International Journal of Sports Physiology and Performance, (Ahead of Print)
https://doi.org/10.1123/ijspp.2023-0276

First Published Online: Oct. 6, 2023

Human Kinetics
ORIGINAL INVESTIGATION

Remote Determination of Critical Speed and Critical Power in Recreational Runners

Ben Hunter,[1,2] Adam Ledger,[1] and Daniel Muniz-Pumares[1]
[1]School of Life and Medical Sciences, University of Hertfordshire, Hatfield, United Kingdom; [2]School of Human Science, London Metropolitan University, London, United Kingdom

***Purpose**:* This study aimed to compare estimations of critical speed (CS) and work completed above CS (D′), and their analogies for running power (critical power [CP] and W′), derived from raw data obtained from habitual training (HAB) and intentional maximal efforts in the form of time trials (TTs) and 3-minute all-out tests (3MTs) in recreational runners. The test–retest reliability of the 3MT was further analyzed. ***Methods**:* Twenty-three recreational runners (4 female) used a foot pod to record speed, altitude, and power output for 8 consecutive weeks. CS and D′, and CP and W′, were calculated from the best 3-, 7-, and 12-minute segments recorded in the first 6 weeks of their HAB and in random order in weeks 7 and 8 from 3 TTs (3, 7, and 12 min) and three 3MTs (to assess test–retest reliability). ***Results**:* There was no difference between estimations of CS or CP derived from HAB, TT, and 3MT (3.44 [0.63], 3.42 [0.53], and 3.76 [0.57] $m \cdot s^{-1}$ and 281 [41], 290 [45], and 305 [54] W, respectively), and strong agreement between HAB and TT for CS ($r = .669$) and CP ($r = .916$). Limited agreement existed between estimates of D′/W′. Moderate reliability of D′/W′ was demonstrated between the first and second 3MTs, whereas excellent reliability was demonstrated for CS/CP. ***Conclusion**:* These data suggest that estimations of CS/CP can be derived remotely, from either HAB, TT, or 3MT, although the lower agreement between D′/W′ warrants caution when using these measures interchangeably.

***Keywords**:* running, intensity domains, testing

The relationship between the intensity of exercise (eg, running speed or power output) and tolerable duration in the severe exercise domain can be characterized by a hyperbolic function[1,2] and therefore defined by its asymptote and curvature constant. The asymptote, critical speed, or critical power (CS/CP, respectively) is generally accepted to denote the transition from heavy to severe exercise domain.[3] Therefore, although CS/CP appears to represent a metabolic rate and not necessarily a unique external mechanical output (power or speed),[4] exercising at intensities below the associated CS/CP, in the heavy domain of exercise, is typically characterized by a metabolic steady state, as denoted as by the steady state observed in oxygen consumption ($\dot{V}O_2$).[5] In contrast, exercise at intensities exceeding CS/CP, in the severe exercise domain, results in an inexorable increase in $\dot{V}O_2$ until the maximum $\dot{V}O_2$ is attained, and task failure ensues soon afterward, as well as concomitant increases in inorganic phosphate and hydrogen ions and decreases in phosphocreatine.[6] The curvature constant of the speed–duration relationship and power–duration relationship (D′ and W′, respectively) represents the upper limit of exercise capacity above CS/CP.

The speed–duration relationship is typically derived from 3 to 5 discrete maximal efforts in the severe exercise domain, each lasting between 2 and 15 minutes.[1] Maximal efforts are performed either as constant-intensity exercise sustained until task failure or as performance tests where a given task is completed in the shortest possible time in the possible time or at the highest intensity (ie, time trials [TT]). Further methodological considerations known to affect estimates of CS/CP and D′/W′ such as the number, duration, or rest between maximal efforts are discussed elsewhere.[1] Crucially, however, due to the time- and effort-intense nature of this approach, alternative protocols have been proposed to determine CS and D′. For instance, the speed–duration relationship has been constructed using raw training data, derived from the best performances recorded for a discrete set of distances over several weeks. However, it is worth highlighting the determination of CS/CP requires maximal efforts performed in the severe domain; therefore, this assumes that best efforts observed during training represent a maximal effort.[7,8] Using raw training data is; therefore, a promising avenue to calculate CS/CP and D′/W′ allowing remote data collection from athletes, but no study so far has compared estimations CS and D′ derived from habitual training (HAB) data and intentional maximal efforts (TT). Moreover, footworn accelerometers (eg, Stryd Power Meter, Stryd Inc) can be used in remote settings to derive running CP and W′.[9]

An alternative approach to determine the parameters of the intensity–duration relationship has been derived from the finite, constant work capacity above CS/CP, denoted as D′ or W′. In the first few seconds of an all-out effort, the intensity of exercise greatly exceeds that associated with the CS/CP and therefore D′/W′ rapidly depletes. If the all-out effort is sufficiently long, D′/W′ will continue to decrease until it is exhausted, and at that point, the intensity of exercise corresponds to CS/CP. In cycling, Vanhatalo et al[10] determined that a 3-minute all-out (3MT) effort was sufficient to fully exhaust W′, as the end-test power output corresponded to CP and the work completed above the end-test power corresponded to W′. Nonetheless, the 3MT requires an all-out effort sustained for 3 minutes, making it challenging to execute correctly and very unpleasant for athletes. With the advancements in wearable technology,[9] new avenues emerge for data collection. For example, Maunder et al[11] demonstrated that the 3MT can be performed remotely, producing results that displayed similar levels of day-to-day variation to what is typically observed for similar

Muniz-Pumares https://orcid.org/0000-0002-6748-9870
Hunter (b.hunter@londonmet.ac.uk) is corresponding author, https://orcid.org/0000-0003-4190-4543

markers of aerobic function (ie, coefficient of variation: 4.5%) but overestimated the maximal metabolic steady state. Furthermore, the 3MT has been adapted to running[12,13] and has been shown to be appropriate for prescription of training sessions and sensitive to training adaptations.[14] However, whether the running 3MT can be performed remotely and unsupervised remains to be determined.

The aim of the current study was, therefore, to compare different approaches to calculate CS and D′, and CP and W′ remotely. Specifically, herein we (1) compared estimations of CS and D′, and CP and W′ derived from HAB data, intentional maximal efforts (TT), and the 3MT; and (2) assessed the test–retest reliability of CS and D′, and CP and W′ determined from the 3MT performed remotely.

Methods

Participants

Twenty-three runners (4 females, group mean [SD] age: 45.3 [7.0] y; stature: 1.78 [0.07] m; mass: 83.3 [22.6] kg) volunteered to participate in the study and completed the protocol. The level of performance of the participants in the study fell under tiers 1 and 2 based on the framework outlined by McKay et al.[15] Following ethical approval by the Health, Science, Engineering and Technology Ethics Committee at the University of Hertfordshire (LMS/PGR/UH/04280), participants were recruited online, through advertising in social media. All methods conformed to the Code of Ethics of the World Medical Association (Declaration of Helsinki).

