INTERNATIONAL JOURNAL OF
-EXERCISE SCIENCE-

Original Research

Associations of Laboratory- and Field- derived Measurements of Critical Power with W'-kinetics during 40-km Cycling Time Trial Performances

D. Luke Wilkins*†[1,2], Julie E. Taylor‡[3], Robert W. Pettitt‡[4], Mark Kramer‡[5]

[1]Campus Recreation/Outdoor Recreation, Utah Tech University, St. George, Utah, USA; [2]Rocky Mountain University of Health Professions (RMUoHP), Salt Lake City, Utah, USA; [3]Kinesiology and Outdoor Recreation Department, Southern Utah University, Utah, USA; [4]School of Health Sciences, Salt Lake Community College (SLCC), Salt Lake City, Utah, USA; [5]Physical Activity, Sport, and Recreation (PhASRec) Research Focus Area, North-West University, Potchefstroom, South Africa

*Denotes student investigator, ‡Denotes established investigator

Abstract

International Journal of Exercise Science 18(8): 757-773, 2025. https://doi.org/10.70252/NEFW5464 This study aimed to establish whether a laboratory (lab) based 3-minute all-out (3MT) protocol and a field-based 3MT protocol would yield similar peak power (P_{max}), critical power (CP), and curvature constant (W') profiles and the implications of parameter estimation for informing a 40-km time trial (TT) performance. Nine competitive male cyclists (mean ± SD: age 36.5 ± 10.42 y, mass = 80.5 ± 10.6 kg, height 1.8 ± 0.1 m) completed two 3MTs on separate days, as well as a 40-km time trial. Both lab and field-based protocols evoked similar CP ($p = 0.160$) and W' ($p = 0.200$) profiles, but P_{max} ($p = 0.012$) may be more sensitive to biomechanical disparities and testing environment. Strong positive associations were observed with W'-kinetics ($r = 0.73$) and W' ($r = 0.83$) and moderate-to-strong negative associations with mean TT power ($r = -0.75$) and CP ($r = -0.68$). TT power outputs occur at 59-65% of CP, and finishing times appear to be informed by CP, W' and P_{max} with high degrees of accuracy (R2 > 0.90). Although TT performances occur predominantly within the moderate-to-heavy intensity domains, the mean intensity from a cardiovascular and core temperature perspective was high (i.e., ~90% HR_{max}; ~39°C). TT performances appear to be accurately informed by CP, W' and P_{max}, with W' dominating the predictive capacity associated with longer TT performances.

Keywords: All-out testing, curvature constant, intensity domains, interchangeability

Introduction

Cycling time trial (TT) races are unique events where cyclists compete alone and against the clock while racing fixed distances as quickly as possible.[1-2] During TT events, cyclists can self-select their pace, intensity, and power output to yield the best possible TT performance which is also dependent on the regulation of energetic output, psychological drive, and tactical patterns influenced by the terrain.[1-3] Cyclists in TT stages were shown to sustain intensities ranging from 75% to 90% of their maximal oxygen uptake ($\dot{V}O_{2max}$), 80 ± 5% of heart rate max (HR_{max}), and speeds ranging from 32.5 to 46.7 km·h^{-1} (20-30 mph) during long-duration time trials (e.g. 40-km),[1,3–5] thereby highlighting the importance of effectively training the bioenergetic pathways such that they are capable of sustaining high intensities for extended durations.

Cycling TT durations typically range from 10 to 60 min and are predominantly aerobic events,[3] which lends itself well to the critical power (CP) concept. The CP concept is based on the hyperbolic relationship between work rate and time to exhaustion where CP is representative of the asymptote.[6] This curvilinear relationship demonstrates that the intensity of exercise above CP (measured in Watts) can only be maintained for a finite duration which is dictated by the magnitude of W-prime (curvature constant: W'; measured in J). Physiologically, CP embodies a given individual intensity that demarcates the boundary between the heavy and severe intensity domains whereas W' is indicative of a depletable energy reserve when exercise intensities exceed CP.[6] Although W' is typically thought to be derived from 'anaerobic' processes, it is more prudent to conceptualise W' as a mechanical work capacity linked to the magnitude between CP and $\dot{V}O_{2max}$ that serves an energy buffer while phosphocreatine concentrations ([PCr]) and pH project towards a nadir at intensities sustained above CP.[7] This heavy-severe boundary is important because it divorces intensities where a steady state is possible from those that are not and is closely associated with TT performance.[8]

Therefore, depending on the distances of a given TT, performances are typically associated with peak power output, $\dot{V}O_{2max}$, lactate threshold, and work economy,[9,10] although the extent to which this is true for longer TT distances (e.g., 40-km) is under-researched.[4] Given the non-linear relationship of the power-duration relationship, sustainable intensities for longer TT durations are more likely to transpire within the heavy intensity domain whereby glycogen depletion, central drive, and hyperthermia are the likely fatigue mechanisms.[4,5,9] Moreover, TT performances are stochastic in nature, implying an undulating transition between intensity domains with intermittent surges in power to elicit more favorable performance outcomes.[2] The extent to which CP and W' would provide serviceable information related to longer TT pacing and performance has not been previously investigated. It is plausible that W' may be diminished to varying extents depending on the proximity of TT power (TTP) to CP (i.e., when TTP is above CP), but it is unclear whether W' is of measurable importance in relation to 40-km TT. Furthermore, performances are likely to depend on the repeated depletion and reconstitution kinetics of W' throughout a TT[10] given that a 40-km TT is achievable within a ~60-minute timeframe, and is thus at the upper limit of the CP paradigm. The application of the CP framework to these upper bounds has not been directly evaluated thereby providing an avenue for gleaning novel insights from both a theoretical and practical perspective.

A better understanding of TT pacing in the context of the CP concept is dependent on derivation of CP/W' (i.e., laboratory vs. field).[6,10] Although both CP and W' are typically derived from a series of laboratory-based (lab-based) constant work rate bouts (i.e., 3-5) lasting between 2-15 minutes of exhaustion time,[8,11] a more time efficient, yet similarly valid method is the 3-min all-out test (3MT).[12] In either instance, accurate modeling is primarily dependent on the level of effort during each trial (i.e., maximal), and the extent of W′ depletion (i.e., W′ should be wholly depleted).[10] Previous research has shown that CP can be reliably obtained from field-based testing, but that W′ variability would likely preclude interchangeability between lab and field-based metrics.[13] Whether CP and W′ derived from either a lab- or field-based 3MT would provide convergent information related to longer TT performances has not been previously investigated. Finally, since a longer TT may be limited by distinctive physiological responses (e.g., changes in core temperature, heart rate)[3,14] it is posited that higher core temperatures may be a limiting factor given that changes in intensity (e.g., 70% $\dot{V}O_{2max}$) and ambient temperature are known to effect metabolic rate, heart rate, elevated thermal sensation, reduced thermal comfort, mental fatigue, and reduced gross cycling efficiency.[15–17] As such, the extent to which core temperatures tend to change throughout the time course of a TT has not been previously researched and requires further investigation.

