***International Journal of Sports Physiology and Performance*, (Ahead of Print)**
https://doi.org/10.1123/ijspp.2019-0214

Human Kinetics
ORIGINAL INVESTIGATION

Maximal Lactate Steady State Versus the 20-Minute Functional Threshold Power Test in Well-Trained Individuals: "Watts" the Big Deal?

Erin Calaine Inglis, Danilo Iannetta, Louis Passfield, and Juan M. Murias

***Purpose**:* To (1) compare the power output (PO) for both the 20-minute functional threshold power (FTP_{20}) field test and the calculated 95% ($FTP_{95\%}$) with PO at maximal lactate steady state (MLSS) and (2) evaluate the sensitivity of $FTP_{95\%}$ and MLSS to training-induced changes. ***Methods**:* Eighteen participants (12 males: 37 [6] y and 6 females: 28 [6] y) performed a ramp-incremental cycling test to exhaustion, 2 to 3 constant-load MLSS trials, and an FTP_{20} test. A total of 10 participants returned to repeat the test series after 7 months of training. ***Results**:* The PO at FTP_{20} and $FTP_{95\%}$ was greater than that at MLSS ($P = .00$), with the PO at MLSS representing 88.5% (4.8%) and 93.1% (5.1%) of FTP and $FTP_{95\%}$, respectively. MLSS was greater at POST compared with PRE training (12 [8] W) ($P = .002$). No increase was observed in mean PO at FTP_{20} and $FTP_{95\%}$ ($P = .75$). ***Conclusions**:* The results indicate that the PO at $FTP_{95\%}$ is different to MLSS, and that changes in the PO at MLSS after training were not reflected by $FTP_{95\%}$. Even when using an adjusted percentage (ie, 88% rather than 95% of FTP_{20}), the large variability in the data is such that it would not be advisable to use this as a representation of MLSS.

***Keywords**:* exercise testing, cycling, performance, FTP, MLSS

Identifying the critical intensity of exercise is a crucial aspect for predicting performance, prescribing exercise training, and evaluating the effectiveness of training interventions.[1,2] This critical intensity is thought to represent the upper boundary of sustainable performance (ie, the boundary separating tolerable and nontolerable exercise) and is often identified by measures including the maximal lactate steady state (MLSS) or critical power (CP).[3] Although the accuracy for determining this intensity is best obtained in a laboratory setting, this is not always feasible due to cost, accessibility, and time constraints. Thus, field-test protocols are popular among cyclists because they are easily conducted with minimal equipment. Given the practical nature of field tests, they do not entail direct measurement of the physiological responses normally used to confirm the level of exertion (eg, blood lactate concentration [BLa], oxygen uptake [$\dot{V}O_2$]), instead they rely on the maximal voluntary performance.

In cycling, power meters are commonly used for monitoring the cyclist's work rate, and it can be used to measure performance during field-test protocols. Specifically, a popular approach among cyclists is to determine their functional threshold power (FTP_{60}), which is defined as the highest mean power output (PO) that can be achieved during a 60-minute time trial.[4] The PO at FTP_{60} is then used as the basis for prescribing training intensities. Because of the length of this test, a more commonly used protocol is the 20-minute FTP (FTP_{20}) test, from which 95% ($FTP_{95\%}$) of the mean FTP_{20} power is calculated as a prediction of FTP_{60}.[4]

With the increased popularity of the FTP test, comparison has been made between the various time trials and other markers of performance. Specifically, MacInnis et al[5] found that the 20-minute time trial (ie, FTP_{20}) is a reliable test with a strong association with the 60-minute time trial (ie, FTP_{60}) and suggested that it may be an appropriate tool for performance assessment and tracking. However, these authors concluded that the use of 95% could result in an overestimation of FTP_{60} and suggested that a reduction in the percentage of the $FTP_{95\%}$—from 95% to 90%—might be a better predictor of this intensity.[5] By contrast, others have found no difference between the $FTP_{95\%}$ and the lactate threshold,[6] nor between the $FTP_{95\%}$ and the individual anaerobic threshold.[7] Moreover, a comparison between $FTP_{95\%}$ and CP (also closely related with the anaerobic threshold[8]) found a strong correlation and no difference[9] between the 2 variables. However, large limits of agreement were reported and it was concluded that CP and $FTP_{95\%}$ should not be considered equivalent nor used interchangeably.[9]

According to Allen and Cogan,[4] FTP_{60} and $FTP_{95\%}$ represent the highest PO that can be maintained for an extended period (~1 h), a duration, which very closely resembles that reported for exercising at MLSS (~55 min).[10] Despite this claim, no study has experimentally investigated whether $FTP_{95\%}$ is equivalent to MLSS. This is pertinent as many regard MLSS to be the criterion measure for tolerable exercise.[2] Furthermore, although cross-sectional comparisons between $FTP_{95\%}$ and other markers of performance have been made,[5–7,9] no study has evaluated the ability or sensitivity of the $FTP_{95\%}$ test to track changes in fitness level on a longitudinal basis.