Study Design

The data collection for this study was carried out remotely over the course of 8 weeks. Participants were asked to record training data using their own Stryd Power Meter (Stryd Inc) and uploaded to a training platform (TrainingPeaks). During the first 6 weeks of the study (weeks 1–6), participants were instructed to continue with their HAB, with no further advice or request provided. During weeks 7 and 8, participants were asked to complete 3 TTs of 3, 7, and 12 minutes (TT) within a week and three 3MT also within a week, but the order of these tests (ie, TT and 3MTs) was randomized (Figure 1).

TTs and All-Out Efforts

All TT and 3MT tests were performed remotely and unsupervised, but participants were requested to follow the following instructions: (1) perform a maximal effort, and only proceed when fully committed to the test and after an appropriate (self-directed) warm-up; (2) allow at least 24 hours of rest between efforts, but 48 hours were recommended; (3) refrain from caffeine for 1 hour before each test; (4) refrain from performing further training during testing weeks (ie, weeks 7 and 8 of the study); and (5) complete all test on the same route, with minimal changes in elevation and sharp corners. The order of the 3 TT was, for all participants, 7-minute TT, 3-minute TT, and 12 minutes. Participants were encouraged to achieve the highest possible distance during each TT. During the 3MT tests, participants were encouraged to avoid any pacing, go flat out, and run as fast as possible, from the beginning of the test, and at all times during the test.

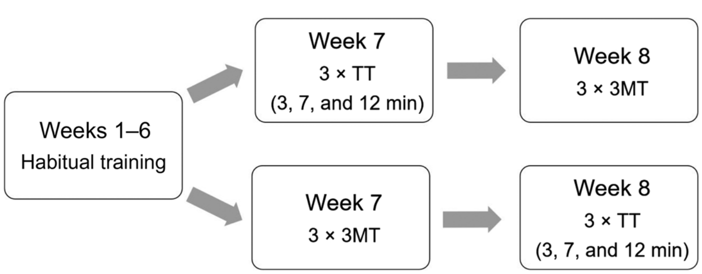

Figure 1 — Schematic overview of the study design. 3MT indicates 3-minute all-out test; TT, time trial.

Data Collection

Running speed and power output were determined with a foot pod power meter (Stryd Inc). Briefly, the foot pod attaches to the shoe at the midfoot, weighing 9.1 g. Based on a 6-axis inertial motion sensor (3-axis gyroscope and 3-axis accelerometer), this device provides 12 metrics to quantify performance: speed, distance, elevation, running power, form power, cadence, ground contact time, vertical oscillation, and leg stiffness. Previous studies have evidenced good reliability for spatiotemporal running characteristics[16] and power output.[17] Participants were encouraged to use Stryd in all training sessions and were reminded to regularly sync training data to the training platform. Data from all participants were downloaded from TrainingPeaks as .fit files and subsequently exported as .csv files using publicly available software (Golden Cheetah, version 3.4).

Data Analysis

All data analyses were carried out using MATLAB (2020b, MathWorks). Once training files were exported data, including speed, running power, and altitude were extracted for each participant. First, speed was grade-adjusted using methods that have been previously described.[18] In brief, this accounts for the dissociation between measured speed and metabolic intensity observed during uphill and downhill running. The 3-, 7-, and 12-minute segments with the highest average speed and power observed in the 6-week HAB phase were extracted to estimate CS (CS_{HAB}) and D′ (D'_{HAB}), and their analogous for running power CP (CP_{HAB}) and W′ (W'_{HAB}), respectively. Similarly, the average speed and power observed during the 3, 7, and 12 minutes of the TTs were used to estimate CS (CS_{TT}) and D′ (D'_{TT}), and CP (CP_{TT}) and W′ (W'_{TT}), respectively.

For both HAB and TT data, CS, CP, D′, and W′ were estimated using three 2-parameter models: (1) the hyperbolic model, (2) the linear distance/power-time model, and (3) the linear inverse-of-time model, as described previously.[19–21] Whichever model resulted in the smallest error of estimate for CS, CP, D′, and W′ was used to give the "best-individual-fit" estimate.[19–21] For each of the 3MTs, CS (CS_{3MT}) and CP (CP_{3MT}) were estimated using the average speed or power during the last 30 seconds of each test.[12,22] The D′ (D'_{3MT}) and W′ (W'_{3MT}) were calculated as the distance covered (in meters) or work completed (in kilojoules) above the CS and CP, respectively. The quality of the 3MT was assessed using the criteria outlined by Muniz-Pumares et al.[1]

Statistical Analysis

First, a one-way analysis of variance, limits of agreement, and within-subject coefficient of variation (CoV) were used to test for agreement between estimates of CS, CP, D′, and W′ in HAB, TT, and 3MT. Where significant main effects were detected, a Bonferroni correction was applied. Statistical significance was accepted at $P < .05$. The first of the 3MT was deemed as familiarization and discarded from further data analysis. Test–retest reliability for the remaining two 3MT was calculated using a 2-way mixed intraclass correlation coefficient with values below .5 indicating poor reliability, between .5 and .75 moderate reliability, between .75 and .9 good reliability, and values above .9 indicated excellent reliability.[23] To test for the strength of the relationship between estimates of CS, CP, D′, and W′ derived from HAB, TT, and 3MT, Pearson product–moment correlations were calculated. Results are reported as mean (SD).

Results

During the HAB phase of the study, participants ran a mean weekly distance of 48.62 (23.38) km, over 5.2 (5.6) weekly sessions, with a mean duration of 51 (12) minutes per session. The best-individual-fit estimates for the speed–duration relationship were obtained from the hyperbolic model in 21 participants with the linear distance-time model and the linear inverse-of-time model accounting for 1 participant each. The power–duration relationship followed a similar pattern, with the hyperbolic model producing the best-individual-fit in 19 participants, with the linear distance-time model and the linear inverse-of-time model accounting for 3 participants and 1 participant, respectively. The parameters of the power- and speed–duration relationship estimated from all models for HAB, TT, and 3MT are given in Supplementary Material (available online).

Habitual Training, Time Trials, and 3-Minute All-Out test

The parameters of the power–duration and speed–duration relationship estimated from HAB, TT, and 3MT are displayed in Table 1. There was no difference in the estimations of either CS and CP using different testing protocols ($P = .222$ and $P = .060$, for CS and CP, respectively). These were accompanied by low CoV and small bias values (Table 2). Moreover, there were no differences between estimations of D′ derived from HAB, TT, and 3MT ($P = .053$). However, large CoV and bias were noted for estimates of D′ using different methods. There was an overall mean difference in W′ ($P = .003$) estimated from these protocols. Bonferroni post hoc tests demonstrated a significantly lower W′ in the 3MT condition when compared with TT ($P = .013$). No further significant differences were noted between testing methods.

