958
- Triska, C., Hopker, J., Wessner, B., Reif, A., Tschan, H., and Karsten, B. (2020). A 30-min rest protocol does not affect W' , critical power, and systemic response. *Med. Sci. Sports Exerc.* doi: 10.1249/MSS.0000000000002477 [Epub ahead of print].
- Valenzuela, P. L., Morales, J. S., Foster, C., Lucia, A., and de la Villa, P. (2018). Is the functional threshold power (FTP) a valid surrogate of the lactate threshold? *Int. J. Sports Physiol. Perform.* 13, 1–20. doi: 10.1123/ijsp.2018-0008
- Vanhatalo, A., Black, M. I., DiMenna, F. J., Blackwell, J. R., Schmidt, J. F., Thompson, C., et al. (2016). The mechanistic bases of the power-time relationship: muscle metabolic responses and relationships to muscle fibre type. *J. Physiol.* 594, 4407–4423. doi: 10.1113/JP271879
- Vanhatalo, A., Jones, A. M., and Burnley, M. (2011). Application of critical power in sport. *Int. J. Sports Physiol. Perform.* 6, 128–136. doi: 10.1123/ijsp.6.1.128
- Whipp, B. J., Huntsman, D. J., Storer, T., Lamarra, N., and Wasserman, K. (1982). A constant which determines the duration of tolerance to high-intensity work. *Federat. Proc.* 41, 1591–1591.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Karsten, Petrigna, Klose, Bianco, Townsend and Triska. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.