MLSS vs the 20-minute FTP test in well-trained individuals: “Watts” the big deal?

Erin Calaine Inglis\textsuperscript{1}, Danilo Iannetta\textsuperscript{1}, Louis Passfield\textsuperscript{1,2} and Juan M. Murias\textsuperscript{1}

\textsuperscript{1}Faculty of Kinesiology, University of Calgary, Calgary, AB, Canada

\textsuperscript{2}University of Kent, Chatham Maritime, UK.

Running head: MLSS vs 20-minute FTP

Author Contribution: All experimentation was performed in the Exercise Physiology Laboratory at the University of Calgary. ECI, DI, LP and JMM conceived and designed the research. ECI and DI performed the experiments. ECI and DI analyzed the results. ECI, DI, LP and JMM interpreted the results of the experiment. ECI drafted the manuscript. ECI, DI, LP and JMM edited and revised the manuscript. ECI, DI, LP and JMM approved the final version of manuscript.

Corresponding author:
Dr. Juan M Murias
Assistant Professor,
Faculty of Kinesiology, University of Calgary
KNB 434, 2500 University Drive NW, Calgary, Alberta, Canada, T2N 1N4
+1 (403) 220-7955
jmmurias@ucalgary.ca

Key Words: exercise testing; functional threshold power; cycling; performance; maximal lactate steady state

Abstract word count: 206; Text-only word count: 3705

Figures: 3; Tables 3
Abstract

PURPOSE This study aimed to i) compare the power output (PO) for both the 20-minute functional threshold power (FTP\textsubscript{20}) field-test and the calculated 95% (FTP\textsubscript{95%}) with power at maximal lactate steady state (MLSS); and ii) evaluate the sensitivity of FTP\textsubscript{95%} and MLSS to training induced changes.

METHODS

Eighteen participants (12 males- 37± 6 years; 6 females- 28 ± 6 years) performed a ramp-incremental cycling test to exhaustion, 2 to 3 constant load MLSS trials, and a FTP\textsubscript{20} test. 10 participants returned to repeat the test series after 7-months of training.

RESULTS

PO at FTP\textsubscript{20} and FTP\textsubscript{95%} was greater than that at MLSS (p=0.00), with the PO at MLSS representing 88.5 ± 4.8% and 93.1 ± 5.1% of FTP and FTP\textsubscript{95%}, respectively. MLSS was greater POST compared to PRE training (12 ± 8 W) (p=0.002). No increase was seen in mean PO at FTP\textsubscript{20} and FTP\textsubscript{95%} (p=0.75).

CONCLUSIONS

The results indicate that the PO at FTP\textsubscript{95%} is different to MLSS, and that changes in MLSS PO after training were not reflected by FTP\textsubscript{95%}. Even using an adjusted percentage (i.e., 88% rather than 95% of FTP\textsubscript{20}), the large variability in the data is such that it would not be advisable to use this as a representation of MLSS.
Introduction

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 (i.e., the boundary separating tolerable and non-tolerable exercise) and is often identified by measures including the maximal lactate steady state (MLSS) or critical power (CP)\(^3\). While 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 amongst cyclists as 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 (e.g., blood lactate concentration ([BLa]), oxygen uptake (\(\text{VO}_2\))), instead they rely on the maximal voluntary performance.

In cycling, power meters are commonly used for monitoring the cyclist’s work rate and can be used to measure performance during field-test protocols. Specifically, a popular approach amongst 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\(^4\). FTP\(_{60}\) PO 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 test (FTP\(_{20}\)), 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 (i.e. FTP\(_{20}\)) is a reliable test with a strong association with the 60-minute time trial (i.e. 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\(^5\). In
contrast, others have found no difference between the FTP\textsubscript{95} and the lactate threshold\textsuperscript{6}, nor between the FTP\textsubscript{95} and the individual anaerobic threshold\textsuperscript{7}. Moreover, a comparison between FTP\textsubscript{95} and CP (also closely related with the anaerobic threshold\textsuperscript{8}) found a strong correlation and no difference\textsuperscript{9} between the two variables. However, large limits of agreement were reported and it was concluded that CP and FTP\textsubscript{95} should not be considered equivalent nor used interchangeably\textsuperscript{9}.

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

Thus, the aims for this study were to assess whether FTP\textsubscript{95} PO is similar to that at MLSS and whether the FTP\textsubscript{20} is sensitive to changes in fitness status over a 7-month training period. Based on the fact that i) the FTP is a performance-based test that could be subject to external factors; ii) the FTP\textsubscript{95} is a fixed percentage that does not consider inter-individual variations; and iii) it has previously been shown to be an overestimation of FTP\textsubscript{60}, we hypothesized that FTP\textsubscript{95} PO would be different to PO at MLSS. Additionally, we hypothesized that PO at FTP\textsubscript{95} would be sensitive to changes in fitness level.