Table 1 Power–Duration and Speed–Duration Parameters Derived From HAB Data, TT, and 2 Identical 3MTs

	HAB	TT	First 3MT	Second 3MT
CP, W	281 (41)	290 (44)	305 (53)	307 (52)
W′, kJ	7.16 (4.34)	9.79 (5.31)[a,b]	6.06 (3.02)[a]	5.67 (2.84)[b]
CS, $m \cdot s^{-1}$	3.44 (0.63)	3.42 (0.53)	3.76 (0.57)	3.77 (0.60)
D′, m	100 (78)	134 (81)	94 (50)	83 (41)

Abbreviations: CP, critical power; CS, critical speed; HAB, habitual training; 3MT, 3-minute all-out test; TT, time trial. Note: Data are given as mean (SD). Superscript letters denote a significant difference between pairs. For HAB and TT, parameters from best individual fit models are displayed.

The agreement between estimates derived from different testing methods was better for estimates of CS and CP than for estimates of D′ and W′ (Figures 2–5). For CS, the mean difference ranged from 0.02 $m \cdot s^{-1}$ (95% confidence interval [CI], –0.92 to 0.96 $m \cdot s^{-1}$) to –0.34 $m \cdot s^{-1}$ (95% CI, –1.19 to 0.50 $m \cdot s^{-1}$), and for CP ranged from –9 W (95% CI, –44 to 26 W) to –24 W (95% CI, –76 to 28 W). Similarly, estimates of CS and CP derived from different methods were significantly and strongly ($P \geq .644$) correlated, whereas estimates of D′ and W′ were not correlated (Table 2).

Three-Minute All-Out Tests

No 3MT resulted in 5% drop below end-test power or speed for >5 seconds. For the first and second 3MT, peak speed was 5.92 (1.09) $m \cdot s^{-1}$ and 5.69 (1.04) $m \cdot s^{-1}$ occurring at 12.5 (7.9) seconds and 14.4 (9.2) seconds, respectively. Peak power was 451 (101) W and 454 (106) W, occurring at 12.6 (3.6) seconds and 17.0 (10.8) seconds for the first and second 3MT, respectively. Rapid depletion of W′ (ie, >90% of W′ depletion within the first 90 s of the test) occurred in 5 participants (21.7%) in the first 3MT and 7 participants (30.4%) in the second 3MT. Similar results were shown for D′, with 5 (21.7%) participants depleting >90% of D′ in the first 3MT and 6 (26.1%) in the second 3MT. An example of 3MT conducted in accordance with the criteria of Muniz-Pumares et al[1] and an example 3MT that exhibits pacing are shown in Figure 6. Moderate reliability was demonstrated between first and second 3MT in both W′ ($r = .716$, $P < .001$) and D′ (.698, $P < .001$), and excellent reliability was demonstrated for CP (.965, $P < .001$) and CS (.940, $P < .001$).

Table 2 Agreement Between Methods of Power–Duration and Speed–Duration Relationship Parameters Derived From HAB Data, TT, and the First 3MT

		95% LoA			Pearson coefficient	
	Bias	LL	UL	CoV%	*r*	*P*
CP, W						
HAB-TT	–9	–44	26	4.8	.916**	<.001
HAB-3MT	–24	–76	28	8.5	.878**	<.001
TT-3MT	–15	–78	48	8.3	.802**	<.001
W′, kJ						
HAB-TT	–2.63	–15.35	10.09	57.3	.107	.626
HAB-3MT	1.10	–10.10	12.30	61.0	–.180	.411
TT-3MT	3.73	–9.30	16.76	66.9	–.214	.326
CS, $m \cdot s^{-1}$						
HAB-TT	0.02	–0.92	0.96	9.7	.669**	<.001
HAB-3MT	–0.32	–1.32	0.67	11.6	.644**	<.001
TT-3MT	–0.34	–1.19	0.50	10.7	.695**	<.001
D′, m						
HAB-TT	–33	–201	134	54.1	.481*	.047
HAB-3MT	7	–197	211	74.2	–.306	.155
TT-3MT	40	–167	247	68.8	–.267	.218

Abbreviations: CoV, coefficient of variation; CP, critical power; CS, critical speed; HAB, habitual training; LoA, limits of agreement; 3MT, 3-minute all-out test; TT, time trial; UL, upper limit; LL, lower limit. Note: Data are given as mean (SD). For HAB and TT, parameters from best individual fit models are displayed. For 3MT, parameters from the first trial are displayed.
*$P < .05$. **$P < .01$.

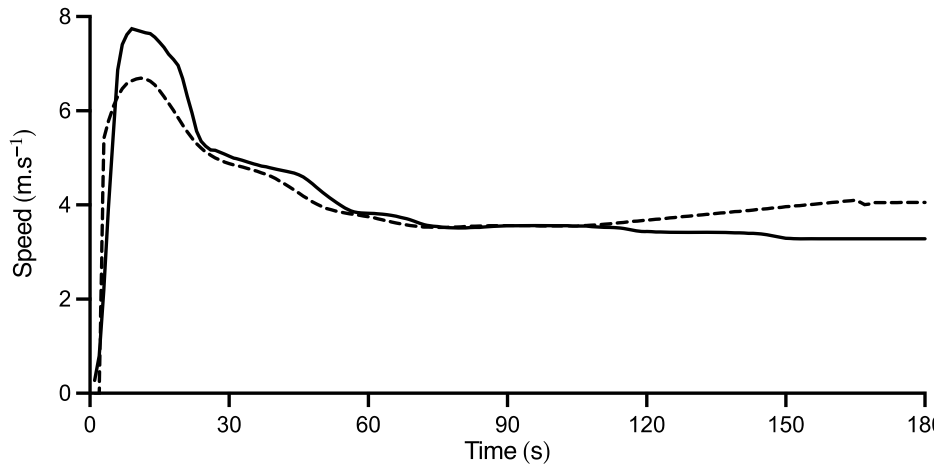

Figure 2 — Example 3MT speed–time profile conducted in accordance with the criteria of Muniz-Pumares et al[1] (solid black line) and an example 3MT speed–time profile that exhibits pacing strategies (dotted black line). Power–time profiles exhibited the same characteristics. 3MT indicates 3-minute all-out test.