Given the gaps identified in the literature, the objectives of the present study were 5-fold, namely to: (i) determine whether the parameters derived from a lab and field-based 3MT were interchangeable, (ii) define TT performances in relation to CP and W′, (iii) to model the W′-balance kinetics during a 40-km TT, (iv) evaluate the changes in HR and core temperature throughout a 40-km TT, and (v) determine which 3MT parameters, if any, are most predictive of TT performances. We hypthesised that (i) there would be significant differences between laboratory and field-derived CP and W′ parameters which may limit their interchangeability, (ii) cycling TT performances would occur near but below CP that would indicate intensities predominantly within the heavy intensity domain, (iii) W′ would demonstrate periodic depletion and partial recovery rates that would be reflective of the stochastic nature of the TT, (iv) both HR and core temperatures would elevate considerably throughout the TT, with core temperature reaching values that would be associated with potential performance impairment, and (v) CP or the relative proportion of CP sustained would the best predictor of overall TT performance.

Methods

Participants

Following approval by the university's Institutional Review Board, 10 male participants were recruited from local cycling/triathlon clubs. The expectation was that the laboratory and field tests would be predictive of each other, therefore a power analysis (G*Power, 3.1.9.2), using a correlation model determined that the total participants (n) needed to be 11 based on the following inputs: (i) alpha error probability of 0.05, (ii) power of 0.80, and (iii) a correlation H_1 of 0.70 (H_0 of 0.0)[6]. One participant was eliminated from the final analysis due to incomplete data for the TT, resulting in a final sample of nine participants.

These volunteer athletes were considered competitive cyclists/triathletes, however, none of them were considered professionals (mean ± SD: age 36.50 ± 10.42yr, body mass 80.5 ± 10.32 kg, height 1.82 ± 0.05 m). All participants completed a physical activity screening questionnaire, gave voluntary verbal and written approved informed consent, and had previous high-intensity exercise experience. Moreover, this research was carried out fully in accordance with the ethical standards of the *International Journal of Exercise Science*.[18]

To be considered for inclusion to the present study, participants needed to be: (i) ≥ 18 yr of age, (ii) classified as a competitive cyclist with ≥ 1-yr competitive race experience (CAT 4 or better, USA cycling race categories), and (iii) currently training as a competitive cyclist. If a participant had a previous diagnosis of a heart-related illness/disease they were excluded from this study due to the high-intensity nature of the 3MT protocol.

Protocol

Testing was conducted in a randomized, counter-balanced, and cross-over design with no more than seven days between trials. Trials were conducted around the same time of day to reduce bias associated with diurnal variations.

Participants were asked to keep a food and fluid intake journal for 24 hr prior to the first testing session and to replicate it for each subsequent testing session. Participants were asked to refrain from strenuous physical activity, irregular caffeine intake, alcohol, nutritional supplements, and anti-inflammatory drugs for 48 hr before all testing sessions. Furthermore, participants were asked whether they adhered to these standards before each testing session.

Pre-Trial and Three-Min All-Out Exercise Testing Procedures.

Participants were required to visit the testing facility on three separate occasions to obtain the relevant anthropometric and physiological data. During their first visit, participant's age, height, and body mass were obtained, and the informed consent document was completed. Height and mass were measured using an eye-level physician's beam scale (439, Cardinal Detecto, Ellicott, Maryland). The second visit entailed the completion of a laboratory-based 3MT with room temperature of ~24°C). The test was conducted on a Wahoo Kickr (power measurement accurate up to ±3%) and indoor bike trainer (Wahoo Fitness, Atlanta, GA). The third visit, separated by a minimum of 48 hrs for adequate between-session recovery, required participants to complete a 3MT on an open road course (temp ~24°C; humidity ~ 52%). Power data were collected via pedal-based power meters (Garmin Rally XC100; ± 1% accuracy, 2.4 GHz; Garmin, Olathe, KS). All tests were performed around the same time of day (i.e., ± 2 hr).

Participants were instructed to bring their own bicycle and kit including riding apparel, helmet, and shoes to all testing sessions. Participants were also asked to adhere to the following prior to testing: (i) refrain from vigorous exercise 24-48 hr prior, (ii) avoid caffeine intake 4 hr prior, (iii) avoid alcohol consumption 24 hr prior, (iv) arrive in a well hydrated state and be ~2 hr post-prandial. Clark et al[19] reported that 3MTs performed on smart cycling trainers evoke mode-specific estimations for W_{peak}, CP and W′.

In a counterbalanced random order, two separate 3MTs were performed on two different days. The tests were preceded with a 15-minute warm up consisting of low resistance, high cadence efforts interspersed with low intensity (i.e, RPE_{10} = 2) cycling. This was followed by 5-minutes of dynamic stretches of the major lower extremity muscle groups whereafter the relevant testing equipment was fitted to the participant. Participants were instructed to begin the test with an all-out effort and to continue pedaling as hard as possible throughout the test and were encouraged to sustain the highest power output they could during the entire 3-min duration. However, they were allowed to change gears as needed and were allowed to stand in order to sustain an all-out effort. Critical power (Equation 1) and W′ (Equation 2) were calculated as follows:

CP = average power output for the last 30 sec of the test[20] [Equation 1]

$W' = 150 \times (P_{150s} - CP)$[21,22] [Equation 2]

P_{150s} is the average power output for the first 150 seconds of the test. CP and P150s are expressed in watts (W) and W′ is expressed in joules (J). For the outdoor trials, all weather data were recorded for consistency (wind: 3 $m \cdot s^{-1}$; heat: 30°C; relative humidity: 17%). All data were collected on the V̇O2 Master app (V̇O2 Master Health Sensors Inc., Vernon, BC, Canada) to obtain synchronised power and heart rate data throughout the entirety of the test.

40-km Time Trial.

The 40-km TT took place on a section of road approximately 8.5 km in length and only contained 165 m (541ft.; mean slope 3.4%, max slope 22.3%) of elevation gain. More specifically, the TT was segmented into 5 laps (intervals) of 8km each to more carefully control potentially confounding effects of gradation, wind, and rolling resistance. The average summer temperature was 36.25° C and an average humidity of 20% made for a stable environment to conduct this study. Temperature, humidity, and wind speed were measured (Kestrel 3000HS, Boothwyn, PA) on experimental days and checked for statistical differences.

Eight to 10 hr before testing, participants were asked to ingest a telemetric core body temperature measuring pill (CorTemp, Palmetto, FL), which could be remotely tracked. Upon arrival at the designated testing area, participants were fitted with a HR monitor (Polar, Accurex Plus, Finland), and had their bikes fitted with a Garmin Rally XC100 pedal-based power meter (± 1% accuracy, 2.4 GHz; Garmin, Olathe, KS). Power meters were calibrated according to the manufacturer's recommendations before all trials.

Following the protocol used by Takeshima et al,[23] participants were allowed a 15-min self-selected warm-up before the intervention. Participants were asked to record their warm-up and repeat it for the second experimental trial. Mimicking the first trial's warm-up protocol minimized differences in core body temperature between trials.[24] Ambient temperature water (20 oz. bottles) was available at an aid station every 8 km, and ad libitum consumption was encouraged. All TTs were completed individually and at the same approximate time of day (~1 hr). Participants were recruited from local affiliates and were therefore already familiar with the section of road, ambient temperatures, and elevation.

Data Collection.