Thus, the aims for this study were to assess whether the PO at $FTP_{95\%}$ is similar to that at MLSS and whether the FTP_{20} is sensitive to changes in a fitness status over a 7-month training period. Based on the fact that (1) the FTP is a performance-based test that could be subject to external factors, (2) the $FTP_{95\%}$ is a fixed percentage that does not consider interindividual variations, and (3) it has previously been shown to be an overestimation of FTP_{60}, we hypothesized that the PO at $FTP_{95\%}$ would be different to the PO at MLSS. In addition, we hypothesized that the PO at $FTP_{95\%}$ would be sensitive to changes in fitness level.

Inglis, Iannetta, Passfield, and Murias are with the Faculty of Kinesiology, University of Calgary, Calgary, Alberta, Canada. Passfield is also with the University of Kent, Chatham Maritime, United Kingdom. Murias (jmmurias@ucalgary.ca) is corresponding author.

Methods

Participants

A total of 18 participants (values in mean [SD]; 12 males: 37 [6] y, 180 [6] cm, 79 [8] kg and 6 females: 28 [6] y, 171 [6] cm, 68 [9] kg) volunteered and provided written informed consent to participate in this study. Participants ranged from trained to well-trained athletes. Maximal oxygen uptake ($\dot{V}O_2$max), self-reported training volume, and years of training experience were used to categorize participants with reference to previously established guidelines.[11] All procedures were approved by the Conjoint Health Research Ethics Board at the University of Calgary and complied with the latest version of the Declaration of Helsinki.

Protocol

All testing sessions were performed on an electromagnetically braked cycle ergometer (Velotron Dynafit Pro; Racer Mate, Seattle, WA) in an environmentally controlled laboratory (ie, temperature ~21°C, relative humidity ~36%) over the span of 4 to 5 sessions. For each participant, the time of day was kept consistent, and each session was separated by at least 48 hours and no longer than 72 hours. Participants were asked to refrain from performing vigorous-intensity exercise the day before each session, while also maintaining a similar diet over the course of the testing. Testing sessions included a ramp-incremental test, constant-load trials, and a maximal effort 20-minute time trial (see following sections for details).

The study was separated into 2 separate parts with identical testing procedures. The first part included all participants (n = 18), whereas the second part included 10 returning participants (9 males and 1 female; 39 [5] y, 178 [8] cm; PRE 76 [10] kg, POST 76 [11] kg). For these 10 participants, the first and second parts corresponded to before (PRE) and to the end (POST) of a 7-month cycling season.[12] In addition, for these 10 participants, PRE-season testing corresponded to the 2 months prior to the start of racing season (a period of time during which training consisted predominantly of prolonged endurance sessions). Over the course of the cycling season, these 10 participants trained on average 5 to 6 days per week, for ~1.5 to 4 hours per session. The necessary sample size for sufficient statistical power was n = 10 and was calculated based on the observed differences in a similar study.[5]

Ramp-Incremental Test. The initial visit consisted of a ramp-incremental test to exhaustion to determine $\dot{V}O_2$max and to predict the initial load for determination of MLSS.[13] The ramp-incremental test began with a 4-minute baseline at 50 W followed by a 30 $W{\cdot}min^{-1}$ ramp for males and a 25 $W{\cdot}min^{-1}$ for females.

Constant-Load Trials. The successive visits after the ramp-incremental test included 30-minute constant-PO trials for the determination of MLSS. Participants were instructed to cycle at their preferred cadence that was recorded and kept consistent for the constant-PO trials. MLSS was defined as the highest PO at which a stable blood lactate concentration ([BLa]) ($\Delta \leq$ 1.0 $mmol{\cdot}L^{-1}$) was measured between the 10th and 30th minute of the constant-PO trial.[14] Multiple trials were performed until this criterion was satisfied. Prior to the MLSS trial, a 4-minute baseline ride was performed at 80 W before the PO was instantaneously increased to a predetermined value. Throughout all the testing sessions, the participants were blinded to the PO and elapsed time.

FTP_{20} Test. The Velotron 3D software (Racer Mate, Seattle, WA) was used for the FTP_{20} during which the participants controlled the gearing of the ergometer. Participants were familiarized with the gearing system prior to the test. The test was preceded by an 8-minute baseline at 80 W. For the FTP_{20} test, the participants were familiar with the goal of achieving the highest average PO possible across the 20 minutes, and no verbal encouragement was provided. During the test, participants were blinded to the PO, but they were allowed to see time and cadence to allow for individual pacing strategies.

Measurements

A metabolic cart (Quark CPET; COSMED, Rome, Italy) measured breath-by-breath gas exchange and ventilatory variables. Expired gases were sampled at the mouth and analyzed for fractional concentrations of oxygen (O_2) and carbon dioxide (CO_2), and a low dead space turbine assessed inspired and expired flow rates. Gas and flow calibrations were performed prior to each testing session according to manufacturer recommendations. Heart rate (HR) was continuously monitored (Garmin, Chicago, IL).

Blood lactate concentration was measured with a portable device (Lactate Scout; SensLab GmbH, Leipzig, Germany) from a finger prick during baseline and at 5-minute increments during the 30-minute constant-load trials and the FTP_{20} test.