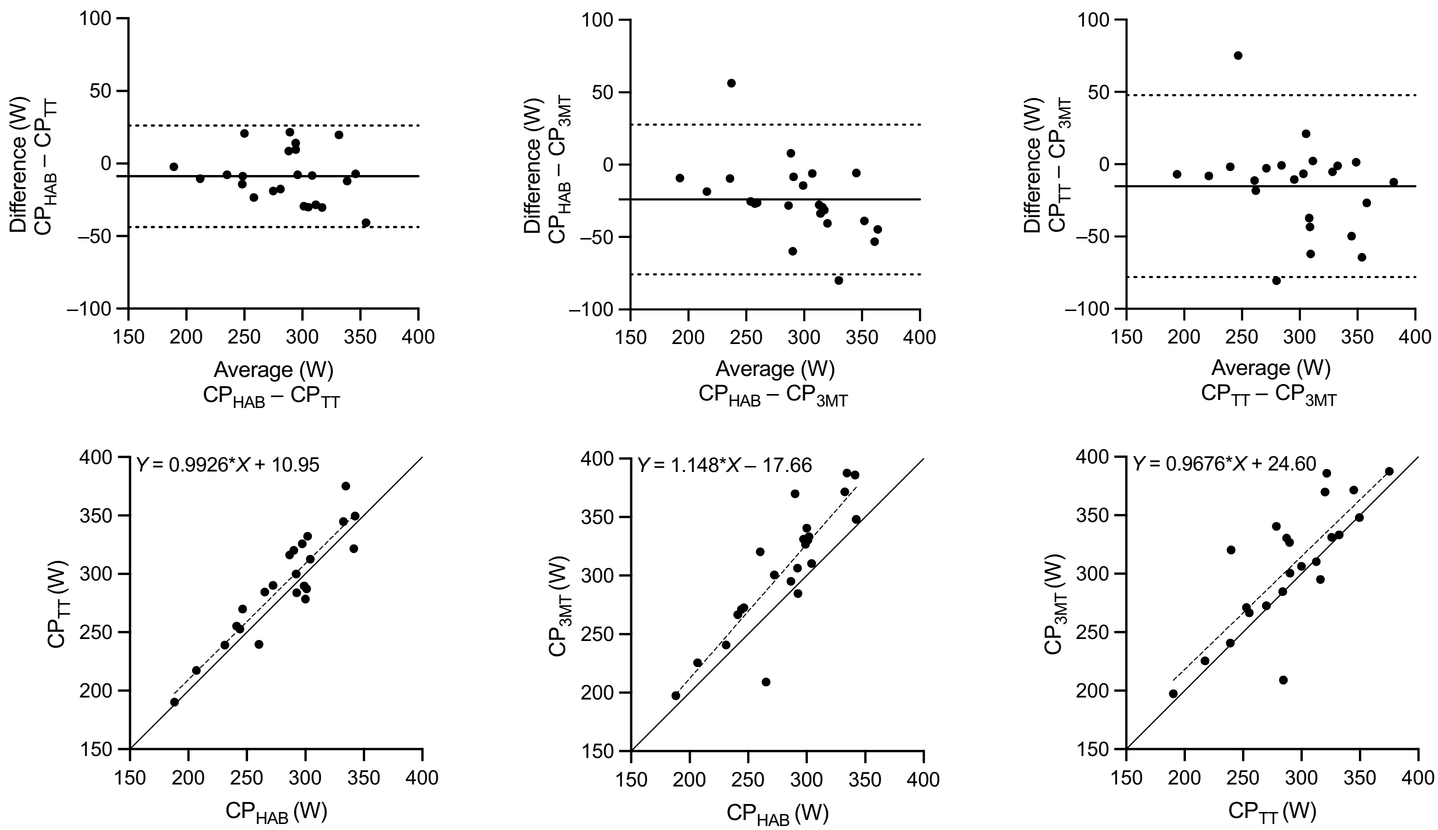

Figure 3 — Bland–Altman plots of the LoA (top panels) and the relationship (bottom panels) between different approaches to determine CP. In the top panels, the horizontal lines represent the mean difference between different approaches to determine CP, and the dashed line represents the 95% LoA. In the bottom panels, the solid lines represent $x = y$. Where a significant relationship was demonstrated, a dotted regression line and formula are given. CP indicates critical power; HAB, habitual training; LoA, limits of agreement; 3MT, 3-minute all-out test; TT, time trial.

Discussion

The primary aim of this study was to compare estimations of CS and D′ and running CP and W′ derived from 3 different remote protocols: HAB, TT, and 3MT. The secondary aim was to assess the reliability of these parameters derived from 3MT. The key findings are: (1) CS and CP estimated remotely through a range of methods (HAB, TT, and 3MT) showed good agreement (CoV of between 4.7% and 10.8%); (2) the correlation between estimates of CS and CP derived from HAB, TT, and 3MT was generally strong ($r \geq .664$ and $r \geq .802$, respectively); (3) the agreement between estimates of D′ and W′ from different testing methods,