All physiological variables were collected continuously, on the VO2 Master app (VO_2 Master Health Sensors Inc., Vernon, BC, Canada) throughout the TT, with performance time collected every 8 km lap. Lap times and overall finishing times were collected with an iPad mini 2 (native stopwatch app) (Apple Inc., Cupertino, CA). Heart rate was captured telemetrically (Polar H9 HR sensor, Polar, Accurex Plus, Finland). Core body temperature (TC) was captured via an ingestible telemetric pill (CorTemp, Palmetto, FL). Power was captured by a set of pedal-based power meters (Garmin Rally XC100 (± 1% accuracy, 2.4 GHz; Garmin, Olathe, KS).

W'-Balance kinetics.

We utilized the W'_{bal}-int model[25] to show the amount of W' remaining as a function of time during the 40-km TT. More specifically we used the discrete form of the equations expressed by Skiba et al. (Equations 3-4) such that the integrals are expressed as sums for working with digital data:[25]

$$W'_{bal-int,j} = W'_0 - \sum_{i=1}^{j}\left[e^{\frac{-(j-i)}{\tau_{W'}}}\right] W'_{exp,i} \cdot \Delta u_i \quad \text{[Equation 3]}$$

where, W'_0 indicates the initial W' at the initiation of the TT, 'e' is Euler's constant that is approximately equal to 2.17, $\tau_{W'}$ is the time constant of W' reconstitution, Δu_i is a time segment for measuring changes in power output (typically 1-second), 'i' is the i^{th} segment of the total time subdivided into n segments, j is the segment for which $W'_{bal-int}$ is calculated, and $W'_{exp,i}$ refers to the linear depletion of W' when the power output exceeds CP (see Equation 4):

$$W'_{exp,i} = \left\{\frac{0, P_i \leq CP}{(P_i - CP)\cdot \Delta u_i, P_i > CP}\right\} \quad \text{[Equation 4]}$$

From this analysis the W'_{bal} remaining at the end of the TT as well as the maximal change in W' ($\Delta W' = W'_{max} - W'_{min}$) were retained for analysis.

Statistical Analysis

Unless otherwise stated, all data are reported as means and standard deviations (SD). Data were evaluated for normality using the Shapiro-Wilk test, whereby deviations from normality were accepted at $p < 0.05$. Due to the small sample size and to minimise the type-1 error rate, the Wilcoxon signed rank test was used to compare 3MT parameters from the laboratory (criterion method) to that of the field test (reference method). A Bland-Altman analysis was used to compare the bias and limits of agreement (LoA) of the reference method to the criterion method for each parameter. A correlation analysis was conducted to evaluate the association between 3MT derived parameters and TT performance. Correlations were evaluated using the Spearman Rank correlation coefficient due to the smaller sample size which was then quantitatively interpreted in absolute terms as follows: r: 0.00-0.10 = negligible; 0.10-0.39 = weak; 0.40-0.69 = moderate; 0.70-0.89 = strong; 0.90-1.00 = very strong.[26] To minimise the type-1 error rates associated with the correlation coefficients, the Holm correction was implemented. To determine which 3MT parameters were most associated with TT finishing times three separate multiple regression models were evaluated. The first model incorporated 3MT parameters such as CP, W′ and P_{max}; the second model utilized additional parameters such as ΔW′, ΔCT; the third model used stepwise regression to retain only those parameters that significantly contributed to the model outcome. Model effectiveness was evaluated using the coefficient of determination (R^2) as well as the adjusted-R^2 to compensate for model complexity (i.e., accommodate for additional parameters). Statistical significance was set at $p < .05$. All statistical analyses were completed using R (RStudio [version 22.12.0 Build 353]: Integrated Development for R. RStudio, PBC, Boston, MA URL: http://www.rstudio.com).[27,28]

Results

The descriptive statistics for 3MT derived parameters as well as TT performance parameters are highlighted in Table 1.

Table 1. Descriptive Statistics.

Variable	Condition	Mean	Std. Deviation	Shapiro-Wilk	P-value of Shapiro-Wilk
CP (W)	Lab	312.292	54.483	0.895	0.222
CP (W)	Field	280.941	65.457	0.909	0.306
W′ (J)	Lab	12730.802	8639.912	0.788	0.015
W′ (J)	Field	14801.560	5171.920	0.884	0.172
P_{max} (W)	Lab	776.267	168.795	0.960	0.801
P_{max} (W)	Field	880.429	196.076	0.973	0.918
TTP_{avg} (W)	All	183.119	44.777	0.950	0.689
TTP_{avg} slope (W/s)	All	-0.010	0.01	0.619	<0.001
$TT_{P/CP}$ (%)	Lab	58.896	12.576	0.910	0.315
$TT_{P/CP}$ (%)	Field	65.493	7.614	0.917	0.368
$Time_{>CP}$ (s)	Lab	151.222	192.164	0.796	0.018
$Time_{>CP}$ (s)	Field	372.556	396.112	0.882	0.166
W′-Bal_{end} (J)	Lab	11967.201	9055.374	0.867	0.113
W′-Bal_{end} (J)	Field	14089.889	5240.546	0.892	0.207
ΔW′ (J)	Lab	1016.721	1407.021	0.707	0.002
ΔW′ (J)	Field	4881.090	8855.802	0.628	<0.001
T_{finish} (s)	All	4669.222	393.672	0.820	0.035

$TT_{P/CP}$ (time trial power relative to CP); TTP_{avg} (average power output during time trial); P_{max} (max power achieved during 3MT); CP (critical power); $Time_{>CP}$ (time spent above CP); W′ (W-prime); W′-Bal_{end} (W′ available at end of TT from the W′-bal model); ΔW′ (change in W′ from maximum to minimum throughout the TT); T_{finish} (finishing time for the TT); TTP_{avg} slope = mean change in power output across the TT duration

The paired individual comparisons for 3MT parameters for each testing condition (i.e., lab vs. field-based) are shown in Figure 1 (panels A 1-3). The results of the Bland-Altman analysis showing the bias and LoA between criterion method (laboratory) and reference method (road) are also shown in Figure 1 (panels B 1-3).