Psychological measures including the feeling scale (FS) for measures of "affect" and the felt arousal scale (FAS) for "arousal" level were administered before and after each testing session, while a rating of perceived exertion (RPE) scale (0–10) was administered at the same 5-minute intervals as [BLa] was measured.

Data and Analysis

Gas Exchange and Ventilatory Variables. Breath-by-breath $\dot{V}O_2$ data from each test was processed (aberrant data points that were >3 SD from the local mean were removed), time aligned (such that time "zero" represented the onset of the ramp- or constant-load exercise), and then linearly interpolated to 1-second intervals.

Ramp-Incremental Test. The highest $\dot{V}O_2$ computed from a 30-second rolling average was defined as $\dot{V}O_2$max, whereas peak PO was the highest PO value achieved at the end of the ramp-incremental test.

MLSS and FTP. Oxygen uptake at MLSS was determined from the average of the last 10 minutes of the constant-load trial. PO at FTP_{20} was calculated as the average of the entire 20-minute test from which the $FTP_{95\%}$ was derived (95% of the 20-min average). The PO was interpolated into 1-second intervals and to 5-minute bins for statistical comparison (ie, PRE vs POST).

HR, RPE, FS, FAS, and [BLa]. Heart rate was taken as the average of the last 2 minutes of exercise, and the RPE collected during the final minute of exercise was used for comparison of PRE with POST. Presession FS and FAS measurements were used for analysis. End [BLa] represents the sample taken in the 30th minute and 20th minute for MLSS and FTP_{20}, respectively.

Pacing. Changes in pacing strategy were evaluated from PRE to POST by finding the average PO within 5-minute segments during the FTP_{20}.

Statistics

All data processing and modeling were performed with a commercially available computer software (OriginLab, Northampton, MA), and statistical analysis was performed using SPSS (version 23; SPSS, Chicago, IL) with statistical significance set at a $P < .05$. Descriptive data are presented as mean (SD). Paired samples *t* tests were used to evaluate differences in the PO at FTP_{20} and $FTP_{95\%}$ compared with MLSS, in addition to differences in physiological and psychological measures (HR, [BLa], RPE, FS, and FAS) from MLSS and FTP_{20} within the same testing period. A repeated-measures analysis of variance was used to evaluate these differences from PRE to POST. Bland–Altman analyses were used to test for agreement between the PO at MLSS and $FTP_{95\%}$, whereas the association between values of $\dot{V}O_2$ and the PO were tested by linear regression analysis and Pearson's product–moment correlations. Paired samples *t* tests were used to evaluate changes in pacing strategy from PRE to POST at each 5-minute segment.

Results

Full-Group PRE

Overall, $\dot{V}O_2max$ was 4.00 (0.68) $L \cdot min^{-1}$ with a peak PO of 394 (67) W. The PO at FTP_{20} and $FTP_{95\%}$ was greater than that at MLSS ($P < .05$; Table 1), with PO at MLSS representing 88.5% (4.8%) and 93.1% (5.1%) of FTP and $FTP_{95\%}$, respectively. There was a strong correlation between MLSS and $FTP_{95\%}$ (Figure 1, right panel), with a significant mean difference (ie, bias) between the PO observed at MLSS compared with $FTP_{95\%}$ (Figure 1, left panel). Mean HR and RPE were 162 (8) $beats \cdot min^{-1}$ and 5.0 (1.7) at MLSS and 175 (8) $beats \cdot min^{-1}$ and 8.4 (1.4) at FTP_{20}, respectively. Mean change in [BLa] for MLSS was 0.7 (0.3) $mmol \cdot L^{-1}$. Mean end [BLa] was 4.3 (1.2) $mmol \cdot L^{-1}$ for MLSS and 12.3 (2.6) $mmol \cdot L^{-1}$ for FTP_{20}.

Table 1 Power Output Values During MLSS, the FTP_{20}, and 95% FTP_{20}

	MLSS	FTP_{20}	$FTP_{95\%}$
Power output, W	243 (48)	275 (48)*	261 (45)*

Abbreviations: FTP_{20}, 20-Minute functional threshold power test; MLSS, maximal lactate steady state. Note: Values are in mean (SD).
*Significantly greater than MLSS ($P < .05$).

PRE to POST Responses

For the 10 participants who completed both phases of the study, no increase in $\dot{V}O_2max$ was observed from PRE (4.32 [0.53] $L \cdot min^{-1}$, 56.6 [4.3] $mL \cdot kg \cdot min^{-1}$) to POST (4.37 [0.60] $L \cdot min^{-1}$, 57.7 [7.9] $mL \cdot kg \cdot min^{-1}$) ($P = .45$). Mean change in [BLa] for the MLSS trials (Δ [BLa] from the 10th to 30th minute) was 0.7 (0.3) $mmol \cdot L^{-1}$ at PRE and 0.7 (0.3) $mmol \cdot L^{-1}$ at POST. Table 2 presents mean HR, end [BLa], RPE, FS, and FAS for MLSS and FTP_{20} at PRE and POST.