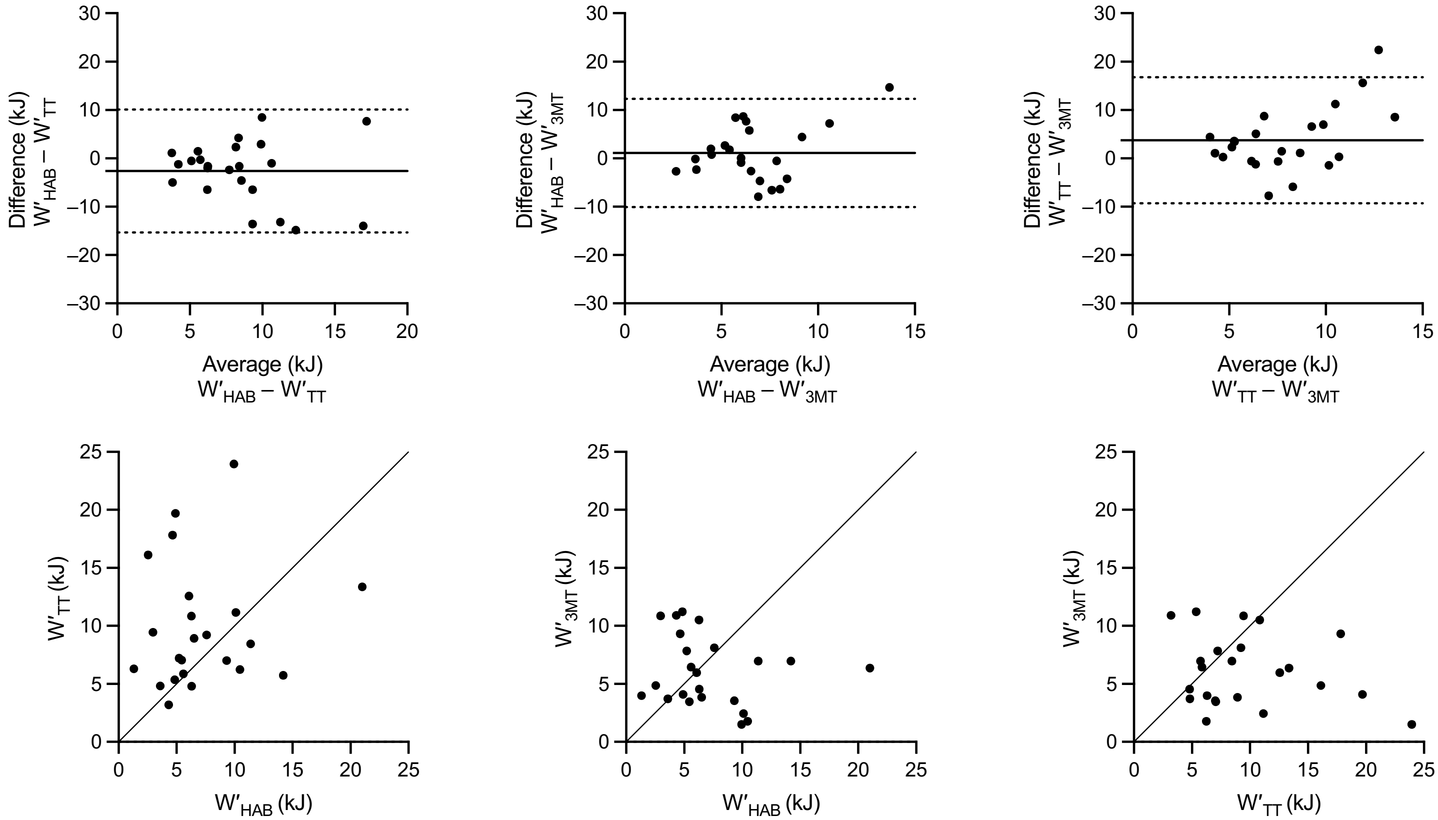

Figure 4 — Bland–Altman plots of the LoA (top panels) and the relationship (bottom panels) between different approaches to determine W′. In the top panels, the horizontal lines represent the mean difference between different approaches to determine W′, and the dashed line represents the 95% LoA. In the bottom panels, the solid lines represent $x = y$. Where a significant relationship was demonstrated, a dotted regression line and formula are given. W′ indicates work completed above critical power; HAB, habitual training; LoA, limits of agreement; 3MT, 3-minute all-out test; TT, time trial.

however, was lower; and (4) the 3MT exhibited fair test–retest reliability.

The estimations of CS and CP appear to be unaffected by the method of estimation (HAB, TT, or 3MT), as evidenced by the lack of significant difference, strong, positive correlation, and the low CoV between methods, as reported in Tables 1 and 2 and Figures 2 and 4. The results presented herein demonstrate a particularly strong agreement between estimations of CP and CS derived from HAB and TT methods, as evidenced by the positive correlation ($r = .916$ and $r = .669$), low mean difference (~0.02 $\text{m} \cdot \text{s}^{-1}$ and ~9 W) and low CoV (4.8% and 9.7% for CP and CS, respectively). Combined, these data suggest that CS/CP, an important marker of endurance performance and exercise tolerance,[24,25] can be estimated from HAB training data. This result supports the determination of CS from HAB to, for example, analyze training data, retrospectively, similar to what has been done from analysis of performance measures [7,8] and opens new avenues for research. For example, the determination of CS from HAB training offers the possibility to design remote interventions and monitor adaptations to training programs, by means of monitoring changes in CS over time.[14,26] Nonetheless, there are circumstances where the determination of CS requires higher precision, and the use of intentional maximal efforts (eg, TT in the field or laboratory testing) may be still preferred. In addition, it is also important to highlight that the transition between heavy and severe exercise domains occurs gradually and not suddenly.[27] Thus, we would advise athletes and practitioners to be consistent with the methodology adopted to estimate CS/CP and prescribe exercise outside the confidence limits when seeking to elicit heavy- or severe-domain-specific responses to exercise.

There were no differences between estimations of CS and CP from HAB and TT, and those derived from 3MT, as well as moderate to excellent reliability scores during the repeated 3MT tests. However, it is notable that the agreement between HAB, TT, and 3MT was generally poor and considerably lower than the agreement between HAB and TT. The 3MT requires an all-out effort sustained for 3 minutes. We emphasized to participants the importance of correctly executing the test, avoiding pacing, and to adhere to the guidelines set out in the literature.[1,22] However, despite these instructions, and concurrent with other literature,[11] some evidence of pacing was noted in some participants. For example, the rapid depletion of W′ or D′ only occurred in a maximum of 7 (30.4%) and 6 (26.1%) participants, respectively. These findings suggest that the 3MT may not represent a maximal effort in some participants, thus conserving some capacity above the CP/CS throughout the first 150 seconds of the test. This may explain the elevated end-test parameters (CS/CP) measured in the current study, as well as lower capacity above the CP/CS when compared with HAB and TT trials. This may tentatively be attributed to several factors. In the 3MT, participants are typically unaware of elapsed time as clocks and other time-displaying devices are removed,[10,22] but in the current study, elapsed time may have also been known to the participants during the trial. Furthermore, to the researchers' knowledge, no strong verbal or other encouragement was provided to the participants, resulting in more limited extrinsic motivation to complete the task. However, no 3MT resulted in a