**Methods**

**Participants**

Eighteen participants (12 males – mean ± SD; 37 ± 6 years; 180 ± 6 cm; 79 ± 8 kg; 6 females – mean ± SD; 28 ± 6 years; 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. VO\textsubscript{2max}, self-reported training
volume and years of training experience were used to categorize participants with reference to previously established guidelines\textsuperscript{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, USA) in an environmentally controlled laboratory (i.e., temperature \(\sim 21^\circ\text{C}\), relative humidity \(\sim 36\%\)) over the span of 4-5 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 below for details).

The study was separated into two separate parts with identical testing procedures. The first part included all participants (\(n=18\)) whereas the second part included 10 returning participants (9 males, 1 female-mean \(\pm\) SD: 39 \(\pm\) 5 years; 178 \(\pm\) 8 cm; PRE 76 \(\pm\) 10 kg, POST 76 \(\pm\) 11 kg). For these ten participants the first and second parts corresponded to before (PRE) and to the end (POST) of a seven-month cycling season\textsuperscript{12}. Additionally, for these ten 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 ten participants trained on average 5-6 days per week, for \(\sim 1.5\)-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\textsuperscript{5}.

*Ramp-incremental test.* The initial visit consisted of a ramp-incremental test to exhaustion to determine maximal oxygen uptake (\(\dot{\text{V}}\text{O}_{2}\text{max}\)) and to predict the initial load for determination of MLSS\textsuperscript{13}. The ramp-incremental test began with a 4-min baseline at 50W followed by a 30W\(\cdot\text{min}^{-1}\) ramp for males and a 25W\(\cdot\text{min}^{-1}\) for females.
Constant-load trials. The successive visits after the ramp-incremental test included 30-min constant-PO trials for the determination of MLSS. Participants were instructed to cycle at their preferred cadence which 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·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 80W before PO was instantaneously increased to a predetermined value. Throughout all the testing sessions the participants were blinded to PO and elapsed time.

20-minute Functional Threshold Power Test. The Velotron 3D software (Racer Mate, Seattle, WA, USA) 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 80W. 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 PO but 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 as per manufacturer recommendations. Heart rate was continuously monitored (Garmin, Chicago, USA).

[BLa] was measured with a portable device (Lactate Scout, LS, SensLab, GmbH, 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 1s intervals.

Ramp-Incremental test. The highest \( \dot{V}O_2 \) computed from a 30 s rolling average was defined as \( \dot{V}O_2_{\text{max}} \), while peak PO was the highest PO value achieved at the end of the ramp-incremental test.

MLSS & FTP. \( \dot{V}O_2 \) at MLSS was determined from the average of the last 10 minutes of the constant-load trial. FTP\(_{20}\) PO was calculated as the average of the entire 20-minute test from which the FTP\(_{95}\%\) was derived (95% of the 20-minute average). PO was interpolated into 1 s intervals and to 5 min bins for statistical comparison (i.e. PRE vs POST).

HR, RPE, FS, FAS, [BLa]. HR was taken as the average of the last two minutes of exercise and the RPE collected during the final minute of exercise was used for comparison of PRE to POST. Pre-session FS and FAS measurements were used for analysis. End [BLa] represents the sample taken in the 30\(^{th}\) minute and 20\(^{th}\) minute for MLSS and FTP\(_{20}\), respectively.

Pacing. Changes in pacing strategy were evaluated PRE to POST by finding the average PO within 5-minute segments during the FTP\(_{20}\).

Statistics
All data processing and modelling were performed with a commercially available computer software (OriginLab, Northhampton, MA) and statistical analysis was performed using SPSS version 23 (SPSS, Chicago, USA) with statistical significance set at a $P < 0.05$. Descriptive data are presented as mean ± SD. Paired samples t-tests were used to evaluate differences in PO at FTP$_{20}$ and FTP$_{95\%}$ compared to 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 ANOVA was used to evaluate these difference PRE-POST. Bland-Altman analyses were used to test for agreement between PO at MLSS and FTP$_{95\%}$ while the association between values of $\dot{V}O_2$ and 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 $\dot{V}O_2_{\text{max}}$ was 4.00 ± 0.68 L/min with a peak PO of 394 ± 67W. PO at FTP$_{20}$ and FTP$_{95\%}$ was greater than that at MLSS ($p<0.05$; 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 (i.e., bias) between the PO observed at MLSS compared to FTP$_{95\%}$ (Figure 1, left panel). Mean HR and RPE was 162 ± 8 bpm and 5.0 ± 1.7 at MLSS and 175 ± 8 bpm and 8.4 ± 1.4 at FTP$_{20}$, respectively. Mean change in [BLa] for MLSS was 0.7 ± 0.3 mmol·L$^{-1}$. Mean end [BLa] was 4.3 ± 1.2 mmol·L$^{-1}$ for MLSS and 12.3 ± 2.6 mmol·L$^{-1}$ for FTP$_{20}$.