Table 3 presents PRE and POST values of MLSS, FTP_{20}, and $FTP_{95\%}$. MLSS was greater at POST compared with PRE for both the PO (+12 [8] W; range +2 to 28 W) ($P = .00$) and $\dot{V}O_2$ (PRE 3.63 [0.51] $L \cdot min^{-1}$, POST 3.77 [0.51] $L \cdot min^{-1}$; +0.14 [0.13] $L \cdot min^{-1}$; range −0.01 to + 0.37 $L \cdot min^{-1}$) ($P = .01$). No increase was observed in the mean PO at FTP_{20} (range −18 to +26 W) and $FTP_{95\%}$ (−17 to +25 W) ($P = .75$). Bland–Altman and a correlation analysis of changes in the PO at MLSS and $FTP_{95\%}$ from PRE to POST are shown in Figure 2. At PRE, $FTP_{95\%}$ represented 88% (6%) of the PO at MLSS (range −8 to +51 W; bias = −20 W), whereas at POST, $FTP_{95\%}$ values represented 92% (5%) of the PO at MLSS (range −31 to +28 W; bias −9 W). No difference in PO was found at any of the 5-minute segments from PRE to POST ($P = .48$–.96) (Figure 3).

Discussion

The main goal of this study was to evaluate the ability of the FTP_{20} test to predict the PO associated with MLSS. As hypothesized, despite a strong correlation between the PO at $FTP_{95\%}$ and MLSS, the calculated $FTP_{95\%}$ overestimated the PO corresponding to MLSS (ie, bias = −17 W) with a large variability between the measures (ie, differences ranging from −8 to +51 W for $FTP_{95\%}$). A second goal of this study was to evaluate the ability of the $FTP_{95\%}$ to reflect changes in fitness on a longitudinal basis. Contrary to our hypothesis, the results of this study indicate that the PO at $FTP_{95\%}$ was not sensitive to changes in MLSS, as improvements in this marker were not reflected in the $FTP_{95\%}$.

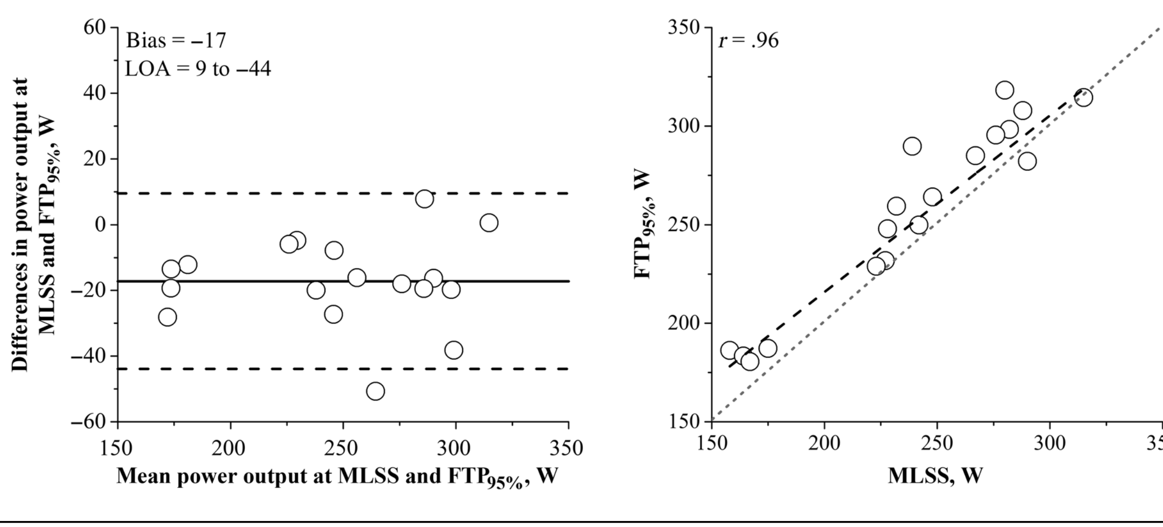

Figure 1 — Bland–Altman plot analysis (left) showing differences in power output at MLSS and $FTP_{95\%}$ for all participants. Correlation graph (right) between power output at MLSS and $FTP_{95\%}$ (dashed gray line indicates line of identity). $FTP_{95\%}$ indicates 95% of the 20-minute functional threshold power test; MLSS, maximal lactate steady state.

Table 2 Mean HR, End [BLa], RPE, and Pretest FS, and FAS for the MLSS and the 20-Minute FTP at PRE and POST for 10 Participants

	MLSS		FTP_{20}	
	PRE	POST	PRE	POST
HR, beats·min^{-1}	162 (6)	161 (5)	174 (8)**	170 (8)*,**
End [BLa], mmol·L^{-1}	4.6 (1.2)	6.0 (1.3)*	13.1 (2.4)**	12.3 (2.5)**
RPE (1–10)	4.1 (0.9)	4.6 (1.0)	7.7 (1.3)**	7.4 (0.7)**
FS (+5 to −5)	3.4 (1.4)	3.0 (1.3)	2.3 (1.3)**	2.8 (1.9)
FAS (1–6)	2.0 (1.3)	2.3 (1.5)	2.2 (1.3)	2.5 (1.5)

Abbreviations: [BLa], blood lactate concentration; FAS, Felt Arousal Scale; FS, Feeling Scale; FTP, functional threshold power test; HR, heart rate; MLSS, maximal lactate steady state; POST, end; PRE, before; RPE, ratings of perceived exertion. *Significantly different than PRE. **Significantly different than MLSS.