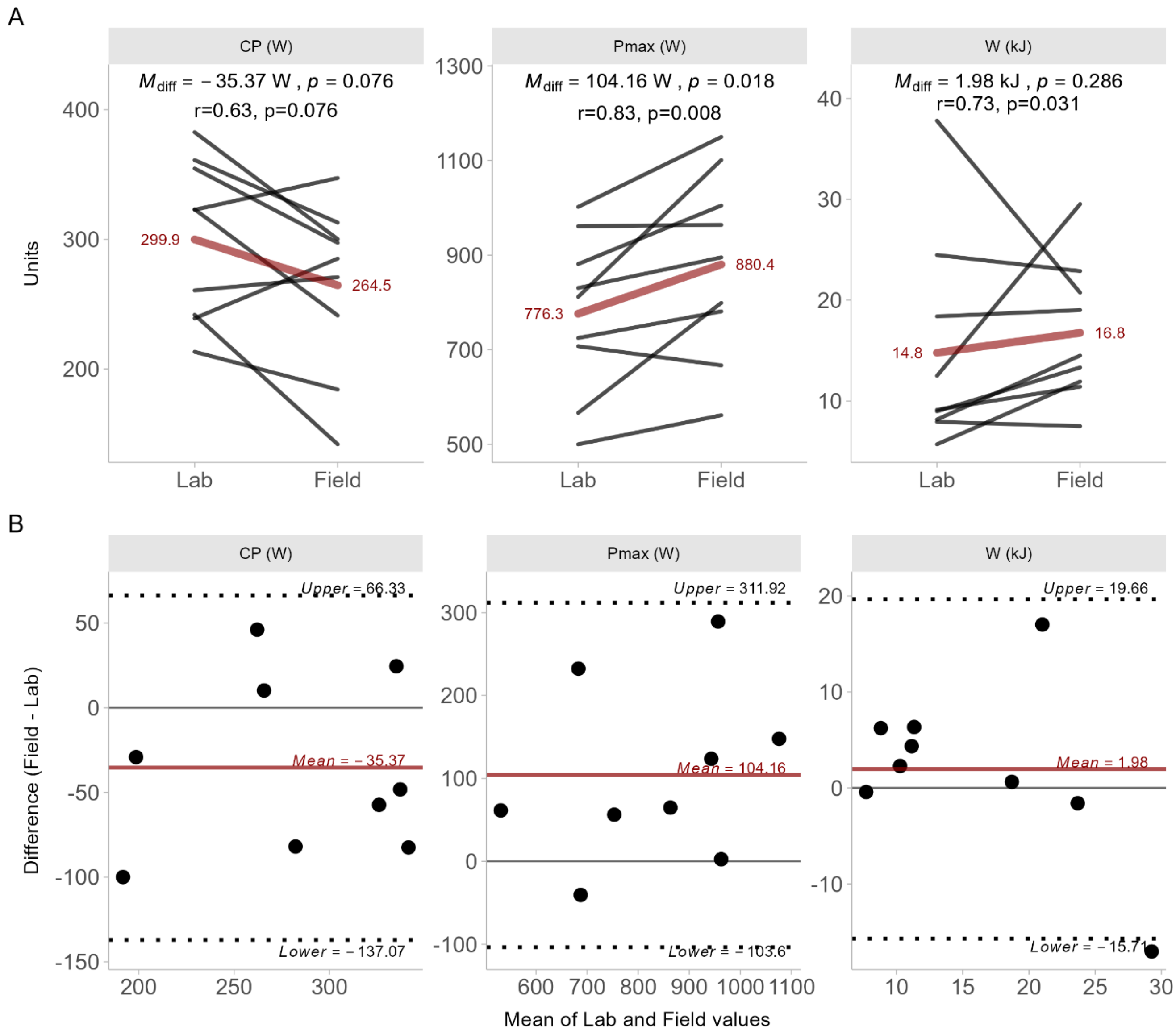

Figure 1. Differences between laboratory- and field-derived parameters for the 3MT. Panels A-C show paired comparisons for each participant. Panels D-F show the bias in the estimate between Road and Lab parameters (mean differences ± 95% CI).

The correlations between 3MT parameters from both testing conditions (laboratory vs. field) and TT performance parameters are shown in Figure 2 (panels A and B respectively). The magnitude and direction of the correlations are represented by the given opacity and color for easier interpretation.

The instantaneous power and associated W'_{bal} for a representative participant is highlighted in Figure 3. Whenever power output exceeds CP, the W′ available to a participant would be diminished, whereas W′ recovery would be initiated as soon as power output drops below CP.

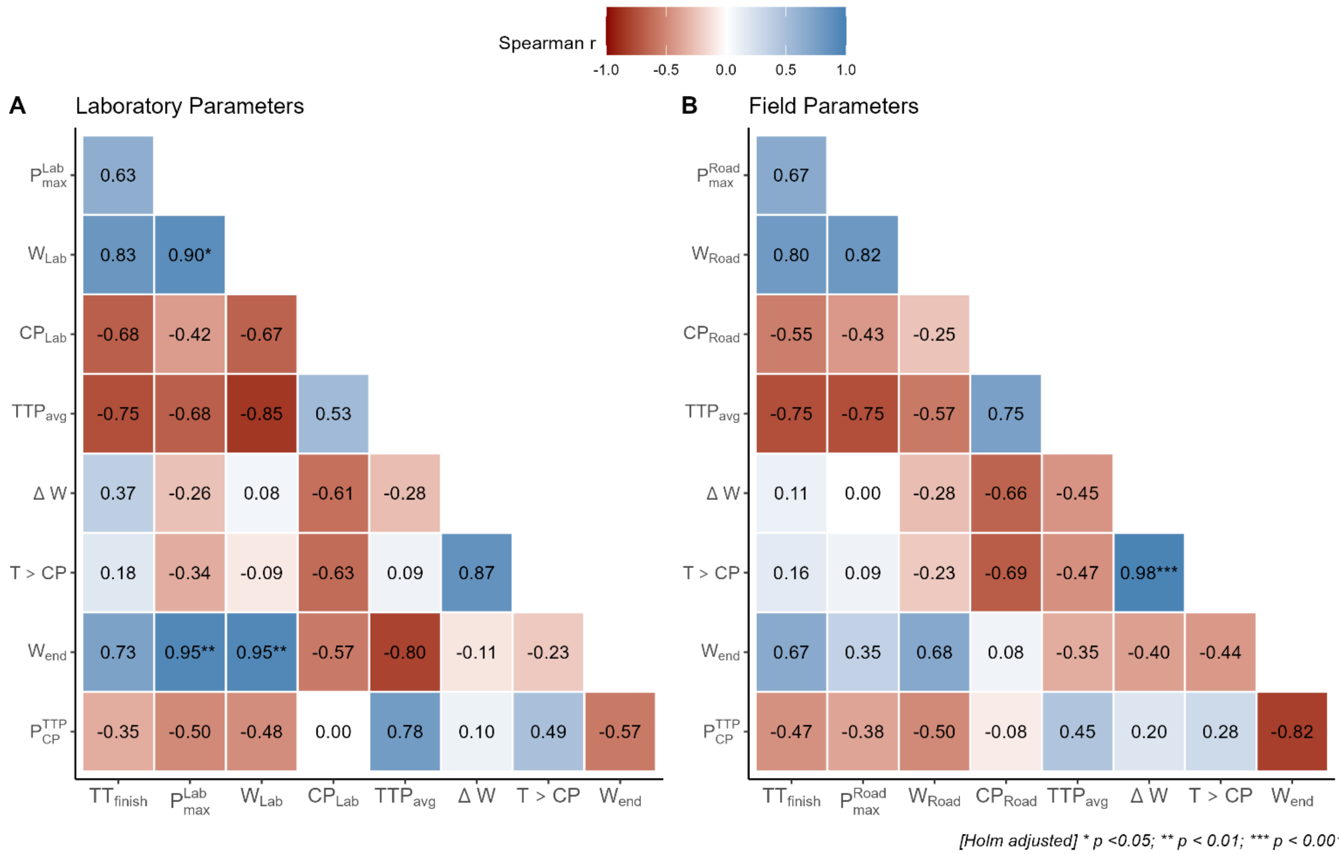

Figure 2. Correlation plots for 3MT-derived parameters and TT performance. Panel A: laboratory 3MT-derived parameters; panel B: field 3MT-derived parameters. Note: TTP/CP = TT power relative to CP; CP = critical power; W′ = W-prime; W′end = W′ at the end of the TT based on the W′-bal model; T>CP = time cycling at power outputs above CP; P_{max} = maximal power attained during the 3MTTable 1; TimeTT = finishing time of the TT.