**PRE to POST responses**

For the 10 participants that completed both phases of the study, no increase in $\dot{V}O_2_{\text{max}}$ was seen from PRE (4.32 ± 0.53 L·min$^{-1}$, 56.6 ± 4.3 ml·kg·min$^{-1}$) to POST (4.37 ± 0.60 L·min$^{-1}$, 57.7 ± 7.9 ml·kg·min$^{-1}$) ($p=0.45$). Mean change in [BLa] for the MLSS trials ($\Delta$ [BLa] from the 10$^{\text{th}}$ to 30$^{\text{th}}$ minute) was 0.7 ± 0.3
mmol·L⁻¹ PRE and 0.7 ± 0.3 mmol·L⁻¹ POST. Table 2 displays mean HR, end [BLa], RPE, FS and FAS for MLSS and FTP20 at PRE and POST.

Table 3 displays PRE and POST values of MLSS, FTP20 and FTP95%. MLSS was greater at POST compared to PRE for both PO (+12 ± 8W; range +2 to 28W) (p=0.00) and \( \dot{V}O_2 \) (PRE 3.63 ± 0.51 L·min⁻¹, POST 3.77 ± 0.51 L·min⁻¹; +0.14 ± 0.13 L/min; range -0.01 to + 0.37 L·min⁻¹) (p=0.01). No increase was seen in mean PO at FTP20 (range -18 to +26W) and FTP95% (-17 to +25W) (p=0.75). Bland-Altman and a correlation analysis of changes in PO at MLSS and FTP95% from PRE to POST are shown in Figure 2. At PRE, FTP95% represented 88 ± 6% of PO at MLSS (range: -8 to +51W; bias = -20W) whereas at POST, FTP95% values represented 92 ± 5% of PO at MLSS (range: -31 to +28W; bias -9W). No difference in PO was found at any of the 5-minute segments from PRE-POST (p=0.48-0.96) (Figure 3).

Discussion

The main goal of this study was to evaluate the ability of the 20-minute FTP test to predict PO associated with MLSS. As hypothesized, despite a strong correlation between PO at FTP95% and MLSS, the calculated FTP95% overestimated PO corresponding to MLSS (i.e., bias = -17W) with large variability between the measures (i.e., differences ranging from -8 to +51W for FTP95%). A second goal of this study was to evaluate the ability of the FTP95% to reflect changes in fitness on a longitudinal basis. Contrary to our hypothesis, the results of this study indicate that the PO at FTP95% was not sensitive to changes in MLSS, as improvements in this marker were not reflected in the FTP95%.

Relevance of FTP20 testing from PRE training data: The FTP95% derived from the FTP20 test has recently become a widespread approach thought to be able to estimate PO associated to the critical intensity of exercise. This study compared the FTP95% PO derived from the FTP20 test to that at MLSS, which represents the upper limit for metabolic steady-state during continuous exercise. The results of this study demonstrate that a PO lower than the recommended 95% of the FTP20 was associated MLSS. While the results of this study indicate that 88.5% (± 4.8%) of the FTP20 is more likely to reflect PO at MLSS, the
large amount of variability in the agreement for these measures (LOA = 9 to -44W) prevents the use of this percent value with any confidence, as a superior approximation of MLSS. In this regard, MacInnis et al. previously reported that FTP95% exceeds PO for the FTP60 test and that the FTP60 PO represented 90% (CI 88-92%) of that achieved during a FTP20 test, which is in good agreement with the present study. Taken together, these data are in accordance with our hypothesis, and indicate not only that using a PO of 95% of the FTP20 seems to be an overestimation of the actual PO associated to MLSS, but also that even by using a lower percentage of the FTP20, there is large inter-individual variability inherent in this prediction. This may be partly related to the fact that both oxidative and non-oxidative energetic pathways contribute to the overall FTP20 performance but that their proportional contributions may vary between individuals. In this context, the discrepancy between these measures is concerning if trying to use FTP95% as a proxy for MLSS, as previous research has demonstrated that exercising at only 10 W above MLSS profoundly reduces subsequent performance ability (i.e. time to task failure) is substantially reduced.