Table 3 Power Output Values During MLSS, the 20-Minute Functional Threshold Power Test (FTP_{20}), and 95% of the FTP_{20} ($FTP_{95\%}$) at PRE and POST for 10 Participants

	MLSS	FTP_{20}	$FTP_{95\%}$
Power output, W			
PRE	252 (39)	286 (37)**	272 (35)**
POST	264 (38)*	288 (41)**	273 (38)

Abbreviations: FTP, functional threshold power test; MLSS, maximal lactate steady state; POST, end; PRE, before. Values are in mean (SD). *Significantly different than PRE. **Significantly greater than MLSS.

Relevance of FTP_{20} Testing From PRE Training Data

The $FTP_{95\%}$ derived from the FTP_{20} test has recently become a widespread approach thought to be able to estimate the PO associated with the critical intensity of exercise. This study compared the PO at $FTP_{95\%}$ derived from the FTP_{20} test with that at MLSS, which represents the upper limit for metabolic steady state during continuous exercise.[2] The results of this study demonstrate that a PO lower than the recommended 95% of the FTP_{20} was associated with MLSS. Although the results of this study indicate that 88.5% (4.8%) of the FTP_{20} is more likely to reflect the PO at MLSS, the large amount of variability in the agreement for these measures (limits of agreement = 9 to −44 W) prevents the use of this percent value with any confidence as a superior approximation of MLSS. In this regard, MacInnis et al[5] previously reported that $FTP_{95\%}$ exceeds the PO for the FTP_{60} test, and that the PO at FTP_{60} represented 90% (confidence interval 88%–92%) of that achieved during an FTP_{20} test, which is in good agreement with this present study. Taken together, these data are in accordance with our hypothesis and indicate not only that using a PO of 95% of the FTP_{20} seems to be an overestimation of the actual PO associated with MLSS, but also that even when using a lower percentage of the FTP_{20}, there is large interindividual variability inherent in this prediction. This may be partly related to the fact that both oxidative and nonoxidative energetic pathways contribute to the overall FTP_{20} performance but that their proportional contributions may vary between individuals. In this context, the discrepancy between these measures is concerning if trying to use $FTP_{95\%}$ as a proxy for MLSS, as previous research has demonstrated that exercising at only 10 W above MLSS profoundly reduces subsequent

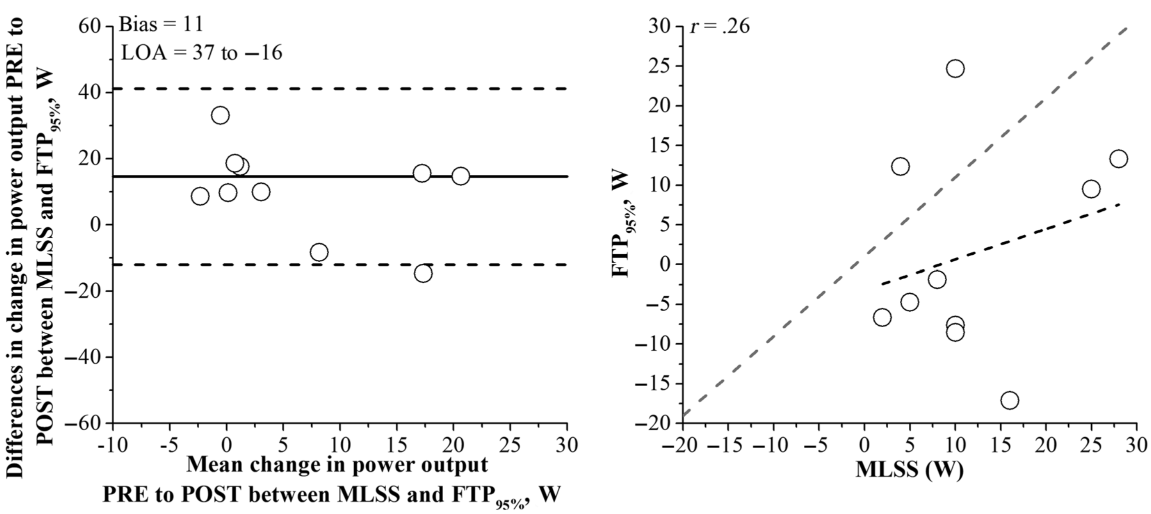

Figure 2 — Bland–Altman plot analysis (left) showing changes in power output from PRE to POST at MLSS and 20-minute $FTP_{95\%}$. Correlation graph (right) between change in power output at MLSS and $FTP_{95\%}$ (dashed gray line indicates line of identity). FTP indicates functional threshold power test; MLSS, maximal lactate steady state; POST, end; PRE, before.