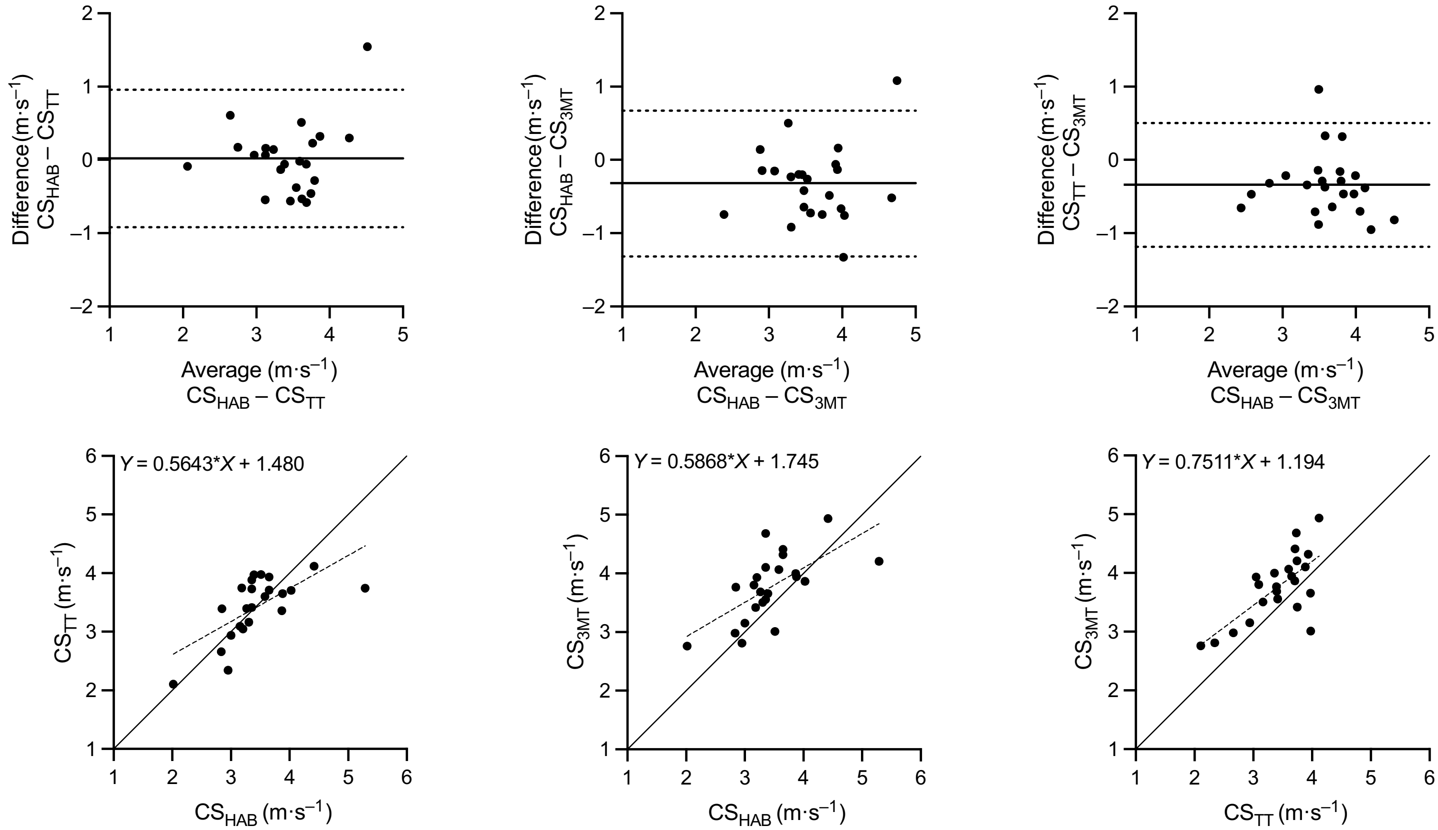

Figure 5 — Bland–Altman plots of the LoA (top panels) and the relationship (bottom panels) between different approaches to determine CS. In the top panels, the horizontal lines represent the mean difference between different approaches to determine CS, and the dashed line represents the 95% LoA. In the bottom panels, the solid lines represent $x = y$. Where a significant relationship was demonstrated, a dotted regression line and formula are given. CS indicates critical speed; HAB, habitual training; LoA, limits of agreement; 3MT, 3-minute all-out test; TT, time trial.

5% drop below end-test power or speed for >5 seconds, which means that replenishment of either D′ or W′ is unlikely to have occurred.

It could be postulated that the 3MT in running requires alterations to the traditional methods of executing a 3MT. Indeed, similar estimates of the heavy- and severe-domain transition have been shown when using an average speed of the final 20 seconds of the 3MT.[13] However, this approach also resulted in an underestimation of D′. To derive better estimates of CS/CP and D′/W′ from one test, it could be posited that a longer test is needed, similar to the 5 minutes required for knee-extension exercise.[28] Differences between the 3MT and TTs were similar to existing literature using the same parameters[29] and exhibited 5% to 10% difference in the estimation of the power/speed–duration relationship parameters when compared with HAB and TT. Despite the apparent overestimation of the heavy–severe boundary, the parameters of the 3MT generally exhibited a moderate test–retest reliability for W′ ($r = .716$, $P < .001$), D′ (.698, $P < .001$), and excellent reliability was demonstrated for CP (.965, $P < .001$) and CS (.940, $P < .001$). In this study, following criteria devised for the correct execution of a cycling 3MT,[1] as running-specific criteria are currently lacking, both tests exhibited similarities in mean time to peak speed and power, mean peak speed and power, number of participants to deplete >90% of W′ depletion within the first 90 seconds of the test, and drops below end-test values for >5 seconds. However, the reliability of the 3MT must be considered in the context of the plausible overestimation of CS/CP and the fact that 3MT has been shown to overestimate the maximal metabolic steady state previously.[11]

Estimates of D′ were not different between methods, while W'_{3MT} was lower than W'_{HAB} and W'_{TT}, and W'_{TT} was the higher than both W'_{TT} and W'_{HAB}. Furthermore, the agreement between these estimates and estimates for D'_{3MT} compared with other methods (D'_{HAB} and D'_{TT}) was generally weak (see Tables 1 and 2, Figure 5). Only D'_{HAB} and D'_{TT} showed a significant relationship, but the relationship was only moderate. Given the sensitivity of both D′ and W′ to changes in methodology, evidenced by the poor agreement shown in the current study, caution is warranted when determining D′/W′ remotely. Incongruency between measures of the curvature constant has been documented previously and is associated with a greater error when compared with CS/CP.[1]

Limitations and Methodological Considerations

Both TT and 3MT necessitate either a series of maximal efforts or one exhaustive trial, respectively. In laboratory settings, it is possible to assess whether efforts were maximal by checking whether maximal $\dot{V}O_2$ has been attained.[1] However, it has been shown that unintended, and uninstructed, efforts from training can be used to estimate CP and W′, which show a high level of agreement with laboratory-based measures.[30] Our results are also in agreement with Maunder et al,[11] who demonstrated the 3MT overestimates the maximal metabolic steady state, suggesting that