Furthermore, inaccurate estimations of this intensity could change prescription of intended training intensity zones.

While the present study adopted MLSS as the criterion measure for the upper limit of metabolic stability and compared this with the FTP95%, other studies have investigated the correspondence between FTP tests and different markers of critical intensity. For example, Morgan et al. found a close relationship between CP (275 ± 42W) and FTP95% (278 ± 42W). However, similar to our results, they found that the corresponding limits of agreement (+10.9% to 13.1%) exceeded those that would allow the two measures to be used interchangeably with a high level of confidence. Additionally, in the previously mentioned study of MacInnis et al. the authors found that 95% of the FTP60 (CI 92-98%) was equal to CP. Although this may be in contrast to our results, it should be noted that the authors utilized a two-trial linear model (including a four-minute trial) which might have overestimated PO at CP. It is important to highlight that, although CP and MLSS share a similar definition, they reflect two different methods to derive PO at critical intensity. As briefly mentioned, estimates of CP are affected by the testing protocol, the
mathematical model used, and the data fitting strategy\textsuperscript{16,19} and in some circumstances have been shown to elicit POs greater than that at MLSS\textsuperscript{20}. Thus, caution is warranted when comparing POs at MLSS and CP in relation to that derived from FTP testing. Given the great variability in measures of FTP compared to other markers of critical intensity, caution should be exerted before using FTP\textsubscript{95} as one size fits all approach to predicting critical intensities of exercise\textsuperscript{7}.

\textit{Effects of training on FTP and MLSS:} This study found that PO at MLSS was greater at POST compared to PRE. Surprisingly, the increase in PO associated with MLSS did not translate into an improvement in the FTP\textsubscript{20} test, as evidenced by the fact that the improvements in MLSS from PRE to POST correlated poorly to changes (or the lack thereof) for FTP\textsubscript{20} from PRE to POST, and did not translate to improvements in FTP\textsubscript{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 valid and reliable\textsuperscript{21,22}, the present data question the ability of the FTP\textsubscript{20} test to accurately track those changes. It could be argued that the average increase of MLSS was relatively small (i.e., 12W) however it is likely that well-trained populations have a smaller opportunity for improvement and that changes are of a smaller magnitude compared to untrained populations\textsuperscript{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\textsubscript{20}, as in this population greater relative improvements in aerobic fitness can be expected\textsuperscript{25}. Therefore, the results of the present study indicate that the FTP\textsubscript{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 which may not encompass all components of performance, which however seems unlikely given the tight association between MLSS and exercise capacity\textsuperscript{2}.
Although the reasons why the increase in PO at MLSS did not translate into an improvement in the FTP\textsubscript{20} test cannot be fully elucidated from this study, it should be acknowledged that the FTP\textsubscript{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\textsuperscript{26}. While 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 (e.g., motivation) cannot be controlled. Even though there were no differences 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 towards 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 to a laboratory-based test (i.e., MLSS).

Additionally, it has been showed that in some circumstances experience and training status also can influence the reliability of a time-trial test\textsuperscript{27}; however, it is unlikely that this played a role in our study as the cyclists involved in the post-measurements were the most familiar with the FTP\textsubscript{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\textsuperscript{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\textsuperscript{29}. Regardless, we did not find differences in the pacing strategy employed by the participants between PRE and POST measurements of the FTP\textsubscript{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\textsubscript{95\%} have been shown in our results and those presented by others\textsuperscript{5}, there may be the need to develop an alternative approach to the FTP\textsubscript{95\%} that is reliable, valid and convenient for cyclists. Based
on these data, it could be suggested that the use of the FTP95% 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 PO at MLSS.

**Conclusions**

The results from this study indicate that the FTP95% does not provide an accurate representation of PO at MLSS. Even with an adjusted percentage (i.e., 88% rather than 95% of FTP20 representing a value for FTP60), the large variability in the data is such that it would not be advisable to use the FTP95% test to estimate MLSS. Furthermore, the results demonstrated that POs from the FTP95% 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.
Disclosure of interest

The authors report no conflict of interest.

Acknowledgements

We would like to express our gratitude to the participants in this study.

Erin Calaine Inglis 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. Juan M. Murias 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.
References


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

**Figure 2.** Bland-Altman plot analysis (*left*) showing changes in power output PRE to POST at maximal lactate steady state (MLSS) and 95% of the 20-minute functional threshold power test (FTP_{95%}). Correlation graph (*right*) between change in power output at MLSS and FTP_{95%} (dashed grey line indicates line of identity).

**Figure 3.** Power output during the 20-minute functional threshold power test (FTP_{20}) at PRE (grey circles, positive SD) and POST (white circles, negative SD).