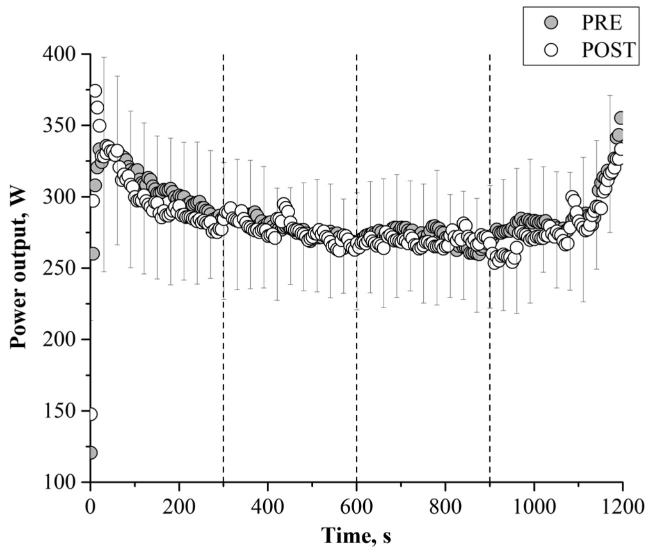

Figure 3 — Power output during the FTP_{20} at PRE (gray circles, positive SD) and POST (white circles, negative SD). FTP_{20} indicates 20-minute functional threshold power test; POST, end; PRE, before.

performance ability (ie, time to task failure is substantially reduced).[15] Furthermore, inaccurate estimations of this intensity could change prescription of intended training-intensity zones.

Although this study adopted MLSS as the criterion measure for the upper limit of metabolic stability and compared this with the $FTP_{95\%}$, other studies have investigated the correspondence between FTP tests and different markers of critical intensity. For example, Morgan et al,[9] found a close relationship between CP (275 [42] W) and $FTP_{95\%}$ (278 [42] W). However, similar to our results, they found that the corresponding limits of agreement (+10.9% to 13.1%) exceeded those that would allow the 2 measures to be used interchangeably with a high level of confidence. In addition, in the previously mentioned study of MacInnis et al,[5] the authors found that 95% of the FTP_{60} (confidence interval 92%–98%) was equal to CP. Although this may be in contrast to our results, it should be noted that the authors used a 2-trial linear model (including a 4-min trial) that might have overestimated the PO at CP.[16] It is important to highlight that although CP and MLSS share a similar definition, they reflect 2 different methods to derive the PO at critical intensity.[17,18] As briefly mentioned, estimates of CP are affected by the testing protocol, the mathematical model used, and the data fitting strategy,[16,19] and in some circumstances it has been shown to elicit POs greater than that at MLSS.[20] Thus, caution is warranted when comparing the POs at MLSS and CP in relation to those derived from FTP testing. Given the great variability in measures of FTP compared with other markers of critical intensity, caution should be exerted before using $FTP_{95\%}$ as one-size-fits-all approach to predicting critical intensities of exercise.[7]

Effects of Training on FTP and MLSS

This study found that the PO at MLSS was greater at POST compared with PRE. Surprisingly, the increase in the PO associated with MLSS did not translate into an improvement in the FTP_{20} test, as evidenced by the fact that the improvements in MLSS from PRE to POST correlated poorly with changes (or the lack thereof) for FTP_{20} from PRE to POST, and did not translate to improvements in $FTP_{95\%}$. This is an important finding as it would be expected that a greater PO from MLSS, a physiologically validated test that determines the highest intensity corresponding with stable metabolic responses, should be related to performance improvements during a similarly challenging FTP test. Given that it is important that measurements are sensitive to small but meaningful changes in performance, as well as being valid and reliable,[21,22] the present data question the ability of the FTP_{20} test to accurately track those changes. It could be argued that the average increase of MLSS was relatively small (ie, 12 W); however, it is likely that well-trained populations have a smaller opportunity for improvement, and that changes are of a smaller magnitude compared with untrained populations.[23,24] From this perspective, it could be possible that individuals of lower fitness level undergoing training programs may display greater changes in MLSS that may also better relate to performance changes in FTP_{20}, as in this population greater relative improvements in aerobic fitness can be expected.[25] Therefore, the results of this study indicate that the $FTP_{95\%}$ may not be sensitive enough to detect small physiological training adaptations occurring in well-trained individuals. Alternatively, improvements in MLSS may solely indicate changes in physiology that may not encompass all components of performance, which however seems unlikely given the tight association between MLSS and exercise capacity.[2]