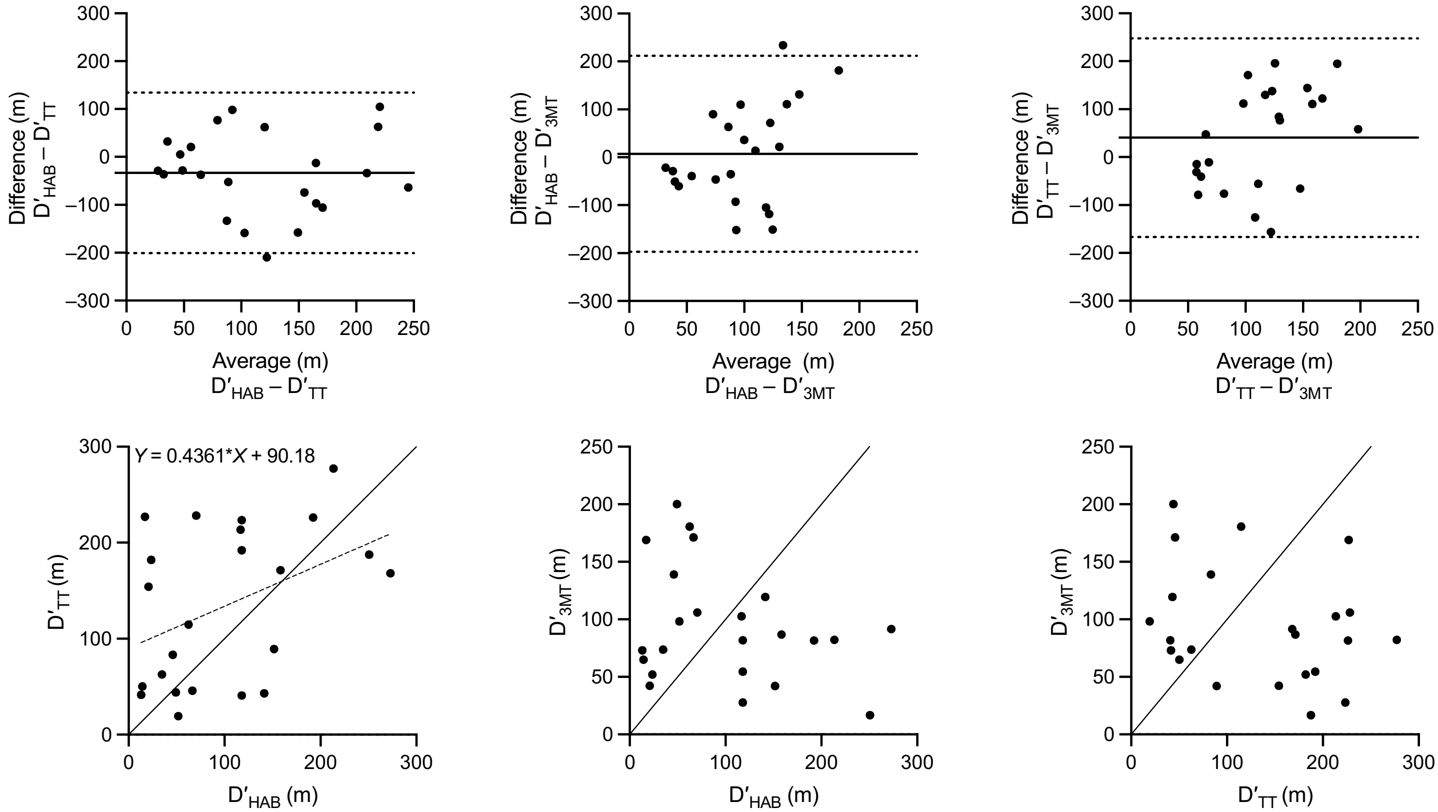

Figure 6 — Bland–Altman plots of the LoA (top panels) and the relationship (bottom panels) between different approaches to determine D′. In the top panels, the horizontal lines represent the mean difference between different approaches to determine D′, and the dashed line represents the 95% LoA. In the bottom panels, the solid lines represent $x = y$. Where a significant relationship was demonstrated, a dotted regression line and formula are given. D′ indicates distance covered above critical speed; HAB, habitual training; LoA, limits of agreement; 3MT, 3-minute all-out test; TT, time trial.

the TT or HAB parameters provide a more accurate estimation. It is worth considering that in the current study, HAB was collected from 6 weeks, so that the best performance for the discrete times corresponding to 3, 7, and 12 minutes was derived from ~30 training sessions. Previous studies have used 16 weeks of habitual training data to derive estimates of CS and D′.[7] It remains unclear what is the minimum period of HAB required to obtain reasonable estimates of CS/CP from HAB. The surfaces used to conduct TT and 3MT trials were uncontrolled and subject to changes in surface conditions, weather, and so forth. Nonetheless, participants were asked to run on flat route to account for this, and grade adjustment was applied to speed. Although previous studies have evidenced good reliability for spatiotemporal running characteristics[16] and power output[17] for the foot pod used in the current investigation, some error may occur when running style is changed.[31] Although this is unlikely to have occurred, differences in running style between trials may have affected the results.

Practical Applications

Alongside GPS-enabled wearables to track speed, power is also becoming a more widely used tool in training, through the use of accelerometers.[32] The ability to determine meaningful physiological parameters (ie, CS/CP and D′/W′) to control training intensity is paramount. This study has demonstrated that it is possible to determine the CS/CP from habitual training similar to the CS/CP determined by more time-consuming TTs. Such an approach permits 2 important applications: (1) to monitor progress using habitual training instead of, or complementary to, traditional testing methods, and (2) prescription of targeted training intensity by coaches or athletes. However, caution is warranted when trying to determine the curvature constant using this approach.

Conclusions

The determination of CS/CP and D′/W′ has attracted considerable interest due to its practical and physiological significance. Herein, we have shown that computing the best efforts recorded during 6 weeks of habitual training produces estimates of CS/CP comparable to those derived from intentional maximal efforts (ie, TT), as evidenced by the lack of significant difference and low mean difference. However, despite the good reliability of the 3MT, the agreement between estimates of CS/CP derived from habitual training and the 3-minute all-out test was poor, possibly indicative of poor pacing during these trials. Limited agreement was evidenced between estimates of D′/W′; thus, estimates of the curvature constant using different methods should not be used interchangeably.