Both HR and CT were continuously evaluated throughout the TT for all participants. The mean ± SD are highlighted in Figure 4. HR changed rapidly towards ~90% HR_{max} and remained elevated for the duration of the TT, whereas CT increased meaningfully at approximately 35% of the full TT. Intriguingly, some participants were precariously close to the hyperthermic threshold (40.5 oC) whereas others managed to stay comfortably below.

The results of the multiple regression models using field-based metrics are shown in Table 2. The field-based metrics were chosen given that these are more likely to encompass environmental conditions that would more closely mimic the TT conditions. Intriguingly, the best model, as judged by adj-R^2, appears to be solely reliant on W′ (adj-R^2 = 0.93) and shows that those with higher W′ magnitudes could be expected to have slower finishing times (i.e., for each 1J increase in W′ there is an expected 1-sec increase in finishing time). The second-best performing model (adj-R2 = 0.92) incorporated CP, W′ and P_{max}, and showed that those with higher CP and P_{max} values would be expected to have faster finishing times (i.e., for each 1W increase in CP, $Time_{TT}$ would decrease by 0.66-sec; each 1W increase in P_{max}, $Time_{TT}$ would decrease by 0.24-sec).

Table 2. Multiple regression models for TT finishing time.

	Model 1			Model 2			Model 3		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	4487.1	3496.38 - 5477.83	**<0.001**	3892.39	1426.26 - 6358.53	**0.015**	4107.383	3954.29 - 4260.48	**<0.001**
CP	-0.657	-3.82 - 2.5	0.616	0.643	-5.59 - 6.88	0.764			
W′	0.044	0.01 - 0.08	**0.013**	0.051	0 - 0.1	**0.043**	0.044	0.03 - 0.05	**<0.001**
P_{max}	-0.235	-1.4 - 0.93	0.625	-0.278	-2.02 - 1.47	0.647			
ΔW				0.040	-0.12 - 0.2	0.473			
ΔCT				66.201	-199.13 - 331.53	0.485			
Observations	N=9			N=9			N=9		
R^2 / R^2 adj	0.97 / 0.92			0.98 / 0.90			0.97 / 0.93		

Note: CP (critical power); W′ (W-prime); P_{max} (maximal power from the field-based 3MT); ΔW (absolute change in W′ (i.e., depletion) throughout the TT); ΔCT (absolute change in core temp throughout the TT); CI (confidence interval)

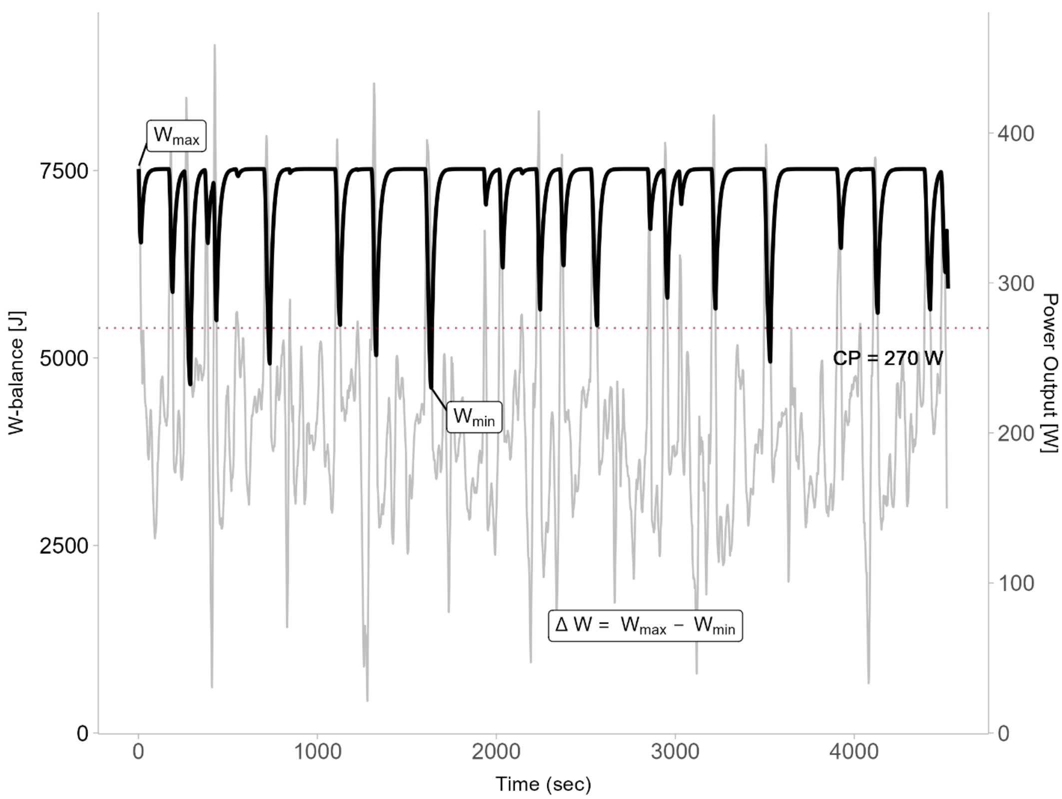

Figure 3. Power output and W′ expenditure during the TT for a representative athlete. W′-balance is shown in black whereby W′ depletion occurs at power outputs above CP and W′ recovery occurs at power outputs below CP (primary y-axis). Light grey line shows the instantaneous power output during the TT (secondary y-axis).