Although the reasons why the increase in the PO at MLSS did not translate into an improvement in the FTP_{20} test cannot be fully elucidated from this study, it should be acknowledged that the $FTP_{95\%}$ itself is a performance-based test, and the ability of the test to track the actual changes in performance relies upon participants exerting maximal effort. In this context, it is important to consider that this performance may be influenced by other factors.[26] Although laboratory-based testing procedures ensure that tests are performed in standardized and well-controlled conditions for the majority of factors that might influence performance, the psychological state of the participants (eg, motivation) cannot be controlled. Even though there were no differences from PRE to POST in the FS and FAS measures, no direct measures of motivation were taken in this study, and thus it is possible that motivation to provide a maximal effort changed toward the end of the season. This may be a limitation of FTP testing as, in addition to the possibility that small changes in fitness are not detected with the test, other factors such as motivation are more likely to jeopardize a performance-based protocol compared with a laboratory-based test (ie, MLSS). In addition, it has been shown that in some circumstances, experience and training status also can influence the reliability of a time-trial test[27]; however, it is unlikely that this played a role in our study as the cyclists involved in the postmeasurements were the most familiar with the FTP_{20} test and were also among the individuals in this study with the highest training status. In fact, in well-trained cyclists, time-trial performance is reported to be highly reproducible, despite the fact that pacing strategy can be subject to variability.[28] Furthermore, it has been shown that even if the time trials performed differ in duration, when the absolute PO and overall pacing strategy are expressed against relative exercise duration, well-trained athletes show minimal differences between conditions.[29] Regardless, we did not find differences in the pacing strategy employed by the participants between PRE and POST measurements of the FTP_{20}, thus it is unlikely that this played a role.

As MLSS testing is not readily and easily accessible to every individual and discrepancies in the predictive ability of the $FTP_{95\%}$ have been shown in our results and those presented by others,[5] there may be the need to develop an alternative approach to the $FTP_{95\%}$ that is reliable, valid, and convenient for cyclists. Based on these data, it could be suggested that the use of the $FTP_{95\%}$ on its own does not closely estimate the critical intensity of exercise and does not seem to effectively monitor changes in performance. Thus, future studies are warranted to develop alternative field-test protocols that produce a closer approximation of the PO at MLSS.

Conclusions

The results from this study indicate that the $FTP_{95\%}$ does not provide an accurate representation of the PO at MLSS. Even with an adjusted percentage (ie, 88% rather than 95% of FTP_{20} representing a value for FTP_{60}), the large variability in the data is such that it would not be advisable to use the $FTP_{95\%}$ test to estimate MLSS. Furthermore, the results demonstrated that the POs from the $FTP_{95\%}$ are not sensitive to small but meaningful and significant changes in fitness level, and thus its use as a tool for monitoring training may be limited.

Acknowledgments

We express our gratitude to the participants in this study. E.C.I. was supported by the Natural Sciences and Engineering Research Council (NSERC) Master's Award. This study was supported by funding from NSERC (RGPIN-2016-03698). Dr J.M.M. is supported by the Heart and Stroke Foundation of Canada National New Investigator Award (# 1047725). The results of the study are presented clearly, honestly, and without fabrication, falsification, or inappropriate data manipulation. No conflicts of interest, financial or otherwise, are declared by the authors. All experimentation was performed in the Exercise Physiology Laboratory at the University of Calgary. E.C.I., D.I., L.P., and J.M.M. conceived and designed the research. E.C.I. and D.I. performed the experiments. E.C.I. and D.I. analyzed the results. E.C.I., D.I., L.P., and J.M.M. interpreted the results of the experiment. E.C.I. drafted the article. E.C.I., D.I., L.P., and J.M.M. edited and revised the article. E.C.I., D.I., L.P., and J.M.M. approved the final version of article.