References

1. 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;33(2):584–596. doi:10.1519/JSC.0000000000002977
2. Poole DC, Ward SA, Gardner GW, et al. Metabolic and respiratory profile of the upper limit for prolonged exercise in man. *Ergonomics*. 1988;31(9):1265–1279. doi:10.1080/00140138808966766
3. Jones AM, Burnley M, Black MI, Poole DC, Vanhatalo A. The maximal metabolic steady state: redefining the 'gold standard.' *Physiol Rep*. 2019;7(10):98. doi:10.14814/phy2.14098
4. Barker T, Poole DC, Noble ML, Barstow TJ, Barstow TJ. Human critical power-oxygen uptake relationship at different pedalling frequencies. *Exp Physiol*. 91:621–632. doi:10.1113/expphysiol.2005.032789
5. Nixon RJ, Kranen SH, Vanhatalo A, Jones AM. Steady-state VO_2 above MLSS: evidence that critical speed better represents maximal metabolic steady state in well-trained runners. *Eur J Appl Physiol*. 2021;121(11):3133–3144. doi:10.1007/s00421-021-04780-8
6. Jones AM, Wilkerson DP, DiMenna F, Fulford J, Poole DC. Muscle metabolic responses to exercise above and below the "critical power" assessed using 31P-MRS. *Am J Physiol Regul Integr Comp Physiol*. 2007;294(2):R585–R593. doi:10.1152/ajpregu.00731.2007
7. Smyth B, Muniz-Pumares D. Calculation of critical speed from raw training data in recreational marathon runners. *Med Sci Sports Exerc*. 2020;52(12):2637–2645. doi:10.1249/MSS.0000000000002412
8. Smyth B, Maunder E, Meyler S, Hunter B, Muniz-Pumares D. Decoupling of internal and external workload during a marathon: an analysis of durability in 82,303 recreational runners. *Sports Med*. 2022;52(9):2283–2295. doi:10.1007/s40279-022-01680-5
9. Ruiz-Alias SA, Olaya-Cuartero J, Ñancupil-Andrade AA, García-Pinillos F. 9/3-minute running critical power test: mechanical threshold location with respect to ventilatory thresholds and maximum oxygen uptake. *Int J Sports Physiol Perform*. 2022;17(7):1111–1118. doi:10.1123/IJSPP.2022-0069
10. Vanhatalo A, Doust JH, Burnley M. Determination of critical power using a 3-min all-out cycling test. *Med Sci Sports Exerc*. 2007;39(3):548–555. doi:10.1249/mss.0b013e31802dd3e6
11. Maunder E, Rothschild JA, Ramonas A, Delcourt M, Kilding AE. A three-minute all-out test performed in a remote setting does not provide a valid estimate of the maximum metabolic steady state. *Eur J Appl Physiol*. 2022;122(11):2385–2392. doi:10.1007/S00421-022-05020-3/FIGURES/4
12. Pettitt RW, Jamnick N, Clark IE. 3-Min all-out exercise test for running. *Int J Sports Med*. 2012;33(6):426–431. doi:10.1055/s-0031-1299749
13. Broxterman RM, Ade CJ, Poole DC, Harms CA, Barstow TJ. A single test for the determination of parameters of the speed-time relationship for running. *Respir Physiol Neurobiol*. 2013;185(2):380–385. doi:10.1016/j.resp.2012.08.024
14. Clark IE, West BM, Reynolds SK, Murray SR, Pettitt RW. Applying the critical velocity model for an off-season interval training program. *J Strength Cond Res*. 2013;27(12):3335–3341. doi:10.1519/JSC.0b013e31828f9d87
15. McKay AKA, Stellingwerff T, Smith ES, et al. Defining training and performance caliber: a participant classification framework. *Int J Sports Physiol Perform*. 2021;17(2):317–331. doi:10.1123/IJSPP.2021-0451
16. García-Pinillos F, Roche-Seruendo LE, Marcén-Cinca N, Marco-Contreras LA, Latorre-Román PA. Absolute reliability and concurrent validity of the stryd system for the assessment of running stride kinematics at different velocities. *J Strength Cond Res*. 2021;35(1):78–84. doi:10.1519/JSC.0000000000002595
17. Cartón-Llorente A, Roche-Seruendo LE, Jaén-Carrillo D, Marcen-Cinca N, García-Pinillos F. Absolute reliability and agreement between Stryd and RunScribe systems for the assessment of running power. *J Sports Eng Tech*. 2021;235(3):182–187. doi:10.1177/1754337120984644
18. Minetti AE, Moia C, Roi GS, Susta D, Ferretti G. Energy cost of walking and running at extreme uphill and downhill slopes. *J Appl Physiol*. 2002;93(3):1039–1046. doi:10.1152/JAPPLPHYSIOL.01177.2001/ASSET/IMAGES/LARGE/DG0921828006.JPEG
19. Black MI, Jones AM, Bailey SJ, Vanhatalo A. Self-pacing increases critical power and improves performance during severe-intensity exercise. *Appl Physiol Nutr Metab*. 2015;40(7):662–670. doi:10.1139/apnm-2014-0442
20. Black MI, Jones AM, Blackwell JR, et al. Muscle metabolic and neuromuscular determinants of fatigue during cycling in different exercise intensity domains. *J Appl Physiol*. 2017;122(3):446–459. doi:10.1152/japplphysiol.00942.2016
21. Hunter B, Greenhalgh A, Karsten B, Burnley M, Muniz-Pumares D. A non-linear analysis of running in the heavy and severe intensity domains. *Eur J Appl Physiol*. 2021;121(5):1297–1313. doi:10.1007/s00421-021-04615-6
22. 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. doi:10.1249/01.mss.0000232024.06114.a6
23. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15(2):155–163. doi:10.1016/J.JCM.2016.02.012
24. Poole DC, Burnley M, Vanhatalo A, Rossiter HB, Jones AM. Critical power: an important fatigue threshold in exercise physiology. *Med Sci Sports Exerc*. 2016;48(11):2320–2334. doi:10.1249/MSS.0000000000000939
25. Meyler S, Bottoms L, Wellsted D, Muniz-Pumares D. Variability in exercise tolerance and physiological responses to exercise prescribed relative to physiological thresholds and to maximum oxygen uptake. *Exp Physiol*. 2023;108(4):581–594. doi:10.1113/EP090878
26. Souza HLR, Bernardes BP, dos Prazeres EO, et al. Hoping for the best, prepared for the worst: can we perform remote data collection in sport sciences? *J Appl Physiol*. 2022;133(6):1430–1432. doi:10.1152/JAPPLPHYSIOL.00196.2022
27. Pethick J, Winter SL, Burnley M. Physiological evidence that the critical torque is a phase transition not a threshold. *Med Sci Sports Exerc*. 2020;52(11):2390–2401. doi:10.1249/mss.0000000000002389
28. Burnley M. Estimation of critical torque using intermittent isometric maximal voluntary contractions of the quadriceps in humans. *J Appl Physiol*. 2009;106(3):975–983. doi:10.1152/japplphysiol.91474.2008
29. Aguiar RAD, Salvador AF, Penteado R, Faraco HC, Pettitt RW, Caputo F. Reliability and validity of the 3-min all-out running test [Confiabilidade e validade do teste de 3 minutos máximo] [Fiabilidad y validez de la prueba de 3 minutos máximos]. *Revista Brasileira de Ciencias do Esporte*. 2018;40(3):288–294. doi:10.1016/j.rbce.2018.02.003
30. 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. doi:10.1007/s00421-014-3001-z
31. Baumgartner T, Held S, Klatt S, Donath L. Limitations of foot-worn sensors for assessing running power. *Sensors*. 2021;21(15):4952. doi:10.3390/S21154952
32. Jaén-Carrillo D, Roche-Seruendo LE, Cartón-Llorente A, Ramírez-Campillo R, García-Pinillos F. Mechanical power in endurance running: a scoping review on sensors for power output estimation during running. *Sensors*. 2020;20(22):6482. doi:10.3390/S20226482