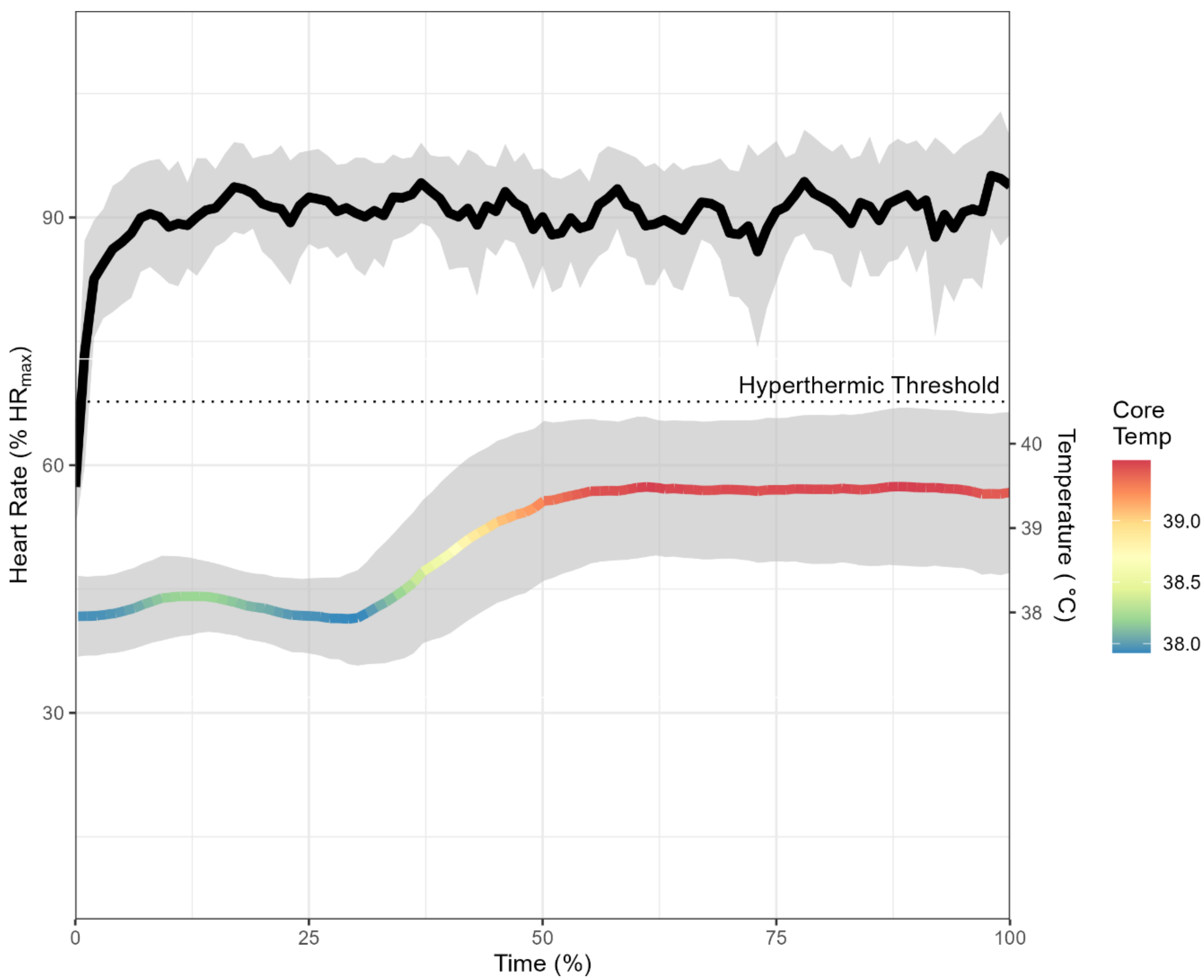

Figure 4. Mean heart rate (HR) and core temperature (CT) as a function of the normalized TT time. Top line (primary y-axis) shows the mean heart rate ± SD; solid horizontal grey line (top) shows the mean HR expressed as a percentage of HR_{max}. Scaled color line (secondary y-axis; bottom) shows the mean ± SD of CT. The coloring is scaled to show how CT changed throughout the TT. Dotted horizontal grey line (bottom) shows the hyperthermic threshold (40.5°C).

Discussion

While laboratory-based testing offers controlled environments and precise measurements, field-based testing provides a more comprehensive and applicable understanding of a cyclist's physiological parameters in real-world conditions. For TT performances, where environmental and biomechanical factors, psychological stressors, and dynamic physical demands play critical roles, field-based testing delivers insights that are essential for optimizing training, enhancing performance, and achieving competitive success.

We proposed that a field 3MT would produce similar peak power, CP, and W′ profiles as a lab 3MT. The results of the present study partially supported this in that both lab and field 3MT's produced CP ($p = 0.160$) and W′ ($p = 0.200$) profiles that were not significantly different, however, it is likely that P_{max} ($p = 0.012$) may be more sensitive to biomechanical discrepancies and the testing environment. The interpretation of the latter finding may be justified by biomechanical differences such as standing out of the saddle, braking/friction components, and potential differences in mechanical efficiency induced by these differences.[29,30] It is important to highlight the bias and LoA in the parameters estimates whereby the field-derived CP was 31.4 W lower, W′ was ~2kJ larger, and P_{max} was also 104 W larger compared to laboratory conditions (see Table 1).

Although both CP and W′ were not statistically different between testing conditions, it is likely that the bias in the estimates are large enough to preclude interchangeability. Although methodologically different, previous research has shown similar performance outcomes between laboratory and field testing in that CP would be statistically similar, whereas W′ would be statistically different.[13,31] Comparably, Bertucci et al[30] reported 6% higher peak power outputs during field than stationary ergometer sprinting (~5s sprints), although these authors ascribed decreased lateral oscillations of the bike in the laboratory as a possible differentiating mechanism. Moreover, for the same power output differences in gross efficiency and cycling economy have also been reported to be ~12% and 11% higher in field compared to lab conditions which have been attributed to differences in aerodynamic positioning and crank inertial loads.[30] Plausible differences in performances for the present study could also be credited to the gradation (M=3.8%, max 21.7%) of the field 3MT compared to the level platform of the laboratory 3MT.

The overarching utility of the 3MT is to derive credible approximations of CP, W′ and P_{max} such that sustainable intensities can be identified and utilized to inform both competition and training practices. In this line, a 40-km TT was completed and the relative functionality of each 3MT parameter was evaluated (see Figure 2). Based on the lab-based parameters, TT finishing times ($Time_{TT}$) were strongly predicted by W'_{end} ($r = 0.73$) and W′ ($r = 0.83$) implying that those with higher W′-related capacities would likely experience slower finishing times. Traditionally W′ is interpreted as a being reflective of predominantly anaerobic energy sources which would contextualize the findings of the present study. Similarly, $Time_{TT}$ showed moderate-to-strong negative associations with TTP_{avg} ($r = -0.75$) and CP ($r = -0.68$) implying that those who could sustain a higher mean power and exhibited a higher CP would likely yield faster finishing times. It is well understood that CP is indicative of aerobic metabolic processes whereby it represents the power output associated with the maximum metabolic steady state.[32–34] Indeed, a recent investigation has shown that CP derived from a laboratory-based 3MT can predict shorter TT performances (16.1-km).[35] A similar, yet slightly different interpretation emerged from the field-derived 3MT parameters in that $Time_{TT}$ was still positively associated with W'_{end} ($r = 0.67$) and W′ ($r = 0.80$), but now also included P_{max} ($r = 0.67$). Strong negative associations were observed only for TTP_{avg} ($r = -0.75$). Previous investigations have shown that P_{max} (derived from graded exercise testing [GXT]) exhibited very strong negative associations ($r = -0.91$) with short TT performances (20-km),[36] but the extent to which GXT-derived P_{max} and 3MT-derived P_{max} are associated has not been implicitly investigated. Nonetheless, the associations between P_{max} (GXT) and $Time_{TT}$ are rational on the basis that there are also very strong correlations between P_{max} and maximal oxygen uptake (VO_{2max}) ($r = 0.97$), showing that a high aerobic ceiling is required to generate and sustain higher power outputs.[36,37]

Research on actual TT performances that are informed by 3MT-derived parameters are severely limited, especially those of longer durations.[35] More specifically, the mean power output in relation to CP, the potential role of W′, and the physiological responses (e.g., core temperature and heart rate) during longer TT had not been previously explored. The present study showed that, on average, cyclists completed the TT at ~65% $TT_{P/CP}$ (road) or ~59% $TT_{P/CP}$ (lab) which more closely approximates power outputs associated the moderate-to-heavy intensity boundary.[38] An intriguing finding was the analytical capacity of W′ for 'predicting' $Time_{TT}$ (adj R^2 = 0.93, SE = 104.69 sec), which was stronger than models incorporating CP, and P_{max}. Such a result potentially indicates that the capacity for speed surges, which effects TTP_{avg}, may be a deciding factor even in longer format TT performances. The extent to which such a finding would truly be predictive of performance in a larger, more heterogenous sample, would however require verification. In part, Black et a[35] observed very strong associations between 16.1-km TT performances and CP (r=-0.83) as well as total work done (3MT, r = -0.86), although they did not explicitly investigate the potential role of W′. The fact that total work done yielded stronger associations with TT performances compared to CP is again an intriguing finding and should be explored to a greater extent in future research since it is unlikely that the maximization of W′ should be the goal; there is likely an ideal balance in the optimal W′.