References

1. Poole DC, Burnley M, Vanhatalo A, Rossiter HB, Jones AM. Critical power: an important fatigue threshold in exercise physiology. *Med Sci Sport Exerc*. 2016;48(11):2320–2334. doi:10.1249/MSS.0000000000000939
2. Billat VL, Sirvent P, Py G, Koralsztein J-P, Mercier J. The concept of maximal lactate steady state. *Sport Med*. 2003;33(6):407–426. doi:10.2165/00007256-200333060-00003
3. Keir DA, Fontana FY, Robertson TC, et al. Exercise intensity thresholds: identifying the boundaries of sustainable performance. *Med Sci Sport Exerc*. 2015;47(9):1932–1940. doi:10.1249/MSS.0000000000000613
4. Allen H, Coggan A. *Training and Racing with a Power Meter*. Boulder, CO: VeloPress; 2006.
5. MacInnis MJ, Thomas ACQ, Phillips SM. The reliability of 4-minute and 20-minute time trials and their relationships to functional threshold power in trained cyclists. *Int J Sports Physiol Perform*. 2018;14(1):1–27.
6. Valenzuela PL, Morales JS, Foster C, Lucia A, de la Villa P. Is the functional threshold power (FTP) a valid surrogate of the lactate threshold? *Int J Sports Physiol Perform*. 2018;13(10):1–6.
7. Borszcz FK, Tramontin AF, Bossi AH, Carminatti LJ, Costa VP. Functional threshold power in cyclists: validity of the concept and physiological responses. *Int J Sports Med*. 2018;39(10):737–742. PubMed ID: 29801189 doi:10.1055/s-0044-101546
8. Moritani T, Nagata A, deVries HA, Muro M. Critical power as a measure of physical work capacity and anaerobic threshold. *Ergonomics*. 1981;24(5):339–350. PubMed ID: 7262059 doi:10.1080/00140138108924856
9. Morgan PT, Black MI, Bailey SJ, Jones AM, Vanhatalo A. Road cycle TT performance: relationship to the power-duration model and association with FTP. *J Sports Sci*. 2018;37(8):902–910. PubMed ID: 30387374 doi:10.1080/02640414.2018.1535772
10. Baron B, Noakes TD, Dekerle J, et al. Why does exercise terminate at the maximal lactate steady state intensity? *Br J Sports Med*. 2008;42:828–833. PubMed ID: 18070803 doi:10.1136/bjsm.2007.040444
11. De Pauw K, Roelands B, De Geus B, Meeusen R. Guidelines to classify subject groups in sport- science research. *Int J Sports Physiol Perform*. 2013;8:111–122. doi:10.1123/ijspp.8.2.111
12. Inglis EC, Iannetta D, Keir DA, Murias JM. Training-induced changes in the RCP, [HHb] BP and MLSS: evidence of equivalence. *Int J Sports Physiol Perform*. 2019;1–23. doi:10.1123/ijspp.2019-0046
13. Iannetta D, Fontana FY, Maturana M, et al. An equation to predict the maximal lactate steady state from ramp incremental exercise test data in cycling. *J Sci Med Sport*. 2018;21(12):1274–1280. PubMed ID: 29803737 doi:10.1016/j.jsams.2018.05.004
14. Beneke R. Methodological aspects of maximal lactate steady state—implications for performance testing. *Eur J Appl Physiol*. 2003;89(1):95–99. PubMed ID: 12627312 doi:10.1007/s00421-002-0783-1
15. Iannetta D, Inglis EC, Fullerton C, Passfield L, Murias JM. Metabolic and performance-related consequences of exercising at and slightly above MLSS. *Scand J Med Sci Sports*. 2018;28(12):2481–2493. PubMed ID: 30120803 doi:10.1111/sms.13280
16. Mattioni Maturana F, Fontana FY, Pogliaghi S, Passfield L, Murias JM. Critical power: how different protocols and models affect its determination. *J Sci Med Sport*. 2018;21(7):742–747. doi:10.1016/j.jsams.2017.11.015
17. Keir DA, Mattioni Maturana F, Murias JM. When is it appropriate to compare critical power to maximal lactate steady-state? *Appl Physiol Nutr Metab*. 2018;43(1):96–97. doi:10.1139/apnm-2017-0552
18. Jones AM, Vanhatalo A, Burnley M, Morton RH, Poole DC. Critical power: implications for determination of VO_2max and exercise tolerance. *Med Sci Sport Exerc*. 2010;42(10):1876–1890. doi:10.1249/MSS.0b013e3181d9cf7f
19. Bishop D, Jenkins DG, Howard A. The critical power function is dependent on the duration of predictive tests chosen. *Int J Sports Med*. 1998;19:125–129. PubMed ID: 9562222 doi:10.1055/s-2007-971894
20. Mattioni Maturana F, Keir DA, McLay KM, Murias JM. Can measures of critical power precisely estimate the maximal metabolic steady-state? *Appl Physiol Nutr Metab*. 2016;41(11):1197–1203. PubMed ID: 27819154 doi:10.1139/apnm-2016-0248
21. Hopkins WG. Measures of reliability in sports medicine and science. *Sport Sci Rev*. 2000;30(1):1–15.
22. Hopkins WG, Schabort EJ, Hawley JA. Reliability of power in physical performance tests. *Sport Med*. 2001;31(3):211–234. doi:10.2165/00007256-200131030-00005
23. Sjodin B, Jacobs I, Svedenhag J. Changes in onset of blood lactate accumulation (OBLA) and muscle enzymes after training at OBLA. *Eur J Appl Physiol Occup Physiol*. 1982;49:45–57. PubMed ID: 6213407 doi:10.1007/BF00428962
24. Denis C, Fouquet R, Poty P, Geyssant A, Lacour J. Effect of 40 weeks of endurance training on the anaerobic threshold. *Int J Sports Med*. 1982;3:208–214. PubMed ID: 7152767 doi:10.1055/s-2008-1026089

25. Hickson RC, Bomze HA, Holloszy JO. Linear increase in aerobic power induced by a strenuous program of endurance exercise. *J Appl Physiol*. 1977;42(3):372–376. PubMed ID: 838658 doi:10.1152/jappl.1977.42.3.372
26. Faria EW, Parker DL, Faria IE. The science of cycling factors affecting performance – part 2. *Sport Med*. 2005;35(4):313–337. doi:10.2165/00007256-200535040-00003
27. Zavorsky GS, Murias JM, Gow J, et al. Laboratory 20-km cycle time trial reproducibility. *Int J Sports Med*. 2007;28(9):743–748. PubMed ID: 17455116 doi:10.1055/s-2007-964969
28. Thomas K, Stone MR, Thompson KG, St Clair Gibson A, Ansley L. Reproducibility of pacing strategy during simulated 20-km cycling time trials in well-trained cyclists. *Eur J Appl Physiol*. 2012;112(1):223–229. PubMed ID: 21533808 doi:10.1007/s00421-011-1974-4
29. Nicolò A, Marcora SM, Sacchetti M. Respiratory frequency is strongly associated with perceived exertion during time trials of different duration. *J Sports Sci*. 2016;34(13):1199–1206. doi:10.1080/02640414.2015.1102315