Despite TT power outputs being well below CP in the present study, and therefore sustainable with limited metabolic perturbations, the intensity from a cardiovascular perspective was high (i.e., ~90% HR_{max}).[39] Changes in CT also showed that some individuals were precipitously close to the hyperthermic threshold indicating a significant thermal challenge during the TT. It seems unlikely that either factor (i.e., HR or CT) alone impacted overall performance, yet it is important to highlight that the TT_{avg} showed a mean negative slope (-0.01 W/sec) over the course of the TT indicating that fatigue was certainly a limiting factor. Moreover, from a pacing perspective, most participants seem to have used 'variable' pacing most likely to account for fluctuations in course geography, temperature and environmental conditions.[1,40,41] The variability in TT strategies, as gauged by W′bal, W′end, ΔW′ and TTP/CP, were intriguing, especially given that the second and third place finishers were within 3% (134-sec) and 4% (155-sec) of first place respectively. Including speed surges where power exceeds CP would have implications for metabolic stress and perceptions of effort/fatigue unless this were to be coupled with a sufficiently high aerobic capacity such that when power descends below CP, recovery can be sufficiently rapid.[42] The rapidity of W′-repletion is again governed by the magnitude of W′ where a larger W′ would be coupled with longer repletion kinetics. Thomas et al[4] showed that fatigue during a 40-km TT is likely to be mediated by central factors (e.g., reduced voluntary muscle activation, motivation) to the extent that psychological rather than physiological factors likely limit performance. Although the latter was not directly evaluated in the present study, such an explanation would seem plausible given that (i) the mean power output throughout the TT was sufficient below CP to not be excessively taxing from a metabolic perspective, (ii) there was substantial variability in pacing strategy, and (iii) the ability to sustain a high overall mean power or to have a higher CP was less predictive of performance.

Although there were several strengths associated with the present study it is also appropriate to highlight several limitations. Firstly, the current study had a limited sample size (n=9), which would imply that the generalizability of the findings should be interpreted with some caution. Secondly, all participants were male, therefore, the study should be replicated with the addition of a female cohort. Finally, the effects of environmental conditions (e.g., temperature, temperature perception, rolling friction etc.) on parameter estimations should be explored in greater detail.

This study was the first to examine whether a 3MT conducted in a laboratory- and field setting would produce similar peak power, CP, and W′ profiles. While the laboratory and field 3MTs produced 'comparable' CP and W′ profiles, P_{max} was substantially different between conditions. Given such a finding, it is important to be cognizant of the bias and LoA between testing conditions as this may preclude interchangeability of parameters. The latter is highlighted by the different parameter associations with $Time_{TT}$. For example, $Time_{TT}$ is associated with W'_{end}, W', and CP from the laboratory, and with W'_{end}, W' and P_{max} from the field to different extents.

The present study showed that mean TT power outputs tend to occur at 59-65% of CP, and finishing times appear to be informed by CP, W' and P_{max} with relatively high degrees of accuracy. It is, however, important to note that such a finding should be replicated on a larger sample to verify the associations. Pacing strategies were also highly variable between participants whereby some would employ power surges such that W' would be expended to appreciable extents whereas others would not exceed their CP and therefore never diminish their W' reserves.

Finally, despite finding that the mean power output of the TT was considerably below CP which would minimize metabolic perturbations, the mean HR and CT showed high levels of effort. It is likely that psychological factors rather than physiological factors may have accounted for aspects related to fatigu during the TT although future research should include specific psychological measures to account for distinctive afferent perceptions of effort.

Acknowledgements

The authors thank Dr. Susan Hart and Dr. Steve Bui, Health & Human Performance, Utah Tech University, for granting access to the Exercise Physiology Lab. The authors have no conflicts of interest. A Graduate Research Support Fund grant from Rocky Mountain University of Health Professions supported this work.

References

1. Foster C, de Koning JJ, Hettinga FJ, et al. Competition between desired competitive result, tolerable homeostatic disturbance, and psychophysiological interpretation determines pacing strategy. *Int J Sports Physiol Perform*. 2023:1-12. https://doi.org/10.1123/ijspp.2022-0171

2. Jacobs RA, Rasmussen P, Siebenmann C, et al. Determinants of time trial performance and maximal incremental exercise in highly trained endurance athletes. *J Appl Physiol*. 2011;111(5):1422-1430. https://doi.org/10.1152/japplphysiol.00625.2011

3. Santalla A, Earnest CP, Marroyo JA, Lucia A. The Tour de France: An updated physiological review. *Int J Sports Physiol Perform*. 2012;7(3):200-209. https://doi.org/10.1123/ijspp.7.3.200

4. Thomas K, Goodall S, Stone M, Howatson G, Gibson ASC, Ansley L. Central and peripheral fatigue in male cyclists after 4-, 20-, and 40-km time trials. *Med Sci Sports Exerc*. 2015;47(3):537-546. https://doi.org/10.1249/MSS.0000000000000448

5. Burnley M, Jones AM. Oxygen uptake kinetics as a determinant of sports performance. *Eur J Sport Sci*. 2007;7(2):63-79. https://doi.org/10.1080/17461390701456148

6. Nimmerichter A, Prinz B, Gumpenberger M, Heider S, Wirth K. Field-derived power–duration variables to predict cycling time-trial performance. *Int J Sports Physiol Perform*. 2020;15(8):1095-1102. https://doi.org/10.1123/ijspp.2019-0621

7. Jones AM, Vanhatalo A, Burnley M, Morton RH, Poole DC. Critical power: Implications for determination of VO2max and exercise tolerance. *Med Sci Sports Exerc*. 2010;42(10):1876-1890. https://doi.org/10.1249/MSS.0b013e3181d9cf7f

8. Muniz-Pumares D, Karsten B, Triska C, Glaister M. Methodological approaches and related challenges associated with the determination of critical power and curvature constant. *J Strength Cond Res*. 2019;32(2):584-596. https://doi.org/10.1519/JSC.0000000000002977

9. Sanders D, van Erp T, de Koning JJ. Intensity and load characteristics of professional road cycling: Differences between men's and women's races. *Int J Sports Physiol Perform*. 2019;14(3):296-302. https://doi.org/10.1123/ijspp.2018-0190

10. Chorley A, Lamb KL. The application of critical power, the work capacity above critical power (W′), and its reconstitution: A narrative review of current evidence and implications for cycling training prescription. *Sports*. 2020;8(9):123. https://doi.org/10.3390/sports8090123

11. Leo P, Spragg J, Podlogar T, Lawley JS, Mujika I. Power profiling and the power-duration relationship in cycling: a narrative review. *Eur J Appl Physiol*. 2022;122:301-316. https://doi.org/10.1007/s00421-021-04833-y

12. Dekerle J, Vanhatalo A, Burnley M. Determination of critical power from a single test. *Sci Sports*. 2008;23(5):231-238. https://doi.org/10.1016/j.scispo.2007.06.015

13. Karsten B, Jobson S, Hopker J, Jimenez A, Beedie C. High agreement between laboratory and field estimates of critical power in cycling. *Int J Sports Med*. 2013;35(04):298-303. https://doi.org/10.1055/s-0033-1349844

14. Cheshire WP. Thermoregulatory disorders and illness related to heat and cold stress. *Auton Neurosci*. 2016;196:91-104. https://doi.org/10.1016/j.autneu.2016.01.001

15. Otani H, Kaya M, Goto H, Tamaki A. Rising vs. falling phases of core temperature on endurance exercise capacity in the heat. *Eur J Appl Physiol*. 2020;120(2):481-491. https://doi.org/10.1007/s00421-019-04292-6

16. González-Alonso J, Teller C, Andersen SL, Jensen FB, Hyldig T, Nielsen B. Influence of body temperature on the development of fatigue during prolonged exercise in the heat. *J Appl Physiol*. 1999;86(3):1032-1039. https://doi.org/10.1152/jappl.1999.86.3.1032

17. Naito T, Saito T, Morinaga H, Eda N, Takai Y. Elevated core temperature in addition to mental fatigue impairs aerobic exercise capacity in highly trained athletes in the heat. *J Physiol Anthropol*. 2024;43(1):30. https://doi.org/10.1186/s40101-024-00377-0

18. Navalta JW, Stone WJ. Ethical issues relating to scientific discovery in exercise science. *Int J Exerc Sci*. 2019;12(1):1-8. https://doi.org/10.70252/EYCD6235

19. Clark IE, Gartner HE, Williams JL, Pettitt RW. Validity of the 3-Minute All-Out Exercise Test on the CompuTrainer. *J Strength Cond Res*. 2016;30(3):825-829. https://doi.org/10.1519/JSC.0000000000001169

20. Burnley M, Doust JH, Vanhatalo A. A 3-min all-out test to determine peak oxygen uptake and the maximal steady state. *Med Sci Sports Exerc*. 2006;38(11):1995-2003. https://doi.org/10.1249/01.mss.0000232024.06114.a6

21. Earnest CP, Wharton RP, Church TS, Lucia A. Reliability of the Lode Excalibur Sport Ergometer and applicability to Computrainer electromagnetically braked cycling training device. *J Strength Cond Res*. 2005;19(2):344. https://doi.org/10.1519/R-15714.1

22. Clark IE, Murray SR, Pettitt RW. Alternative procedures for the three-minute all-out exercise test. *J Strength Cond Res*. 2013;27(8):2104-2112. https://doi.org/10.1519/JSC.0b013e3182785041

23. Takeshima K, Onitsuka S, Xinyan Z, Hasegawa H. Effect of the timing of ice slurry ingestion for precooling on endurance exercise capacity in a warm environment. *J Therm Biol*. 2017;65:26-31. https://doi.org/10.1016/j.jtherbio.2017.01.010

24. Yeo Z, Fan PW, Nio AQ, Byrne C, Lee JK. Ice slurry on outdoor running performance in heat. *Int J Sports Med*. 2012;33(11):859-866. https://doi.org/10.1055/s-0032-1304643

25. Skiba PF, Clarke DC. The W′ Balance Model: Mathematical and methodological considerations. *Int J Sports Physiol Perform*. 2021;16(11):1561-1572. https://doi.org/10.1123/ijspp.2021-0205

26. Schober P, Schwarte LA. Correlation coefficients: Appropriate use and interpretation. *Anesth Analg*. 2018;126(5):1763-1768. https://doi.org/10.1213/ANE.0000000000002864

27. Patil I. Visualizations with statistical details: The "ggstatsplot" approach. *JOSS*. 2021;6(61):3167. https://doi.org/10.21105/joss.03167

28. Min SH, Zhou J. smplot: An R package for easy and elegant data visualization. *Front Genet*. 2021;12:802894. https://doi.org/10.3389/fgene.2021.802894

29. Bertucci W, Grappe F, Groslambert A. Laboratory versus outdoor cycling conditions: Differences in pedaling biomechanics. *J Appl Biomech*. 2007;23(2):87-92. https://doi.org/10.1123/jab.23.2.87

30. Bertucci WM, Betik AC, Duc S, Grappe F. Gross efficiency and cycling economy are higher in the field as compared with on an Axiom stationary ergometer. *J Appl Biomech*. 2012;28(6):636-644. https://doi.org/10.1123/jab.28.6.636

31. Karsten B, Jobson SA, Hopker J, Stevens L, Beedie C. Validity and reliability of critical power field testing. *Eur J Appl Physiol*. 2015;115(1):197-204. https://doi.org/10.1007/s00421-014-3001-z

32. Jones AM, Burnley M, Black MI, Poole DC, Vanhatalo A. The maximal metabolic steady state: redefining the ' gold standard .' *Physiol Rep*. 2019;7(10):1-16. https://doi.org/10.14814/phy2.14098

33. Burnley M, Jones AM. Power – duration relationship : Physiology , fatigue , and the limits of human performance. *Eur J Sport Sci*. 2016;18(1):1-12. https://doi.org/10.1080/17461391.2016.1249524

34. Craig JC, Vanhatalo A, Burnley M, Jones AM, Poole DC. Critical power: Possibly the most important fatigue threshold in exercise physiology. In: Zoladz A, ed. *Muscle and Exercise Physiology*. Elsevier Inc.; 2019:159-182. https://doi.org/10.1016/B978-0-12-814593-7.00008-6

35. Black MI, Durant J, Jones AM, Vanhatalo A. Critical power derived from a 3-min all-out test predicts 16.1-km road time-trial performance. *Eur J Sport Sci*. 2014;14(3):217-223. https://doi.org/10.1080/17461391.2013.810306

36. Hawley JA, Noakes TD. Peak power output predicts maximal oxygen uptake and performance time in trained cyclists. *Europ J Appl Physiol*. 1992;65(1):79-83. https://doi.org/10.1007/BF01466278

37. Faria EW, Parker DL, Faria IE. The science of cycling: Physiology and training - Part 1. *Sports Med*. 2005;35(4):285-312. https://doi.org/10.2165/00007256-200535040-00002

38. Collins J, Leach O, Dorff A, et al. Critical power and work-prime account for variability in endurance training adaptations not captured by VO2max. *J Appl Physiol*. 2022;133(4):986-1000. https://doi.org/10.1152/japplphysiol.00344.2022

39. American College of Sports Medicine. *ACSM's Guidelines for Exercise Testing and Prescription*. (Ehrman JK, ed.). Wolters Kluwer; 2018.

40. Abbiss CR, Laursen PB. Describing and understanding pacing strategies during athletic competition. *Sports Med*. 2008;38(3):239-252. https://doi.org/10.2165/00007256-200838030-00004

41. Castronovo AM, Conforto S, Schmid M, Bibbo D, D'Alessio T. How to assess performance in cycling: the multivariate nature of influencing factors and related indicators. *Front Physiol*. 2013;4. https://doi.org/10.3389/fphys.2013.00116

42. Brooks GA, Arevalo JA, Osmond AD, Leija RG, Curl CC, Tovar AP. Lactate in contemporary biology: A phoenix risen. *J Physiol*. 2021;0:1-23. https://doi.org/10.1113/JP280955